

UNIVERSIDADE CATÓLICA DE BRASÍLIA

PROGRAMA DE PÓS-GRADUAÇÃO
STRICTU SENSU EM CIÊNCIAS GENÔMICAS E
BIOTECNOLOGIA

MESTRADO

Desenvolvimento e Mapeamento de
Marcadores Moleculares RGA em
Espécies Silvestres de *Arachis*

Ana Carolina Vilarinho Fernandes José
Orientador: David John Bertioli

BRASÍLIA

2006



Ana Carolina Vilarinho Fernandes José

Desenvolvimento e Mapeamento de Marcadores Moleculares RGA em Espécies Silvestres de *Arachis*

Dissertação apresentada ao programa de Pós-Graduação *Strictu Sensu* em Ciências Genômicas e Biotecnologia, da Universidade Católica de Brasília como requisito para obtenção do Título de Mestre em Ciências Genômicas e Biotecnologia.

Orientador: David John Bertioli

Brasília
2006

Termo de Aprovação

Dissertação defendida e aprovada como requisito parcial para obtenção do Título de Mestre em Ciências Genômicas e Biotecnologia, defendida e aprovada em 30 de Março de 2006 pela banca examinadora constituída por:

Dr. Rinaldo Wellerson Pereira

Dra. Soraya Leal-Bertioli

Dra. Gláucia Salles Cortopassi Buso

Dr. David John Bertioli

Brasília
UCB

*"As dificuldades, como as montanhas, aplainam-se
quando avançamos por elas".
Émile Zola (1840-1902)
Escritor francês*

Agradecimentos

À FUNARBE e Comunidade Européia pelo suporte financeiro que viabilizou o desenvolvimento desta Dissertação.

Ao Professor Dr. David John Bertioli pela orientação, paciência, ensinamentos, incentivo e confiança, assim como pela eterna batalha por equipamentos, materiais e resultados...

À Dra. Soraya Leal-Bertioli pela enorme orientação e ajuda essencial para o término do trabalho.

À Dra. Patrícia Messenberg pela imensa ajuda em experimentos e ótima orientação na no Laboratório LPPII na EMBRAPA-Recursos Genéticos e Biotecnologia.

Ao MSc. Márcio Moretzsohn pelo empréstimo de materiais, instalações e contribuição prestada à minha formação assim como pela cordialidade e imensa ajuda na inclusão dos marcadores moleculares aqui identificados em um mapa genético por ele construído.

Aos Drs. Gláucia Salles Cortopassi Buso e Rinaldo Wellerson Pereira que cordialmente aceitaram participar da banca examinadora desta Dissertação.

A todos do Laboratório LPPII pela amizade, ajuda, horas de descontração e fofocas...

Ao Laboratório de Ciências Biotecnológicas da Universidade Católica de Brasília pela disponibilidade, organização, materiais, equipamentos e agradável ambiente de trabalho.

A todos os professores da Universidade Católica de Brasília em especial aos professores Betânia Quirino, Danielle Cordeiro, Dario Grattapaglia, David Bertioli, Eduardo Leonardecz, Eliane Noronha, Francisco Aragão, Georgios Pappas, Maria Fátima Grossi, Octávio Franco, Ricardo Kruger, Rinaldo Pereira, Robert Miller, Rosane Collevatti, Ruy Caldas e Tatsuya Nagata pelas excelentes disciplinas ministradas e convenientes “puxadas de orelha”.

Aos queridos colegas de turma: Camila, Patrícia, Fred, Gustavo, Lílian, Alexandre, André, Jorge e aos mestres recém-formados Suzana, Sandra, Letícia, Alexandre e todos os alunos do curso que comigo compartilharam momentos de descontração e tensão nos sempre apertados prazos.

Aos funcionários do laboratório da Universidade Católica, Alessandra Reis, André, Willian, Idacuí, Márcia e Senadara pela prestativa ajuda quando necessária e amizade e ao secretário Fábio pela simpatia e “quebradas de galho” sempre concedidas.

Ao meu querido Alexandre pela compreensão, paciência, falta de roupa passada no armário, ajuda com o computador nas horas de “pane” pelo carinho, interesse, solidariedade e principalmente apoio nas minhas decisões.

Aos meus pais pela luta para que tudo sempre desse certo em minha carreira estudantil, apoio nas horas difíceis, compreensão nas minhas muitas horas de “chilique” e, por tudo...

Obrigada.

Sumário

Lista de Figuras e Tabelas.....	viii
Abreviaturas	x
Resumo	xi
Abstract.....	xii
1- Introdução	1
O gênero <i>Arachis</i>	1
Espécies Silvestres.....	4
Genes de Resistência	5
Marcadores Moleculares.....	7
Mapeamento Genético	10
2- Objetivos	12
Objetivos Gerais	12
Objetivos Específicos	12
3- Material e Métodos	13
Material vegetal e população de mapeamento.....	13
RFLP.....	14
RGA Display	16
Sequenciamento e processamento das amostras.....	19
Alinhamento das Seqüências e Árvore de Similaridade.....	21
4- Resultados	23
Extração de DNA e qualidade das amostras.....	23
Southern-blot	24
RGA-Display	29
Sequenciamento e análise das bandas polimórficas	38
Mapeamento genético dos marcadores.....	52

5- Discussão.....	57
Geração de marcadores através de RFLP	57
A técnica RGA-Display.....	58
Mapeamento Genético	62
Considerações Gerais	63
6- Conclusões	65
7- Referências Bibliográficas	67
ANEXOS	71
Anexo 1	73
Anexo 2	77
Anexo 3	87
Artigo.....	96
Artigo Publicado.....	119

Lista de Figuras, Tabelas e Gráficos

Figura 1 Ilustração de <i>Arachis</i>	1
Figura 2 Domínio NBS-LRR representado esquematicamente.....	7
Figura 3 Representação esquemática da técnica RGA-display.....	21
Figura 4 Gel de agarose contendo DNA genômico de <i>A. duranensis</i> , planta híbrida e <i>A. stenosperma</i>	24
Figura 5 gel de agarose contendo digestões das amostras de <i>A. duranensis</i> , da planta híbrida e de <i>A. stenosperma</i>	25
Figura 6 Auto-radiografia de Southern-blot com parentais.....	26
Figura 7 Auto-radiografia de Southern-blot contendo 15 plantas da população F ₂	29
Figura 8 Gel de Poliacrilamida corado com Nitrato de Prata	31
Figura 9 Gel de poliacrilamida 4% corado com Nitrato de Prata contendo amostras de RGA-Display.....	33
Figura 10 Gel de poliacrilamida 4% corado com Nitrato de Prata contendo amostras de RGA-display amplificadas com o <i>primer</i> P2-19.....	35
Figura 11 Gel de poliacrilamida 4% corado com Nitrato de Prata contendo amostras de RGA-display amplificadas com o <i>primer</i> P2-414.....	38
Figura 12 Gel de agarose 0,8% corado com brometo de etídio contendo bandas retiradas do gel de poliacrilamida e re-amplificadas por PCR.....	38
Figura 13 Processamento de seqüências utilizando a ferramenta Staden Package	39
Figura 14 Imagem de processamento do provável RGA obtida através do programa Motif-Scan.....	42
Figura 15 Imagem de alinhamento feito pelo programa Jalview com tipo de alinhamento ClustalW..... Erro! Indicador não definido.	
Figura 16 Árvore de similaridade baseada em um alinhamento ClustalW de seqüências de proteínas dos RGAs encontrados neste trabalho	47
Figura 17 Imagem Jalview com alinhamento das seqüências RGAs.....	48

Figura 18 Árvore de similaridade das 19 amostras RGAs obtidas nesse trabalho e 69 obtidas em Bertoli et al., 2003.....	49
Figura 19 Alinhamento ClustalW utilizando o programa JalView de seqüências de proteínas do tipo kinase.	50
Figura 20 Alinhamento das seqüências protéicas dos marcadores 14B e 14C homólogos ao gene de β -amilase.	53
Figura 21 Mapa Genético para <i>Arachis</i>	55
Tabela 1 RGAs isolados por Bertoli <i>et al.</i> (2003) e usados como sonda para a hibridização em Southern-blots neste trabalho.....	16
Tabela 2 Lista de adaptadores e <i>primers</i> usados em RGA-Display.....	17
Tabela 3 Seqüência dos <i>Primers</i> utilizados.....	18
Tabela 4 RGAs isolados por Bertoli et al. 2003 utilizados como sonda neste trabalho.....	27
Tabela 5 RGAs usados como sonda para genotipagem da população F ₂ e números de indivíduos correspondentes a cada experimento.....	28
Tabela 6 <i>Primers</i> relativos ao Domínio NBS e suas melhores temperaturas de ligação.....	32
Tabela 7 Combinações de <i>primers</i> utilizadas para genotipagem da população F ₂ , Número de bandas amplificadas, polimórficas, seqüenciadas e RGAs produzidas por cada combinação testada.....	37
Tabela 8 Bandas RGA-display seqüenciadas e suas principais homologias com seqüências conhecidas no GenBank.....	40
Tabela 9 Freqüência de bandas RGAs e outros genes encontrados em RGA-Display.....	41
Tabela 10 Marcadores com homologia a RGAs e comparação entre os Hits encontrados pelo BLASTX e Motif-Scan (Pfam).....	44
Gráfico 1 Representação das principais homologias encontradas no seqüenciamento de bandas polimórficas em RGA-Display.....	42

Abreviaturas

μl	Micro litro	NaCl	Cloreto de Sódio
μM	Micromolar	NaOH	Hidróxido de Sódio
³² P	Fósforo 32	NBS	<i>Nucleotide Binding Site</i>
AFLP	<i>Amplified Fragment Length Polymorphism</i>	°C	Graus Celsius
cm	Centímetros	P00	<i>Primer complementar ao adaptador Pst</i>
cM	Centimorgans	pB	Pares de Bases
CTAB	<i>Cetyltrimethylammonium Bromide</i>	PCR	<i>Polymerase Chain Reaction</i>
DNA	Ácido Desoxirribonucleico	pH	Potencial Hidrogeniônico
dNTPs	<i>Deoxyribonucleoside Triphosphates</i>	<i>Pst</i> I	<i>Providencia stuartii</i>
<i>Eco</i> RI	<i>Escherichia coli</i>	RAPD	<i>Randomly Amplified Polymorphism</i>
EDTA	Ácido Etileno Diamono Tetrecético	RFLP	<i>Restriction Fragment Length Polymorphism</i>
g	Grama	RGA	<i>Resistance Gene Analogue</i>
GL	Grupo de Ligação	RNAse	Ribonuclease
H ₂ O	Água	rpm	Rotações por minuto
HCl	Ácido Clorídrico	SDS	<i>Sodium Dodecyl Sulfate</i>
<i>Hind</i> III	<i>Haemophilus influenza</i> linhagem d III	S-SAP	<i>Sequence Specific Amplified Polymorphism</i>
Kb	Kilobase	SSC	<i>Sodium , Salt Citrate</i>
LRR	<i>Leucine Rich Repeat</i>	SSR	<i>Simple Sequence Repeats</i>
M	Molar	<i>Taq</i>	<i>Thermus aquaticus</i>
M00	<i>Primer complementar ao adaptador Mse</i>	TBE	Tris-Base
mg	Miligramas	UV	Radiação Ultravioleta
MgCl ₂	Cloreto de Magnésio	V	Volts
mM	Milimolar	W	Watts

Resumo

O amendoim cultivado (*Arachis hypogaea*), pertencente à família das leguminosas é extremamente susceptível a doenças causadas por fungos e nematóides, não mostrando defesas naturais a esses patógenos dado a sua baixa variabilidade genética. Uma alternativa para descobrir novos genes de resistência é a construção de mapas genéticos visando a utilização dos marcadores em programas de seleção assistida. Para isso, o uso de espécies silvestres diplóides de *Arachis* é uma ferramenta de grande auxílio, pois as mesmas são fontes naturais de genes de resistência a patógenos. Através da técnica RGA-Display, *primers* degenerados e específicos baseados em regiões conservadas do domínio NBS (Nucleotide Binding Site – que é o maior grupo de genes de resistência em plantas) foram utilizados para isolar um total de 139 novos marcadores moleculares. Destes, 19 marcadores são análogos a genes de resistência (RGAs) em uma população de mapeamento obtida através do cruzamento de duas espécies diplóides silvestres de *Arachis* (*A. duranensis* e *A. stenosperma*). Esses marcadores foram separados em dois grupos baseados na estrutura da porção N-terminal do domínio: TIR e NãoTIR. Esses novos RGAs foram seqüenciados e alinhados juntamente com 67 RGAs previamente isolados para o gênero e a partir deste alinhamento uma árvore filogenética foi construída. Ao mesmo tempo, esses RGAs previamente isolados foram usados na técnica RFLP para a identificação de nove novos marcadores para a população de mapeamento. Um total de 148 novos marcadores foi isolado e 28 deles foram identificados como sendo RGAs. 39 marcadores, incluindo 15 RGAs, puderam ser mapeados em um mapa genético recém construído para o gênero com um LOD mínimo de 5,0 e fração de recombinação de 0,35 em oito dos dez grupos de ligação do mapa.

Palavras-chave: Amendoim; RFLP; RGA-Display; espécies silvestres; mapa genético.

Abstract

Peanut seeds (*Arachis hypogaea*), which belong to leguminosae family are extremely susceptible to diseases caused by fungi and nematodes, not showing typical defenses. An alternative to discovery novel resistant genes is the construction of genetic map of the wild type that is nature resources of pathogenic resistance. Degenerated and specific *primers* base on conserved regions of the nucleotide binding site (NBS) domain (encoded by the largest group of plant diseases resistance genes) were used in RGA-Display to isolate a set of 139 new markers, including 19 resistance gene analogs (RGAs) from a mapping population derived from a cross between two wild diploid species of *Arachis* (*A. duranensis* and *A. stenosperma*). These markers grouped into two classes of N-terminus structure TIR or NãoTIR group. These RGAs markers were sequenced and aligned with 67 previously isolated RGAs from *Arachis* and a phylogenetic tree was constructed. At the same time, these previously isolated RGAs from *Arachis* were used in RFLP technique to identify nine new markers for this population. A total of 148 new markers were identified by the two techniques and 28 were identified as RGAs. 39 markers were mapped in a recently constructed linkage map for the genus and 15 of these were RGAs. 39 markers were mapped in a recently constructed linkage map for the genus and 15 of these were RGAs. These markers were mapped with a minimum LOD of 5,0 and 0,35 of recombination fraction in eight of ten LG of the map.

Keywords: Peanut; RFLP, RGA-Display, Genetic Map; Wild Species.

1- Introdução

O gênero Arachis

O gênero *Arachis* vem sendo estudado com grande intensidade nos últimos anos devido ao grande potencial demonstrado por algumas espécies silvestres para o melhoramento do amendoim. O gênero pertence à Família Leguminosae, subfamília Papionoideae, é nativo da América do Sul, possivelmente no limite do Mato Grosso do Sul com o Paraguai segundo Gregory et al (1980) e Kaprovickas e Gregory (1994). Engloba cerca de 80 espécies (Valls e Simpson,1994), sendo 27 pertencentes à secção *Arachis* (Kaprovickas e Gregory, 1994). Dentro desta secção encontra-se a espécie *Arachis hypogaea*, o amendoim cultivado, que tem grande importância na alimentação humana em países da África e Ásia e os Estados Unidos (Kaprovickas, 1994).



Figura 1 Ilustração de plantas do gênero *Arachis*. Na figura da direita *A. pintoi* e na figura da esquerda *A. hypogaea*. Retirado de <http://www.plantatlas.usf.edu>.

Segundo a FAO (2005), a produção mundial de amendoim foi de 36,5 milhões de toneladas neste ano, com 25 milhões de hectares plantados. No Brasil, 125,5 mil hectares são plantados todo ano rendendo uma produção de 292 mil toneladas de amendoim. Em 2002, os principais países produtores da cultura foram a China (43,9%), Índia (22,9%), Estados Unidos (5,3%), Nigéria (4,5%), Indonésia (3,0%), Senegal (2,7%). O Brasil esteve em 13º lugar com 0,6% da produção mundial (Fávero, 2004). O amendoim é a quinta oleaginosa de maior importância no mundo (Moretzsohn et al., 2004). A importância econômica do amendoim se dá principalmente, pela produção de óleo, pastagem, alimentação animal e consumo *in natura* (Fávero, 2004).

Dados oriundos dos estudos citogenéticos e reprodutivos da secção *Arachis* revelaram a presença de três tipos de genomas distintos, denominados “A”, “B” e “D”. O genoma “A” está relacionado à presença de um par de cromossomos “A”. O genoma “B” é caracterizado pela ausência desse cromossomo “A”, enquanto que o genoma “D” está relacionado à presença de seis pares de cromossomos subtelocêntricos (Stalker et., 1991). A maioria das espécies da secção *Arachis* apresenta o genoma “A”, enquanto que *A. batizocoi*, *A. ipaensis*, *A. magna*, *A. valida*, *A. benesis* e *A. hoehnei* possuem genoma “B”. Somente *A. glandulifera* possui o genoma “D”.

Embora a maioria das espécies do gênero seja diplóide, existem quatro espécies tetraploides, e dentro da secção *Arachis* apenas duas são tetraploides, entre elas o *A. hypogaea*. Essas espécies apresentam simultaneamente os dois genomas “A” e “B” que ocorrem em espécies diplóides distintas da secção *Arachis* (Fernandez e Krapovickas, 1994).

O genoma do amendoim cultivado é tetraplóide e muito extenso ($2n = 4x = 40$) $n=1.74 \times 10^9$ bp, Bennet e Smith, 1976) diferentemente das espécies silvestres que pertencem a sua secção. Existem várias teorias sobre a origem do amendoim cultivado. A mais bem aceita foi proposta por Seijo et al., (2004) onde haveria acontecido um único evento de

alotetraploidização de um híbrido estéril de duas espécies silvestres (*A. duranensis* e *A. ipaensis*). A planta resultante, contendo dois genomas distintos, AA e BB, é reprodutivamente isolada dos seus parentais silvestres. Provavelmente por causa desta origem, *A. hypogaea* tem a sua base genética extremamente estreita (Kochert *et al.*, 1991). Essa baixa variabilidade genética no *A. hypogaea* pode causar um grande problema com a busca de resistência a pragas e patógenos. (Kochert *et al.*, 1996), sendo também muito difícil a geração de marcadores moleculares polimórficos para a construção de mapas genéticos.

Devido ao estreitamento da base genética do amendoim e a perda de genes importantes, principalmente aqueles ligados à resistência contra doenças, têm-se aumentado as preocupações de pesquisadores em relação a fungos, insetos e a patógenos presentes no solo. As doenças foliares causadas por fungos são consideradas limitantes para a cultura do amendoim. Doenças como mancha castanha (*Cercospora arachidicola*), mancha preta (*Cercosporidium personatum*), mancha barrenta (*Phoma arachidicola*), ferrugem (*Puccinia arachidis*) e verrugose (*Sphaceloma arachidis*) são potencialmente destrutivas para a cultura, e juntamente com doença causadas por vírus, nematóides e insetos podem comprometer anualmente até 70% da produção mundial (Subrahmanyam *et al.*, 1985). Outro problema encontrado na produção do amendoim é a aflatoxina produzida pelos fungos *Aspergillus flavus* ou *A. parasiticus*. Essa doença, apesar de não causar morte das plantas, pode causar redução na produção final da cultura, visto que as sementes infectadas devem ser descartadas e também há um aumento do custo de processamento pós-colheita. A ANVISA Agência Nacional de Vigilância Sanitária determina a obrigatoriedade dos testes de detecção da toxina, que é extremamente prejudicial à saúde humana e animal (www.anvisa.gov.br).

Espécies Silvestres

Em contraste ao amendoim cultivado, as espécies silvestres diplóides apresentam alta diversidade genética e são uma grande fonte de resistência natural a fatores bióticos (pragas e patógenos) e abióticos (stress hídrico, toxidez do solo) (Galgaro et al., 1997). Essa alta diversidade está sendo usada para facilitar a construção de mapas genéticos derivados de cruzamentos entre *A. hypogaea* com parentes silvestres e de cruzamentos de parentes silvestres entre si, sendo que as espécies silvestres resistentes têm grande potencial para melhorar o amendoim cultivado (Stalker et al., 1997). O Brasil contém alguns dos centros de diversidade genética mais ricos do mundo, com o centro de origem de um número de plantas de grande importância sócio-econômica (ex. mandioca, amendoim, caju, abacaxi e cacau) assim como parentes silvestres de muitas outras plantas. A Embrapa Recursos Genéticos e Biotecnologia possui um banco de germoplasma de *Arachis* silvestre com mais de 1000 acessos de espécies silvestres e outras 1000 variedades primitivas, algumas das quais já identificadas como fontes de resistência para fungos e nematóides (Nelson et al., 1989, Bertioli et al., 2000, Fávero, 2004 e Guimarães et al., 2005).

A transferência de resistências de espécies selvagens para o amendoim cultivado tem sido extremamente difícil pela diferença de ploidia, e conseqüente barreira de fertilidade. Teoricamente, a transferência de genes de resistência é feita através de cruzamentos complexos, seguida de vários retrocruzamentos com *A. hypogaea* (Simpson et al., 2001). Mas isso só foi conseguido uma vez pelo próprio Simpson, quando introduziu resistência a *M. arenaria* e lançou a cultivar COAN. Para exemplificar esta dificuldade, a cultivar COAN só estava disponível no mercado, cerca de 23 anos após o primeiro cruzamento (Simpson, C. comunicação pessoal).

Genes de Resistência

Ao longo da evolução, as plantas têm desenvolvido vários mecanismos de defesa para sua proteção contra doenças e parasitas. Essas defesas incluem: estruturas pré-existentes, como cutículas e pêlos que são barreiras físicas contra o patógeno, ou seja, não permitem a invasão da planta; Produção de inibidores como o inibidor de α -amilase que se ligam às enzimas digestivas do patógeno impedindo que o mesmo assimile o amido da planta; compostos fenólicos, taninos e lectinas; e enzimas, como as glucanases e quitinases. Outra classe de mecanismos de defesa envolve o reconhecimento específico do patógeno pela planta atacada. Entre os eventos celulares que caracterizam este tipo de resistência estão: a indução da expressão de genes de defesa e a rápida morte celular no local da infecção (Agrios, 1997). Estes genes são chamados Genes de Resistência (R-genes).

Um grande número de R-genes foi identificado na planta modelo *Arabidopsis thaliana* (Deslandes et al., 2002) e em várias outras espécies (Lawrence et al., 1995; Thomas et al., 1997; Wang et al., 1999). Muitas dessas proteínas de resistência são caracterizadas pela presença dos domínios LRR e NB-ARC (Meyers et al., 1999). O domínio LRR (Leucine Rich Repeat, Kobe e Deisenhorfer 1994) aparece como responsável primário pelo elicitor de reconhecimento (Jones e Jones, 1997), ou em um modelo alternativo, atua como “guarda” da maquinaria celular que é susceptível ao ataque de patógenos (Dangl e Jones, 2001). O domínio NB-ARC (“Nucleotide Binding, human; APAF-1, Plant Resistance Genes and *Caenorhabditis elegans* CED-4”, Van der Biezen e Jones, 1998) age nos caminhos de transdução de sinal que operam em resposta ao ataque de patógenos e está envolvido na morte celular programada relatada em diferentes organismos. Em plantas, a única função associada com esse domínio é em resistência a doenças, que é frequentemente manifestada como uma resposta de hipersensibilidade envolvendo morte celular programada. Em genes de

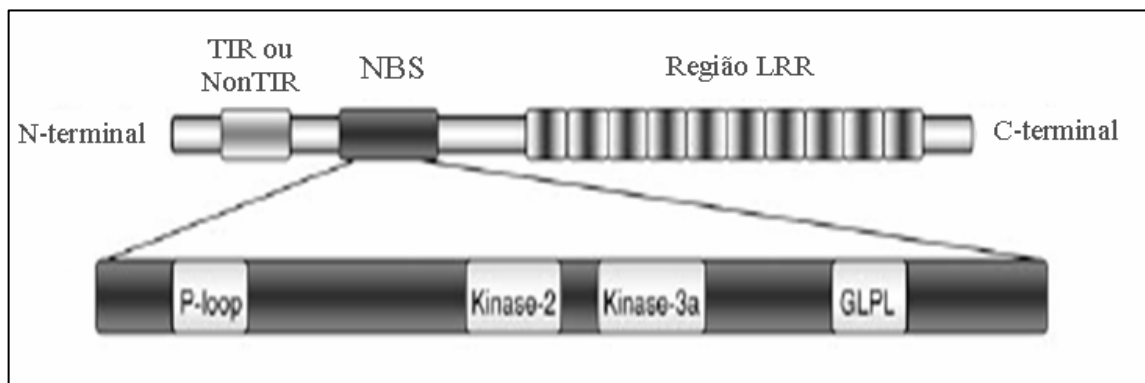
plantas é mais comum nomear este domínio como NBS (“Nucleotide Binding Site”, Tameling et al., 2002).

O domínio NBS-LRR é dividido em duas categorias baseado nas regiões conservadas na porção N-terminal dos seus produtos (Meyers et al., 1999; Young et al., 2000; Yuksel et al., 2005). O primeiro grupo, chamado de NãoTIR, compreende genes que codificam o padrão conservado “Leucine-Zipper-Like” e inclui, por exemplo, o gene *Mi* de tomate, responsável pela resistência ao nematóide das galhas *Meloidogyne incognita* (Milligan et al., 1998). Análises recentes neste grupo indicaram que no lugar onde deveria estar a estrutura TIR se encontra uma outra estrutura em “coiled-coiled” (Young et al., 2000). O segundo grupo, TIR, é composto por genes que codificam produtos contendo uma estrutura homóloga à proteínas receptoras humanas “Interleucina 1” e de *Drosófila* “Toll-Like”. Um exemplo é o gene de resistência ao vírus do mosaico em Tabaco (Whitham et al., 1996). As espécies de *Arachis* possuem as duas formas da estrutura (Bertioli et al., 2003).

É comum a ocorrência de motivos dentro do domínio NBS (Figura 2). Esses motivos são regiões conservadas dentro dos domínios que facilitam a construção de *primers* para a amplificação de RGAs ou genes similares. Alguns desses motivos são: P-loop, Kinase-2, Kinase-3a, GLPL e RNBS-D (Meyers et al., 1999 e Yuksel et al., 2005). Por exemplo, um grande número de RGAs foi amplificado em soja usando *primers* degenerados desenvolvidos para o domínio NBS (Leister et al., e Yu et al., 1996). Similarmente, RGAs de milho (Collins et al., 1998) e alfaca (Meyers e Michelmore, 1998) também foram isolados e caracterizados.

A caracterização mais expressiva de RGAs em *Arachis* spp. foi feita por Bertioli et al. (2003), onde foram gerados 78 RGAs utilizando *primers* degenerados construídos para o domínio NBS de *A. hypogaea* var. Tatu e quatro espécies silvestres (*A. duranensis*, *A. cardenasii*, *A. stenosperma* e *A. simpsoni*). A avaliação comparativa e evolucionária de 63 seqüências ininterruptas com ORFs (Open Reading Frame) indicaram sua proximidade

evolucionária com RGAs de outras leguminosas, incluindo *Glycine max*, *Lotus japonicus*, *Medicago truncatula* e *Phaseolus vulgaris*. A inclusão destas seqüências em um mapa genético de *Arachis* sp. recém construído, possibilitaria a identificação de genes de resistência ligados a estes RGAs, auxiliando os programas de melhoramento da cultura. Recentemente, Yuksel e colaboradores (2005) isolaram um total de 234 novos RGAs para *A. hypogaea* a partir de *primers* desenhados para regiões conservadas no domínio NBS. Com isso, haverá um grande auxílio no entendimento da evolução e organização de genes de resistência no genoma do amendoim.



Adaptado de Yuksel et al., 2005

Figura 2 Domínio NBS-LRR representado esquematicamente. As regiões TIR e NãoTIR (Coiled-coil) também estão ilustradas. Os motivos P-loop, Kinase-2, Kinase-3a e GLPL estão representados na parte inferior da figura.

Marcadores Moleculares

Diversas técnicas de biologia molecular estão hoje disponíveis para a detecção de variabilidade genética ao nível de seqüência de DNA, ou seja, para a detecção de polimorfismo genético. Essas técnicas permitem a obtenção de um número praticamente ilimitado de marcadores moleculares cobrindo todo o genoma do organismo. Tais marcadores podem ser utilizados para as mais diversas aplicações, tanto no estudo de genética como na

prática de melhoramento de plantas. Com o advento das técnicas modernas de biologia molecular, surgiram diversos métodos de detecção de polimorfismo genético diretamente ao nível de DNA.

Inicialmente, a utilização de enzimas de restrição permitiu a análise de polimorfismo de comprimento de fragmentos de restrição de DNA (“Restriction Fragment Length Polymorphism” – RFLP; Botstein et al., 1980). Esta técnica já foi uma das classes de marcadores mais amplamente utilizadas em genética e melhoramento de plantas. O polimorfismo observado na técnica ocorre porque o DNA de indivíduos geneticamente distintos difere na seqüência de nucleotídeos ao longo da fita. A presença ou ausência de seqüências específicas reconhecidas e clivadas pelas diferentes enzimas de restrição, pode variar entre diferentes indivíduos, gerando polimorfismo. Esses marcadores possuem mais uma característica importante para o mapeamento, eles produzem expressão co-dominante, isto é, em cada loco é possível identificar genótipos heterozigotos e homozigotos, gerando mais informação em nível genético e permitindo uma análise detalhada da ação gênica e interação entre alelos (Ferreira e Grattapaglia, 1995).

Mais recentemente, o desenvolvimento do processo de amplificação em cadeia utilizando uma DNA polimerase (PCR, Mullis e Faloona, 1987; Saiki et al., 1988) levou à descrição de outras classes de marcadores moleculares. Aliadas às técnicas de clonagem e seqüenciamento de DNA, estas metodologias têm possibilitado um rápido acúmulo de informações sobre a estrutura de genomas (Ferreira e Grattapaglia, 1995). Os marcadores moleculares baseados em PCR são hoje a melhor ferramenta para a construção de mapas genéticos de plantas. A técnica AFLP (Amplified Fragment Length Polymorphism) publicada por Vos et al., (1995) é baseada na detecção de fragmentos de DNA genômico, clivados com enzimas de restrição, pelo uso de PCR. As amplificações são produzidas sem um conhecimento prévio da seqüência de DNA, usando *primers* genéricos. A característica

principal desta técnica é a capacidade de revelar simultaneamente muitas regiões diferentes distribuídas de maneira aleatória ao longo do genoma (Mueller e Wolfenbarger, 1999). A análise é baseada na amplificação seletiva por PCR de um subconjunto de fragmentos genômicos gerados após a digestão com uma enzima de corte raro combinada com uma enzima de corte freqüente. A principal vantagem desta tecnologia é o grande número de fragmentos produzidos, aumentando o poder de detecção da variabilidade genética. O número de fragmentos detectados em uma reação depende do *primer* utilizado, onde podem ser adicionadas bases seletivas ao final de sua seqüência. A técnica AFLP é muito robusta e reproduzível devido à alta estringência usada na temperatura de anelamento dos *primers*. Em resumo, a técnica combina a reprodutibilidade do RFLP com o poder e praticidade da PCR.

A técnica S-SAP (Sequence Specific Amplified Polymorphism - Waugh *et al.*, 1997) é baseada em AFLP, onde um dos *primers* desenhados ao adaptador do corte da enzima é substituído por um *primer* que reconhece regiões repetidas do genoma, os Retrotransposons. A técnica é fundamentada, então, na detecção de polimorfismo de DNA baseada na distribuição dos retrotransposons em regiões adjacentes aos sítios de restrição das enzimas *Pst*I e *Mse*I (corte raro e freqüente respectivamente). O seqüenciamento das bandas polimórficas é grande particularidade da técnica, onde o resultado pode ser confirmado. O *primer* complementar a regiões repetidas – retrotransposons pode ser marcado com radioatividade para uma melhor detecção dos fragmentos.

Em 2000, Hayes e Shagai Maroof estabeleceram uma técnica de AFLP enriquecido para RGAs, onde um *primer* convencional foi trocado por um degenerado que reconhece uma região conservada dentro do domínio NBS. A técnica ficou conhecida como AFLP-RGA. Recentemente, outro grupo (Linden *et al.*, 2004) aprimorou uma técnica que combina o AFLP e o S-SAP chamada por eles de “NBS-profiling”. Essa técnica recebeu variações e pode ser chamada de RGA-display, RGA-SAP, Motif-display etc. Essa abordagem é similar ao AFLP

onde DNA genômico é cortado com enzimas de restrição e ligado a um adaptador nas extremidades. A amplificação por PCR é feita com um *primer* relativo ao adaptador e outro à região codificante de motivos de proteínas em RGAs (como o P-loop, Kinase-2, etc). O tamanho do produto amplificado varia com a distância do sítio de restrição até o motivo, causando polimorfismo entre indivíduos. Bandas polimórficas podem ser retiradas do gel e seqüenciadas (como no S-SAP), confirmando a similaridade com RGAs já descritos. Essa técnica apresenta como vantagens a praticidade e rapidez inerentes ao uso de marcadores do tipo PCR e a reprodutibilidade do RFLP. Os produtos podem ser detectados por coloração com Nitrato de Prata (Creste et al., 2001) onde há uma diminuição dos custos da técnica.

Mapeamento Genético

Vários tipos de marcadores moleculares como RAPD (Random Amplified Polymorphic DNAs), AFLP, RFLPs e SSRs (Simple Sequence Repeats) têm apresentado baixo polimorfismo dentro do germoplasma de *A. hypogaea* e isto torna quase inviável a construção de mapas usando populações derivadas de cruzamentos intra-específicos (Kochert et al., 1996, Halward et al., 1991 e Herselman et al., 2003). Recentemente, uma maior compreensão sobre a origem de *A. hypogaea* permitiu a re-síntese da espécie a partir de espécies silvestres (Fávero, 2004). Isso, juntamente com um melhor entendimento da estrutura de genes de resistência tem o potencial de acelerar a transferência dos mesmos para cultivares de interesse através do melhoramento tradicional auxiliado pela seleção assistida por marcadores moleculares - MAS (Hash et al., 2003). Para o uso de MAS para acelerar programas de geração de cultivares resistentes a pragas e patógenos, é necessário identificar marcadores proximalmente ligados a genes de resistência.

Em *Arachis* spp., poucos mapas de ligação foram publicados. Halward et al., (1993) desenvolveu um mapa baseado em RFLP usando uma população F₂ derivada do cruzamento entre duas espécies silvestres diplóides com genoma AA (*A. stenosperma* e *A. cardenasii*). Neste trabalho, um total de 117 *loci* foram mapeados em 11 grupos de ligação. Mais tarde, marcadores RFLP e RAPD também foram usados para se conseguir a introgressão de segmentos cromossômicos de *A. cardenasii* (genoma AA) dentro do genoma de *A. hypogaea* (Garcia et al., 1995). Em 2001, Burrow e colaboradores construíram um outro mapa baseado em RFLP para uma população tetraplóide, onde o anfidiplóide sintético [*A. batizocoi* X (*A. cardenasii* e *A. diogoi*)]^{4X} usado como parental doador, foi cruzado com *A. hypogaea* para gerar uma população de mapeamento com 78 indivíduos. 370 *loci* de RFLP foram mapeados, obtendo-se 23 grupos de ligação. Em outro trabalho, marcadores RAPD, SCAR e RFLP mostraram ser úteis na identificação de marcadores ligados a genes de resistência neste mesmo cruzamento (Garcia et al., 1996). Recentemente, marcadores AFLP foram usados para mapear genes que conferem resistência ao afídio vetor do vírus da roseta do amendoim, que é a doença viral mais destrutiva de amendoim na África (Herselman et al., 2004). Em 2005, Moretzsohn e colaboradores publicaram um mapa baseado em Microssatélite (SSR) para uma população F₂ de 93 indivíduos obtidos através do cruzamento de duas espécies silvestres diplóides de genoma AA (*A. stenosperma* e *A. duranensis*). Um total de 271 novos marcadores Microssatélite foram mapeados em 11 grupos de ligação cobrindo uma distância de 1.230,89 cM no mapa e com a distancia média de 7.24 cM entre os marcadores.

2- Objetivos

Objetivos Gerais

O objetivo deste trabalho é o desenvolvimento de marcadores moleculares baseados em seqüências homólogas a genes de resistência (RGAs) através de duas técnicas: RFLP e RGA-display que possam ser incorporados no mapa genético de acessos silvestres de *Arachis*.

Objetivos Específicos

Considerando a importância de marcadores moleculares em programas de melhoramento do amendoim baseados em espécies silvestres, que são fontes naturais de resistências a pragas e patógenos, os principais objetivos do presente trabalho são: (1) desenvolvimento de marcadores moleculares baseados em RGAs através da técnica RGA-display; (2) Construção de marcadores co-dominantes baseados em RGAs através da técnica RFLP; (3) Genotipagem desses marcadores em uma população de mapeamento derivada de dois acessos silvestres de *Arachis* (*A. stenosperma* e *A. duranensis*) contendo 93 indivíduos e (4) inclusão desses marcadores em um mapa de ligação para essas duas espécies que representam o genoma AA de *Arachis* (Moretzsohn et al., 2005).

3- Material e Métodos

Material vegetal e população de mapeamento

Através de bioensaios realizados em trabalhos anteriores (Leal-Bertioli et al., 2000.), dois acessos silvestres de *Arachis* foram escolhidos para a construção da população de mapeamento: *Arachis duranensis* (acesso: K7988) e *A. stenosperma* (acesso: V10309). As duas espécies são diplóides e possuem genoma A, sendo a segunda resistente aos nematóides das galhas *Meloidogyne arenaria* raça 2 e *M. javanica* (Leal-Bertioli et al., 2000) e aos fungos de mancha foliar *Cercospora arachidicola* e *Cercosporidium personatum* (Proite et al., 2002). O cruzamento dessas duas espécies gerou uma planta híbrida fértil (F₁) que foi autofecundada para gerar a população de mapeamento usada nesse trabalho (população F₂). Essa população consiste de 93 plantas mantidas em vasos de concreto, dentro de casas de vegetação na EMBRAPA Recursos Genéticos e Biotecnologia (Brasília-DF).

Para os experimentos realizados neste trabalho, DNA total de *A. duranensis*, *A. stenosperma*, da planta híbrida e de 93 plantas da população F₂ foi extraído de folhas jovens (recém surgidas e antes da expansão foliar) para evitar a extração simultânea de compostos fenólicos, polissacarídeos e proteínas presentes nas folhas adultas. Um protocolo de extração de DNA baseado em CTAB publicado por Doyle et al. (1991) foi modificado para o uso em *Arachis* sp. onde um novo passo de precipitação com CTAB foi adicionado: 200 mg de tecido vegetal foram macerados em cadinho com o auxílio de Nitrogênio líquido e adição de 700 µl de 2x CTAB e 2 µl de 2-mercaptoetanol em cada tubo de microcentrifuga de 2 ml. As amostras foram incubadas a 65°C por 60 minutos. Após esse período foram adicionados 700 µl de clorofórmio-isoamílico (24:1) para cada amostra e misturou-se até formar uma emulsão.

As amostras foram centrifugadas a 14000 rpm por 5 minutos. A fase aquosa foi transferida para dois tubos novos de 1,5 ml. Foram adicionados 600 µl de tampão 1X CTAB, agitou-se lentamente e centrifugou-se a 14000 rpm por 1 minuto. O “pellet” permaneceu aderido ao fundo do tubo. O sobrenadante foi descartado cuidadosamente e o DNA dos dois tubos foi ressuspensionado em 300 µl de 1,2M NaCl em cada tubo e o volume de 600 µl foi transferido para um único tubo de 2 ml. As amostras foram centrifugadas a 12000 rpm por 5 minutos e o sobrenadante foi transferido para novo tubo. O DNA foi precipitado com 1 ml de etanol absoluto e agitado lentamente. As amostras foram centrifugadas a 14000 rpm por 10 minutos e descartou-se o sobrenadante. O “pellet” de cada tubo foi lavado com 500 µl de etanol 70% duas vezes. Os “pellets” foram secos e ressuspensionados em 100 µl de H₂O milli-Q estéril com 0,01mg/ml de RNase-A. As amostras foram quantificadas em gel de agarose 0,8% usando um marcador de massa (High Mass Ladder – Invitrogen) e de tamanho (1 Kb Ladder – Invitrogen) e então foram diluídas e mantidas a -20°C até a realização do experimento.

RFLP

A realização dos Southern-blots foi de acordo com Southern, (1975) onde 10µg de DNA total de cada amostra (K7988, V10309, F₁ e 93 plantas F₂) foram digeridos com 100 unidades de uma das enzimas de restrição: *EcoRI* ou *HindIII*. As duas enzimas foram testadas primeiramente nos parentais (K7988 e V10309) com cada sonda para se escolher a melhor (com maior polimorfismo) para a genotipagem da população F₂. A digestão foi feita a 37°C por oito horas. Os produtos de digestão foram separados em gel de agarose 0,8% a 20 V por aproximadamente 16 h. O gel foi desnaturado em 0,25N de HCl por 30 minutos, neutralizado em 0,5N de NaOH por 15 minutos e transferido para membrana de nylon Hybond N⁺

(Amersham, Pharmacia) por capilaridade em SSC 10X (Anexo 1) com o auxílio de um transferidor a vácuo (BioRad). As amostras foram fixadas à membrana em luz UV em um forno *cross-link* (30 segundos na função Auto Cross Link) por um minuto, identificadas e armazenadas a 4°C até a realização do experimento. Para a pré-hibridização, as membranas foram colocadas em garrafas de vidro contendo solução de pré-hibridização (Anexo 1) e aquecidas em movimentos lentos a 65°C por 2 horas em forno de hibridização.

As sondas utilizadas neste trabalho tiveram como base os RGAs isolados por Bertioli et al., (2003), onde os mesmos foram classificados em diferentes clades, de acordo com sua similaridade. Para o experimento, foi escolhido um RGA representativo de cada clade em questão, de forma a se ter uma amostra representativa de todas as seqüências isoladas. Os RGAs utilizados estão listados na Tabela 1. Para uso como sonda, os RGAs foram re-amplificados por PCR a partir de plasmídios contendo seqüências de RGAs. 50ng de cada RGA foram desnaturados e adicionados ao kit de marcação *Ready-to-Go* d-CTP (Amersham Pharmacia) juntamente com 5 µl de ³²P. A sonda foi incubada por 30 minutos a 37°C em banho-maria. Os nucleotídeos não incorporados foram eliminados da solução utilizando o kit MicrospinTM S-400 HR Columns (Amersham Pharmacia). A sonda já marcada com radioatividade foi então adicionada à garrafa contendo a membrana com a solução de pré-hibridização. A hibridização ocorreu a 65°C por 16 horas. Após o período de hibridização as membranas foram lavadas da seguinte maneira: uma vez com solução de Lavagem I (Anexo 1) a 65°C por 15 minutos e duas vezes com solução de Lavagem II (Anexo 1) a temperatura ambiente por 15 minutos. As membranas foram secas ao ar e colocadas em exposição com filme de raio-X (Kodak-Bio) por aproximadamente 4 dias. Os filmes foram revelados da seguinte forma: 3 minutos em revelador 1X (Dektol – Kodak) e 1 minuto em fixador 1X (Kodak).

Tabela 1 RGAs isolados por Bertoli *et al.* (2003) e usados como sonda para a hibridização em Southern-blots neste trabalho.

RGA	Clade	ID GenBank*	Espécie de origem
S5-A-375	O2	AY157945	<i>A. stenosperma</i>
S5-A-384	N1	AY157942	<i>A. stenosperma</i>
S5C8-AXY-370	Q3	AY157784	<i>A. stenosperma</i>
S1-A-36	M11	AY157808	<i>A. stenosperma</i>
S1-A-37	O1	AY157811	<i>A. stenosperma</i>
C8-V-434	E8	AY157804	<i>A. cardenasii</i>
S4-A-164	R3	AY157813	<i>A. stenosperma</i>
S1S2-A-152	A20	AY157810	<i>A. stenosperma</i>
T-A-44	M1	AY157819	<i>A. hypogaea</i> var. Tatu

Tabela retirada de Bertoli *et al* (2003). * ID consiste no número de acesso de cada RGA depositado no GenBank.

Dos 78 RGAs publicados e depositados no GenBank nove foram escolhidos para serem testados como sonda nos parentais com as duas enzimas (total de 18 testes) e, os que apresentaram maior grau de polimorfismo, foram utilizados na população F₂ contendo 93 indivíduos. As membranas produzidas puderam ser reaproveitadas (hibridizadas com outra sonda) apenas esperando o período de dois meses, equivalente ao decaimento da radioatividade presente no ³²P.

RGA- Display

Para a realização do experimento de identificação de marcadores moleculares pela técnica conhecida como RGA-Display, foram realizados alguns testes e modificações da técnica original (Linden et al., 2004). Essas modificações serão discutidas mais à frente.

Para o experimento, 100 ng de DNA total de *A. stenosperma* V10309, *A. duranensis* K7988, da planta híbrida e das 93 plantas da população F₂ foram digeridos simultaneamente com as enzimas de restrição: *Pst*I e *Mse*I a 37°C por 8 h. Adaptadores (Tabela 2) foram

adicionados e ligados às extremidades cortadas dos fragmentos utilizando a seguinte reação de ligação: 0,5 M de adaptador para *Pst*I, 5 M de adaptador para *Mse*I, 1M de ATP, 5X RL Buffer (Apêndice I), 1 unidade de T4 DNA Ligase, 2,5 unidades de *Pst*I, e 2,5 unidades de *Mse*I. A ligação foi feita a 37°C por 5 horas (Vos et al., 1995; Maughan et al., 1996). A Figura 3 mostra um esquema representativo do preparo das amostras de DNA para o uso no experimento. Os adaptadores e os *Primers* para os mesmos foram baseados nos descritos por Vos et al., (1995).

Logo após a ligação, foi feita uma “pré-amplificação” desse DNA digerido e ligado utilizando *primers* (P00 e M00 – Tabela 2) que reconhecem os adaptadores já ligados aos fragmentos nas duas extremidades. Para um volume de 50 µl de reação foram utilizados: 10% de tampão para PCR sem MgCl₂, 0,9 mM de MgCl₂, 0.6 µM de cada *primer*, 0.25 mM dNTPs, 0.5 U de *Taq* polymerase (Platinum Invitrogen) e 150 ng de cada DNA digerido e ligado. O programa utilizado foi 30 ciclos de: 30s a 95° C, 1 min. a 52° C, 1 min. a 72° C. Antes dos ciclos foi adicionado um passo de 95°C por 15 min para a ativação da enzima. Esses DNAs digeridos, ligados aos adaptadores e amplificados (chamados de “pré-amps”) foram diluídos 20 vezes e mantidos a -20°C, onde serviram de base para as amplificações com os *primers* RGAs.

Tabela 2 Lista de Adaptadores e *primers* para adaptadores usados em RGA-Display

Primer *	Observ.	Seqüência (5' - 3')
Adaptador <i>Pst</i> I (PstAdpt)	Braço curto	TGTACGCAGTCTAC
	Braço longo	CTCGTAGACTGCGTACATGCA
Adaptador <i>Mse</i> I (MseAdapt)	Braço curto	TACTCAGGACTCAT
	Braço longo	GACGATGAGTCCTGAG
<i>Primer</i> para adaptador <i>Pst</i> I P00	Se liga ao PstAdpt	GACTGCGTACATGCAG
<i>Primer</i> para adaptador <i>Mse</i> I M00	Se liga ao MseAdpt	GATGAGTCCTGAGTAA

* Vos et al., 1995

Após a confecção dos “pré-amps”, foi feita então a amplificação dos fragmentos RGAs utilizando-se um *primer Mse* (que reconhece o adaptador ligado ao sítio da enzima *MseI*) no sentido reverso e um *primer* (específico ou degenerado) que reconhece uma região conservada do motivo KinaseII no Domínio NBS para o sentido direto. Os *primers* utilizados nessa técnica complementares ao motivo Kinase-2 no domínio NBS foram sintetizados a partir de seqüências descritas como RGAs para *Arachis* em Bertoli et al., 2003 (*Primers*: P2-12, P2-16, P2-19, P2-378, P2-391 e P2-414) ou descrito por Linden et al., 2004 (*Primer* NBS 5). Os *primers* complementares ao adaptador ligados ao sítio de restrição da enzimas *MseI* foram feitos a partir de *primers* já descritos para AFLP em Vos et al., 1995 (Assim como os adaptadores e *primers* complementares aos adaptadores). Os *primers* foram sintetizados por WMed (World Medical Representações) ou IDT (Integrated DNA Technologies). As seqüências dos *primers* utilizados neste experimento estão listadas na Tabela 3:

Tabela 3 Seqüência dos *primers* utilizados na amplificação dos DNAs pré-amplificados (Pré-amps), local de reconhecimento do *primer* e característica da seqüência.

<i>Primer</i>	Seqüência (5' - 3')	Local	Observ.
P- <i>MseI</i> ***	GATGAGTCCTGAGTAA	Sítio de <i>MseI</i>	Sem base seletiva
P2-391*	IGCIYTIATIGTIYTI GAYGAYGTIGA	Motivo kinase-2	Degenerado
P2-414*	GITTYGTIYTIATGYTIGGAYGAYYTIITGG	Motivo kinase-2	Degenerado
P2-19*	CTTATTGTTCTTGATGATGATGTCGA	Motivo kinase-2	Específico
P2-12*	GTCTCTTGTTCTTGATGATGTGTAAT	Motivo kinase-2	Específico
P2-378*	AAAATTATTTTGATTCTAGATGACATTGA	Motivo kinase-2	Específico
NBS 5**	YYTKRTHGTMITKGATGATGTITGG	Motivo kinase-2	Degenerado
P2-16*	TCTTCTTGTTCTTGATGTAGGT	Motivo kinase-2	Específico

* *Primers* desenhados para RGAs de *Arachis* (Bertoli et al, 2003) ** Linden *et al.* (2004).

*** Padrão para AFLP (Vos et al., 1995) aqui mostrado sem adição de bases seletivas.

Essa segunda amplificação foi conduzida em uma reação de 15 µl contendo: 10% de tampão para PCR sem magnésio; 0,6 µM de cada *primer*; 0,9 mM de MgCl₂; 0,25 mM de dNTPs e 1 unidade de Platinum *Taq* DNA polimerase (Invitrogen). O programa utilizado foi o mesmo da primeira etapa de amplificação. O *primer Mse* foi utilizado com uma, duas ou três

bases seletivas adicionadas ao final de sua seqüência para restringir o número de fragmentos produzindo, diminuindo a complexidade do padrão a ser observado. No total, foram sintetizados 20 *Primers* para *MseI* (com uma, duas ou três bases seletivas) e sete para o motivo Kinase-2 (degenerados ou específicos), somando 140 combinações possíveis.

Os produtos da segunda etapa de PCR foram misturados com Tampão Formamida (Xu et al., 1997. Anexo 1), desnaturados a 95°C por 5 min e separados em gel de poliacrilamida a 4% por 1,5h a 60W. O gel foi corado com Nitrato de Prata Segundo Creste et al.,2001(Anexo 1). As bandas polimórficas foram analisadas e retiradas do gel com o auxílio de uma lâmina e re-hidratadas em 100 µl de H₂O Milli-Q estéril e aquecidas a 100°C por 10 min como descrito por Upender et al., (1995). O DNA eluído foi re-amplificado usando as mesmas condições descritas acima para a amplificação do fragmento original. O produto da re-amplificação foi separado em gel de agarose a 1% e novamente eluído com o auxílio do kit QIAquick Gel Extraction (Quiagen).

Seqüenciamento e processamento das amostras

As bandas polimórficas eluídas foram então seqüenciadas (aproximadamente 60 ng de DNA) nos dois sentidos. As reações de seqüenciamento foram otimizadas utilizando um dos dois *Primers* usado para a amplificação do fragmento polimórfico. Foi usado o Kit de seqüenciamento BigDie Terminator (Applied Biosystems, CA, USA) em um seqüenciador automático ABI377 (Applied Biosystems).

As seqüências obtidas foram processadas e analisadas utilizando-se a ferramenta Staden Package (Staden et al., 2003), onde os dois sentidos de cada seqüência puderam ser montados para se conseguir uma melhor qualidade de leitura. A ferramenta “Pregap4” foi

utilizada para o pré-processamento das amostras como controle da qualidade das seqüências, conversão de formatos e estimativa da qualidade de cada base ($\text{Phred} \geq 20$). Após esse pré-processamento, a ferramenta “Gap4” foi utilizada para o processamento propriamente dito, onde os pares de seqüência foram juntados quando conveniente, onde também se juntou as seqüências provenientes das repetições do seqüenciamento.

As seqüências de DNA processadas foram submetidas aos programas BLASTX e BLASTN (Altschul et al., 1997), contra dados de todas as prováveis proteínas de *Arabidopsis* e ao banco de dados total do *GenBank*. Os “Hits” foram considerados significativos quando tiveram um Nível de Significância “E-value” menor que 1×10^{-1} . Para uma melhor análise, as amostras processadas também foram traduzidas para sua seqüência correspondente protéica, com o auxílio da ferramenta de tradução para aminoácidos “Translate – TrEMBL” no site www.expasy.ch/tools/dna.html. As seqüências traduzidas foram então submetidas à ferramenta “Motif-Scan” (Falquet et al., 2002 - www.expasy.ch/prosite) onde regiões conservadas do genoma são rastreadas e comparadas com seqüências já conhecidas. Essa ferramenta utiliza o banco de dados do “Pfam” para a busca dos motivos e domínios.

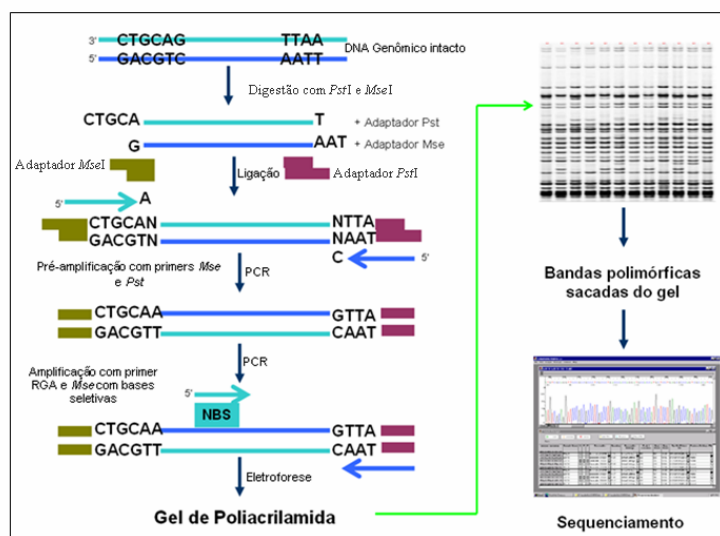


Figura 3 Representação esquemática da técnica RGA-display. DNA genômico é clivado simultaneamente com duas enzimas de restrição (uma de corte raro e uma de corte freqüente).

Adaptadores são ligados às extremidades clivadas para o reconhecimento de *primers*. Uma primeira PCR seleciona os fragmentos com adaptadores ligados corretamente e uma segunda PCR (utilizando *primers* que reconhecem regiões conservadas do motivo NBS) seleciona fragmentos com possíveis seqüências RGAs. Os fragmentos são retirados do gel e seqüenciados.

Alinhamento das Seqüências e Árvore de Similaridade

Para o alinhamento das seqüências o programa de domínio público JalView 2.07 (www.jalview.org) foi utilizado, onde as seqüências protéicas foram alinhadas de acordo com sua similaridade. Sessenta e oito RGAs isolados por Bertoli et al. em 2003 para *Arachis* também foram utilizados para os alinhamentos realizados nesse trabalho. O alinhamento foi do tipo “*ClustalW Multiple Sequence Alignment Realign*” e as regiões conservadas entre as amostras foram coloridas com a ferramenta “*ClustalW*”. Para a construção da Árvore de Similaridade foi utilizada a ferramenta “*Calculate Tree*” com a opção “*Average Distance Using % Identity*”, onde uma primeira árvore foi construída apenas com as seqüências RGAs obtidas neste trabalho e, logo após, uma segunda árvore foi construída onde foram adicionados os RGAs isolados por Bertoli et al., 2003.

Genotipagem da população F₂ e Mapeamento

Cada combinação de *primers* que produziu pelo menos uma banda polimórfica na técnica RGA-Display foi utilizada para genotipar toda a população F₂ (93 indivíduos) juntamente com seus parentais e a planta híbrida. Para cada banda analisada foi construída uma matriz de dados para o processamento da genotipagem. Através do programa Excel (Microsoft) foram geradas tabelas contendo informações relativas à presença ou ausência da

banda polimórfica obtida para cada indivíduo da população F_2 . Esses dados foram atribuídos a partir de notas dadas a determinado padrão: Nota 1 (um) para a presença da banda em um indivíduo; Nota 0 (zero) para a ausência da banda; e nota 9 (nove) para os dados faltantes (amostras que falharam ou impossível de visualizar).

Os marcadores desenvolvidos por RGA-display e os desenvolvidos por RFLP foram separados de acordo com o parental de origem. Esses marcadores polimórficos foram analisados na população de mapeamento (população F_2) consistindo de 93 plantas. Um teste χ^2 foi feito para testar a hipótese nula de segregação de 1:2:1 (no caso dos marcadores co-dominantes) e de 3:1 (no caso dos marcadores dominantes) em todos os marcadores genotipados. Apenas os marcadores que não mostraram distorção ($P > 0,05$) da segregação esperada foram usados na construção inicial do mapa. As análises de ligação foram feitas usando o programa Mapmaker Macintosh versão 2.0 (Lander et al., 1987). Um LOD mínimo utilizado foi de 5,0 e o máximo de fração de recombinação (h) de 0,35 foi escolhido para os grupos de ligação (LG) determinado no comando “group”. O máximo de ligação entre os marcadores dentro de cada grupo de ligação foi estimado pelo método de correlação de matrizes usando o comando “first order” pela análise de três pontos (comando “three point”) para grupos contendo mais de sete marcadores. A ordem dos marcadores foi confirmada por comparação da probabilidade da possível ordem usando a análise de múltiplos pontos (comando “compare”) e permutando todas as ordens das trincas adjacentes (comando “ripple”). *Loci* que puderam ser ordenados com a probabilidade de proporção $\geq 1,00:1$ foram usados para estabelecer a estrutura do mapa. *Loci* de marcadores co-dominantes que não mostraram distorção na segregação em $P > 0,01$ foram usados nesta análise. Em um segundo passo, o LOD foi diminuído para que novos marcadores fossem incluídos em cada grupo. O comando “try” foi, então, usado para determinar a exata posição dos novos marcadores em cada grupo. Em um passo seguinte, marcadores distorcidos e marcadores segregando 3:1

(marcadores RGA-display) foram incluídos usando o comando “group”. A ordem dos novos marcadores foi novamente confirmada com os comandos “first order”, “compare”, e/ou “ripple”. As frações de recombinação foram convertidas em distância de mapa em centimorgans (cM) usando a função “Kosambi’s mapping”. Os tamanhos de genoma no mapa foram estimados de acordo com Hulbert et al., (1988) levando em consideração apenas a comparação dos pares de marcadores colocados na estrutura do mapa.

4- Resultados

Extração de DNA e qualidade das amostras

O protocolo de extração de DNA desenvolvido por Doyle et al. (1991) e modificado neste trabalho atendeu às necessidades requeridas para os experimentos. Entretanto, foi observado que nos meses frios e secos, a qualidade do DNA extraído é muito menor (menor rendimento e maior quantidade de polissacarídeos) do que a daqueles extraídos em meses quentes e úmidos (dados não apresentados).

Para os experimentos de Southern-blot, uma quantidade muito grande de DNA foi requerida (cerca de 15 µg de DNA de cada planta para cada membrana). Para isso, cada planta teve seu DNA extraído várias vezes para se conseguir a quantidade necessária. A degradação do DNA foi um problema encontrado freqüentemente neste trabalho, mostrando que devem existir compostos que atuam nesse processo nas folhas de *Arachis* spp. Outro problema encontrado foi a viscosidade do DNA em concentrações acima de 1 µg/ml. Este excesso de polissacarídeos impediu a migração adequada de algumas amostras,

A Figura 4 mostra um gel de agarose 0,8% contendo amostras de DNA genômico de *Arachis* sp. sendo quantificadas para uso em Southern-blot e RGA-display. O marcador High Mass (Promega) foi utilizado como controle de quantidade de DNA das amostras no gel.

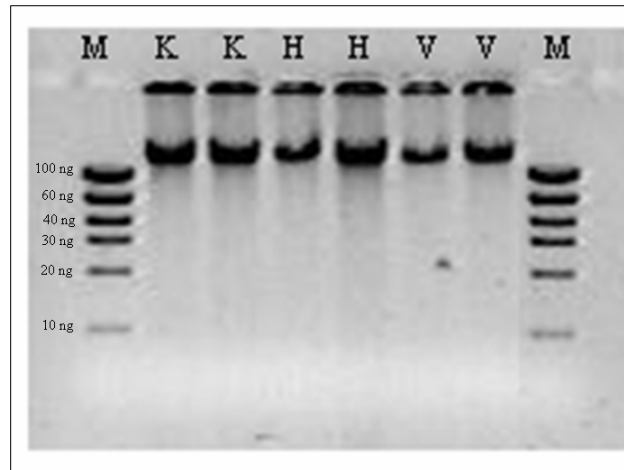


Figura 4 Gel de agarose contendo DNA genômico de *A. duranensis* K7988 (K), planta híbrida (H) e *A. stenosperma*V10309 (V). As amostras estão em duplicata. Foram aplicados 2 μ l de cada amostra no gel. O marcador de massa - High Mass Ladder (Promega) foi utilizado como parâmetro de quantificação (2 μ l).

Southern-blot

Os DNAs utilizados para a confecção das membranas de Southern-blot foram digeridos com duas enzimas de restrição: *Eco*RI ou *Hind*III. Para cada μ g de DNA foram necessárias 10 unidades de enzimas por um período de 8 h ou mais. Em alguns casos, o dobro da quantidade da enzima precisou ser utilizado.

A Figura 5 representa um gel de agarose 0,8% contendo amostras de *A. duranensis* K7988, *A. stenosperma* V10309 e da planta híbrida digeridas com *EcoRI* e *HindIII*. O marcador 1 KB (Promega) foi utilizado como controle do tamanho das bandas no gel.

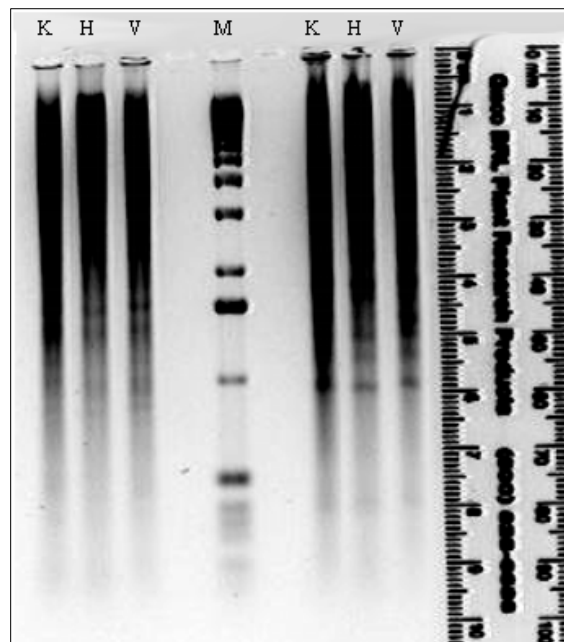


Figura 5 gel de agarose contendo digestões das amostras de *A. duranensis* K7988 (K), da planta (H) híbrida e de *A. stenosperma* V10309 (S) com *EcoRI* (amostras da esquerda) e *HindIII* (amostras da direita). “M” representa o marcador de peso molecular 1 Kb Ladder (Promega). A régua fluorescente colocada ao lado do gel facilita a visualização do tamanho dos fragmentos digeridos.

Foram feitas, no total, dezessete membranas contendo somente os parentais e dezoito membranas contendo a população F_2 para mapeamento. Primeiramente, cada membrana contendo os parentais (digeridos com uma das duas enzimas) foi hibridizada com uma das nove sondas (Tab. 1) diferentes para a seleção de quais seriam polimórficas. Das sondas testadas, todas se mostraram polimórficas para os parentais, sendo elas: S5-A-375, S5-A-384, S1-A-36, S1-A-37, C8-V-434, S4-A-164, S1S2-A-152, T-A-44 e S5C8-AXY-370. A Figura 6 mostra uma auto-radiografia do Southern-blot contendo as plantas *A. duranensis*, *A. stenosperma* e a planta híbrida digeridas com a enzima *EcoRI* (amostras da esquerda) e com a enzima *HindIII* (amostras da direita) hibridizadas com a sonda S1-A-36. As amostras

digeridas com a enzima *EcoRI* neste experimento produziram um padrão co-dominante de bandas, onde se pode observar o genótipo dos parentais (homozigoto) e da planta híbrida (heterozigoto).

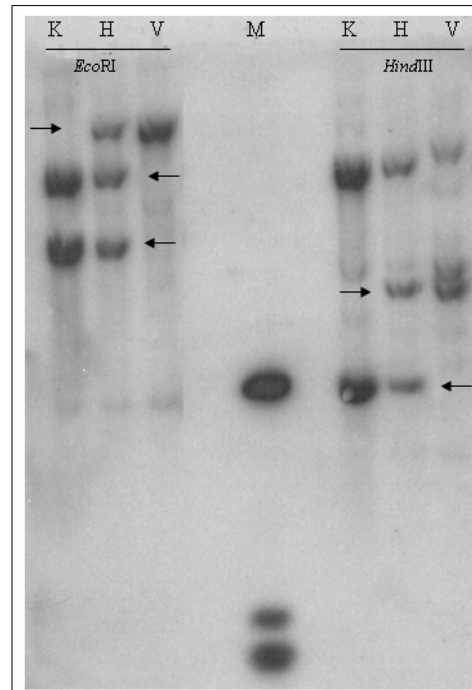


Figura 6 Auto-radiografia de Southern-blot utilizando o RGA S1-A-36 como sonda. Para este teste foram utilizados os parentais *A. duranensis* (K), a planta híbrida (H) e *A. stenosperma* (V). As amostras da esquerda foram digeridas com a enzima de restrição *EcoRI* e as quatro amostras da direita, com a enzima *HindIII*. A letra M representa 1 KB Ladder (Promega). As setas mostram as bandas polimórficas entre os parentais K7988 e V10309. A planta híbrida sempre possui todas as bandas. Pode-se observar o padrão de co-dominância das bandas polimórficas digeridas com *EcoRI* e o padrão dominante das amostras digeridas com *HindIII*.

A Tabela 4 mostra um resumo dos resultados dos testes aplicados nos parentais. A enzima *EcoRI* produziu 12 fragmentos polimórficos, quatro na planta *A. duranensis* e oito na planta *A. stenosperma*, enquanto a enzima *HindIII* produziu 24 fragmentos polimórficos (sete com a planta *A. duranensis* e 17 com a planta *A. stenosperma*) representando o dobro de marcadores. O RGA que produziu mais polimorfismo foi o S1-A-37, onde 12 marcadores poderão ser genotipados na população F_2 entre as duas enzimas. O parental *A. stenosperma* produziu 25

bandas polimórficas no total, enquanto o *A. duranensis* produziu apenas 11 bandas polimórficas.

Tabela 4 RGAs isolados por Bertoli et al. 2003 utilizados como sonda neste trabalho. A tabela representa o número de bandas polimórficas e o número de *loci* produzidos por cada planta digerida com as enzimas *EcoRI* ou *HindIII*.

RGAs	<i>EcoRI</i>				<i>HindIII</i>				Total ^c	
	<i>A. duranensis</i>		<i>A. stenosperma</i>		<i>A. duranensis</i>		<i>A. stenosperma</i>		<i>loci</i>	BP
	<i>loci</i> ^a	BP ^b	<i>loci</i>	BP	<i>loci</i>	BP	<i>loci</i>	BP		
S5-A-375	13	0	13	0	15	0	18	3	59	3
S5-A-384	1	0	1	0	3	1	1	0	6	1
S5C8-AXY-370	5	0	6	1	4	1	3	0	18	2
S1-A-36	2	2	1	1	3	1	3	1	9	5
S1-A-37	2	2	5	3	3	1	11	8	21	12
C8-V-434	1	0	2	1	5	1	4	0	12	2
S4-A-164	2	1	3	1	4	1	4	1	13	4
S1S2-A-152	8	0	8	0	11	1	13	4	40	5
T-A-44	3	1	3	1	2	0	2	0	10	2
Total	37	4	42	8	50	7	59	17	188	36

* a representa o número de *loci* de cada RGA no genoma das plantas *A. duranensis* e *A. stenosperma*. b é o número de bandas polimórficas encontrada na hibridização das plantas com cada um dos RGAs identificados. C é o número total de *loci* ou banda polimórfica para cada RGA usado como sonda nas duas plantas digeridas com as duas enzimas já citadas.

Nos testes preliminares realizados nos parentais também foi possível a contagem do número de *loci* de cada RGA no genoma de cada planta. A Tabela 4 também representa o número de cópias dos RGAs utilizados como sonda nos parentais (*A. duranensis* e *A. stenosperma*). O RGA S5-A-375 possui 13 cópias no genoma da *A. duranensis* digerido com a enzima *EcoRI*, enquanto que no genoma de *A. stenosperma*, clivado com *EcoRI*, só existe 1 cópia do RGA S5-A-384.

Após todos os testes realizados e de posse das informações obtidas sobre as sondas e enzimas com maior potencial de polimorfismo, os Southern-blots foram então realizados com a população F₂ de mapeamento. Noventa e três plantas foram selecionadas e clivadas com uma das enzimas utilizadas neste trabalho (*EcoRI* ou *HindIII*). Cada membrana recebeu DNA digerido de vinte plantas juntamente com os parentais e a planta híbrida, ou seja, para cada

sonda, foram feitas cinco membranas. Três sondas foram genotipadas na população completa e duas sondas foram genotipadas em 20 plantas da população F₂ (Tabela 5).

Tabela 5 RGAs usados como sonda para genotipagem da população F₂ e números de indivíduos correspondentes a cada experimento.

Sonda	Enzima	
	<i>EcoRI</i>	<i>HindIII</i>
S1-A-36	93 plantas	-
S1-A-37	93 plantas	-
S1S2-A-152	-	93 plantas
S4-A-164	-	20 plantas
T-A-44	20 plantas	-

As membranas utilizadas para a genotipagem foram reaproveitadas com sondas diferentes após decaimento da sonda anterior. A Figura 7 mostra um Southern-blot com quinze plantas da população F₂ digeridas com *EcoRI* e hibridizadas com a sonda S1-A-36. Apesar do teste inicial com os parentais hibridizados com essa mesma sonda ter mostrado um padrão de co-dominância entre as bandas polimórficas, a genotipagem revelou mais bandas que não haviam aparecido no teste inicial, tornando difícil a identificação dos genótipos. Por exemplo, a planta 11 mostrou o mesmo padrão de bandas do parental *A. stenosperma*, enquanto a planta 24 possui o mesmo padrão do parental *A. duranensis* (dados não mostrados), mas a planta 14 não mostrou semelhança com nenhum dos dois parentais e tampouco com a planta híbrida. As bandas observadas segregam independentemente, sendo muito difícil a análise como sendo co-dominantes, por isso, neste caso, foram genotipadas como sendo dominantes, ou seja, cada banda representa um marcador.

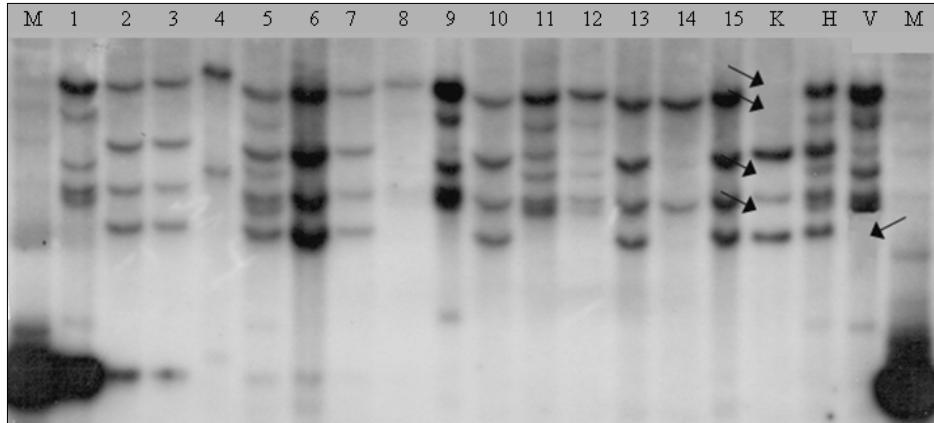


Figura 7 Auto-radiografia de Southern-blot contendo 15 plantas da população F_2 de mapeamento de *Arachis* representadas pelos números. As amostras foram digeridas com a enzima de restrição *EcoRI* e hibridizadas com a sonda RGA S1-A-36. “M” representa o marcador 1 Kb Ladder (Promega), “K” representa a planta *A. duranensis*, “H” representa a planta híbrida e “V” representa a planta *A. stenosperma*. As setas mostram as bandas polimórficas entre os parentais e que segregam nas plantas da população F_2 .

As bandas polimórficas que segregaram na população F_2 foram tratadas como marcadores e o resultado da genotipagem foi adicionado à tabelas onde a presença ou ausência de cada banda foi avaliada. Nove marcadores moleculares foram desenvolvidos através dessa técnica (quatro com o RGA S1-A-36, um com o RGA S1-A-37, dois com o RGA S1S2-A-152 e um com o RGA S4-A-164) e caracterizados na população F_2 oriunda do cruzamento das duas espécies silvestres utilizadas no presente trabalho (*A. duranensis* e *A. stenosperma*).

RGA-Display

A preparação dos DNAs é um passo muito importante na técnica RGA-display. DNAs com baixa qualidade podem fornecer resultados distorcidos, criando bandas que não existem ou não amplificando fragmentos importantes. Para a digestão dos DNAs, duas maneiras foram testadas: digestão dupla simultânea e digestão com uma enzima de cada vez (primeiro a de corte raro- *PstI* e depois a de corte freqüente – *MseI*). Nenhuma diferença foi visualizada nos

produtos finais obtidos da preparação dos DNAs (dados não apresentados), por isso, para esse trabalho, todos os DNAs foram digeridos simultaneamente com as duas enzimas.

A ligação dos fragmentos digeridos aos adaptadores foi realizada a 37°C, para que as enzimas adicionadas à reação de ligação pudessem cortar os fragmentos religados em si mesmos. Essa estratégia permite um maior rendimento de fragmentos potencialmente amplificáveis pelos *Primers* complementares aos adaptadores.

A reação de PCR de pré-amplificação é seletiva e amplifica preferencialmente os fragmentos que possuem os adaptadores diferentes ligados às duas extremidades (*Pst-Mse*), não amplificando os fragmentos *Mse-Mse* ou *Pst-Pst*. Essa foi uma modificação da técnica original, que nessa primeira etapa de PCR utiliza um programa similar ao usado para AFLP onde há um decréscimo de 0,7°C por ciclo na temperatura de ligação do *primer*. Uma diluição de 1:20 destes produtos de pré-amplificados foi utilizada nas reações de amplificação.

Nas reações de amplificação dos fragmentos, de 140 combinações possíveis de *primers* (sete *primers* relativos ao motivo Kinase-2 e vinte relativos ao adaptador ligado ao sítio de *MseI*) 130 foram testadas contra os parentais *A. duranensis* e *A. stenosperma*. Dessas combinações, 100 apresentaram algum grau de polimorfismo entre os mesmos. As combinações que obtiveram pelo menos uma banda polimórfica foram selecionadas para serem genotipadas na população F₂ de 93 indivíduos (Tabela 7).

As amplificações foram testadas com três tipos de *Taq* DNA polimerase: Hot Star *Taq* (Quiagen), *Taq* comum (PhT) e a Platinum *Taq* (Invitrogen). As enzimas Hot Star *Taq* e Platinum mostraram alta fidelidade nos resultados, apresentando o mesmo perfil de bandas. Já a enzima comum, amplificou uma grande quantidade de fragmentos inespecíficos (Figura 8). Esses fragmentos foram seqüenciados e não apresentaram homologia com nenhuma seqüência já depositada no GenBank.

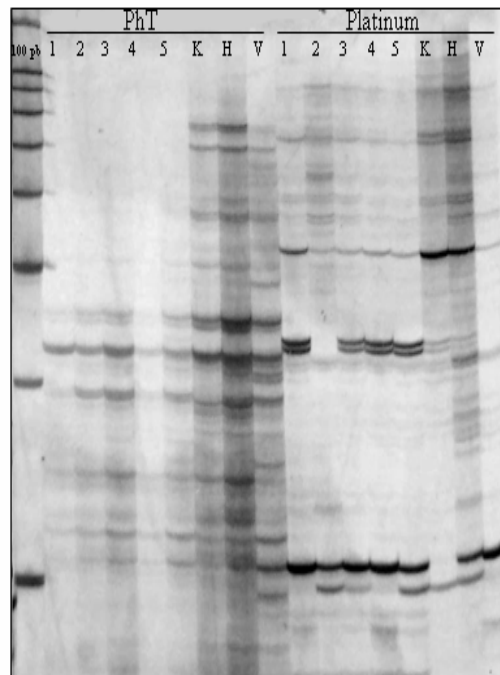


Figura 8 Gel de Poliacrilamida corado com Nitrato de Prata contendo amostras amplificadas com os *primers* P2-414 e Mse-CTC. As primeiras oito amostras foram amplificadas com a *Taq* PhT e as oito últimas amostras foram amplificadas com a *Taq* Platinum Invitrogen. O marcador utilizado foi o 100 Pb da Invitrogen. Os números representam os indivíduos 1 a 5 da população F₂. K é o parental *A. duranensis*, V, o parental *A. stenosperma* e H, a planta híbrida.

A amplificação feita com a *Taq* comum, em geral, apresentou um número muito maior de bandas polimórficas. O perfil encontrado foi o parecido com amplificações de AFLP. As bandas polimórficas provenientes da amplificação com a *Taq* comum não apresentaram homologia com RGAs quando seqüenciadas. Por estes motivos, todas as reações posteriores foram feitas com *Taq* Hot Star ou *Taq* Platinum.

Outro teste realizado foi com as temperaturas de ligação de cada *primer*. Temperaturas muito altas (superiores a 55°C) produziram um padrão de bandas muito restrito, havendo pouca amplificação de fragmentos. Reações com temperaturas de ligação abaixo de 46°C geraram muitos produtos inespecíficos. Contudo, os *primers* degenerados precisam de certa inespecificidade para poder se ligar ao DNA e o decréscimo da temperatura nesses casos

auxiliaram a amplificação. A Tabela 6 contém um resumo das temperaturas de ligação de cada *primer* relativo ao motivo Kinase-2 do Domínio NBS.

Tabela 6 *Primers* relativos ao adaptador de *MseI* e Domínio NBS e suas melhores temperaturas de ligação.

<i>Primer</i>	TM^a	Observ.
P- <i>MseI</i>	42-51	Sem base seletiva ^b
P2-391	49	Degenerado
P2-414	49	Degenerado
P2-19	53	Específico
P2-12	52	Específico
P2-378	52	Específico
NBS 5	50	Degenerado
P2-16	54	Específico

a: Temperatura de ligação dos *Primers*. b: Aqui mostrado sem bases seletivas ao final da seqüência.

Os *primers* desenhados para reconhecer o sítio de restrição da enzima *MseI* possuem bases seletivas ao final de sua seqüência afim de restringir o número de fragmentos produzidos pela PCR. Uma, duas ou três bases foram adicionadas ao final destes *primers*. Testes com esses *primers* contendo essas bases seletivas foram feitos nos parentais e quase todas as combinações produziram polimorfismo entre as espécies estudadas (Tabela 7). A Figura 9 contém um gel de poliacrilamida 4% corado com Nitrato de Prata com amostras de *A. duranensis* e *A. stenosperma* amplificadas com *primers* P2-378 e *Mse* com uma, duas ou três bases seletivas. Pode-se observar que as reações contendo o *primer Mse* com apenas uma base seletiva possuem mais bandas (43 bandas) do que as amplificadas com duas bases seletivas (26), que por sua vez, possuem mais bandas que as amplificadas com três bases seletivas (12 fragmentos). A melhor reação para ser analisada, baseando-se nesse teste, é a amplificada com duas bases seletivas, pois gerou um número satisfatório de fragmentos sem formar um rastro como acontece em alguns casos (dados não apresentados).

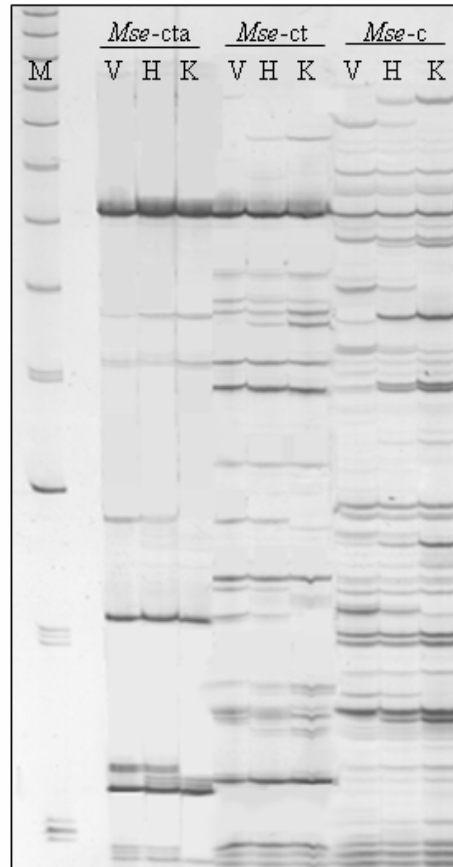


Figura 9 Gel de poliacrilamida 4% corado com Nitrato de Prata contendo amostras de RGA-Display amplificadas com o *primer* P2-378 relativo ao motivo Kinase-2 no Domínio NBS no sentido direto e outro relativo ao sítio de restrição da enzima *MseI* com três, duas ou uma base seletivas. As amostras da esquerda foram amplificadas com o *primer* Mse-cta (3 bases seletivas), as do centro do gel foram amplificadas com o *primer* Mse-ct (2 bases seletivas) e as da direita com o *primer* Mse-c (1 base seletiva). V representa *A. stenosperma*, H representa a planta híbrida e K representa a planta *A. duranensis*. M é o marcador 1 kb plus (Promega).

O padrão mostrado na Figura 8, onde o *primer* com uma base seletiva produziu mais bandas do que o *primer* com três bases seletivas não se repetiu em todas as combinações de *primers*. Em alguns casos, o número de bandas produzido foi tão grande que formou um rastro no gel, não sendo possível a visualização adequada dos fragmentos. Em contraste, algumas combinações usando o *primer* com uma base seletiva quase não houve amplificação dos fragmentos (dados não apresentados).

Após os testes iniciais de temperatura de ligação, *Taq* polimerase e número de bases seletivas, foram feitos testes (*Screenings*) para a detecção de melhores combinações de

primers para serem utilizados na população F₂ de mapeamento. De 140 combinações possíveis, 130 foram testadas nos parentais e na planta híbrida. As reações foram separadas (resolvidas) em gel de poliacrilamida, mas apenas 100 mostraram polimorfismo entre os parentais. A Figura 10 contém um gel com teste de polimorfismo apenas com os parentais amplificados com seis combinações de *primers*: P2-19 (no sentido direto) e M-cat, M-cag, M-cac, M-caa, M-ctt ou M-ctg (no sentido reverso). Todos os *primers* *Mse* neste gel possuem três bases seletivas. Pode ser observado, por exemplo, que na combinação P2-19 e M-cta houve a amplificação de 14 bandas polimórficas entre os parentais, embora esse resultado não tenha sido mantido quando esses mesmos *primers* foram utilizados para a amplificação da população F₂, onde apenas 7 bandas polimórficas puderam ser observadas em virtude da falta de sensibilidade da coloração com Nitrato de Prata (Tabela 7).

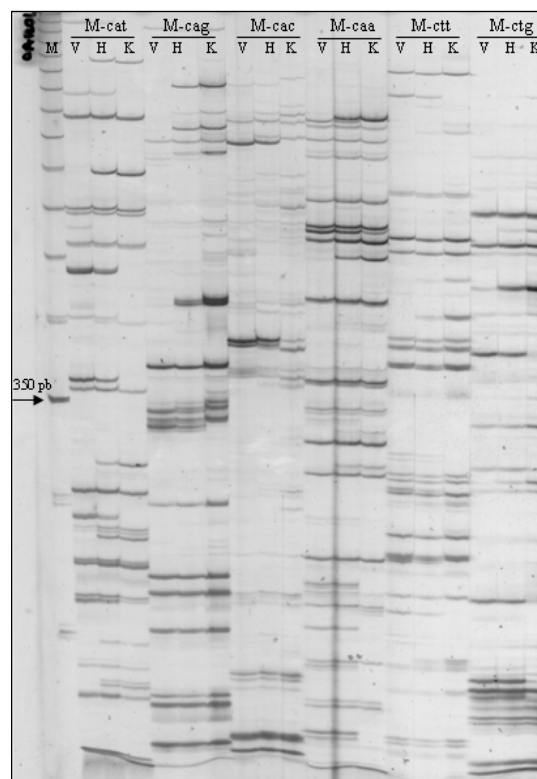


Figura 10 Gel de poliacrilamida 4% corado com Nitrato de Prata contendo amostras de RGA-display amplificadas com o *primer* P2-19 no sentido direto combinado com um dos seis *primers* *Mse* com três bases seletivas (no sentido reverso). Da esquerda para a direita: M-cat, M-cag, M-cac, M-caa, M-ctt e M-ctg combinados com o *primer* P2-19. V representa a planta *A. stenosperma*, H representa a

planta híbrida e K representa a planta *A. duranensis*. M representa o marcador “50 pb DNA Ladder” (Invitrogen). A seta indica o fragmento de aproximadamente 350 pb.

Neste gel apresentado (P2-19 e Mse – com 3 bases seletivas Figura 10), cada combinação de *primer* produziu cerca de 40 fragmentos, o que representa três vezes mais bandas do que nas combinações com o *primer* P2-414 (que possui cerca de 15 bandas com *primers* contendo 3 bases seletivas). O *primer* P2-19 é específico, enquanto que o P2-414 é degenerado. Essa degeneração deve estar restringindo o número de fragmentos produzidos, uma vez que precisa de temperaturas mais baixas para se ligar ao DNA. A *Taq* Platinum também deve estar desfavorecendo a ligação desse *primer* à fita de DNA.

Nas 100 combinações de *primer* testadas nos parentais, foram amplificados 3.095 fragmentos, sendo 293 (9,5%) polimórficos entre os parentais. Para a genotipagem da população F₂, foram escolhidas as 24 melhores combinações de *primers* que produziram polimorfismo entre os parentais. A Figura 10 mostra um gel de poliacrilamida corado com Nitrato de Prata contendo RGA-Display na população F₂. As amostras foram amplificadas com os *primers* P2-414 e M-c e a enzima *Taq* Platinum (Invitrogen). Pode-se observar que oito bandas polimórficas foram visualizadas neste gel. O resultado da genotipagem foi adicionado a uma tabela contendo a mesma informação de outras combinações de *primers*. As amostras amplificadas com o DNA dos parentais mostradas nesta figura, não forneceram exatamente o mesmo resultado do encontrado na população F₂. As bandas principais são as mesmas, mas nos parentais houve uma qualidade menor de amplificação dos fragmentos mais fracos neste caso.

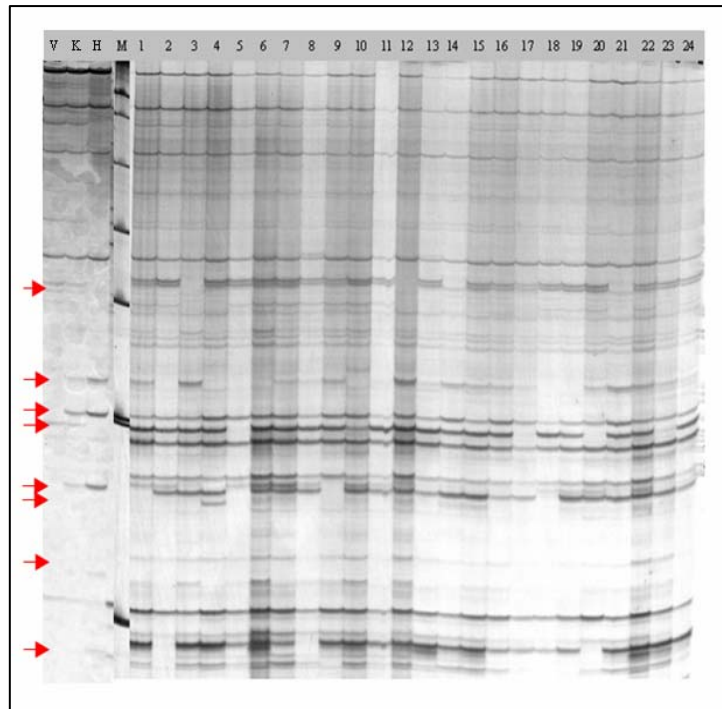


Figura 11 Gel de poliacrilamida corado com Nitrato de prata contendo amostras de RGA-Display com parte da população F_2 de *Arachis*. As amostras foram amplificadas com os *primers* P2-414 (degenerado) e M-c. “M” representa o marcador 1Kb Plus (Promega). A planta *A. duranensis* está representada por “K”, *A. stenosperma* por “V” e a planta híbrida por “H”. Os números na parte superior do gel representam as 24 primeiras plantas da população F_2 genotipadas com essa combinação de *primers*. As setas indicam as bandas polimórficas entre os parentais e que segregaram na população F_2 .

No total, 765 fragmentos (polimórficos e monomórficos) foram amplificados com as 24 combinações de *primers* usadas na população F_2 . Destes, 143 foram considerados polimórficos entre os parentais, mas somente 139 marcadores puderam ser analisados devido à falta de consistência de algumas bandas. Dentre as combinações de *primers* genotipadas, houve uma taxa de polimorfismo de 18,7 %, representando uma média de seis bandas polimórficas por combinação de *primer*.

Tabela 7 Combinações de *primers* utilizadas para genotipagem da população F₂, Número de bandas amplificadas, polimórficas, seqüenciadas e RGAs produzidas por cada combinação testada.

Combinação ^a	Número de bandas				% de	% de
	Amplificadas ^b	Polimórficas ^c	Seqüenciadas ^d	RGAs ^e	Polimorfismo	RGAs*
P2-12 + Mct (9)	13	3	3	3	23,1	100
P2-12 + Mcta (1)	8	1	1	0	12,5	0
P2-16 + Mc (28)	42	9	9	1	21,4	11
P2-16 + Mcaa (26)	17	2	2	2	11,8	100
P2-16 + Mctc (27)	35	7	7	2	20	29
P2-19 + Mcac (6)	45	8	0	0	17,8	0
P2-19 + Mcta (12)	22	7	7	1	31,8	14
P2-19 + Mctc (5)	56	10	0	0	17,9	0
P2-19 + Mgag (7)	49	5	0	0	10,2	0
P2-378 + Mct (20)	43	9	9	0	20,9	0
P2-378 + Mcta (23)	39	6	4	0	15,4	0
P2-378 + Mctc (11)	25	9	8	0	36	0
P2-378 + Mctt (24)	45	15	15	3	33,3	20
P2-391 + Mcta (13)	39	4	4	0	10,3	0
P2-391 + Mgag (16)	31	2	2	0	6,5	0
P2-414 + Mc (17)	40	8	6	0	20	0
P2-414 + Mcta (15)	28	3	3	0	10,7	0
P2-414 + Mctc (10)	14	4	4	1	28,6	25
P2-414 + Mctt (19)	28	3	3	0	10,7	0
NBS5 + Mc (21)	55	9	9	0	16,4	0
NBS5 + Mctc (14)	19	4	4	2	21,1	50
NBS5 + Mgag (22)	36	3	2	0	8,3	0
NBS5 + Mgct (25)	19	2	2	0	10,5	0
NBS5 + Mggc (8)	17	10	5	4	58,8	40
Total	765	143	109	19	18,7	13
Média por combinação	32	6	5	-	-	-

^a refere-se aos dois *primers* utilizados para a amplificação dos fragmentos referentes ao teste realizados nos parentais. O número entre parênteses ao lado das combinações de *primers* diz respeito ao número da combinação utilizada. ^b Número de bandas amplificadas por cada combinação de *primers*. Inclui bandas polimórficas e monomórficas. ^c Número de bandas que diferem entre os parentais e indivíduos da população F₂. ^d refere-se ao número de bandas polimórficas genotipadas que foram seqüenciadas e caracterizadas. ^e Número de bandas seqüenciadas que tiveram homologia com RGAs no GenBank. A comparação com o banco de dados foi feita através de BLASTX e BLASTN. * Porcentagem de RGAs em relação às bandas polimórficas seqüenciadas.

Após a genotipagem da população F₂ com cada combinação de *primer*, as bandas polimórficas foram retiradas do gel com o auxílio de uma lâmina. As bandas foram colocadas em H₂O Milli-Q estéril e aquecidas a 95°C para a liberação do DNA presente nelas. Os fragmentos eluídos foram re-amplificados utilizando a mesma reação de amplificação mencionada na seção Material e Métodos. O produto da re-amplificação foi separado em gel de agarose 0,8% para a confirmação de que apenas uma banda havia sido amplificada. No caso de produtos com mais de um fragmento amplificado, a banda de tamanho certo foi

extraída do gel e purificada conforme já descrito. Os produtos foram levados a seqüenciamento com os mesmos *primers* usados para a sua amplificação (um *primer* para cada reação de seqüenciamento de cada banda). A Figura 12 mostra um gel de agarose contendo amostras re-amplificadas para seqüenciamento. Neste gel, apenas um fragmento visível de cada banda retirada do gel foi amplificado.

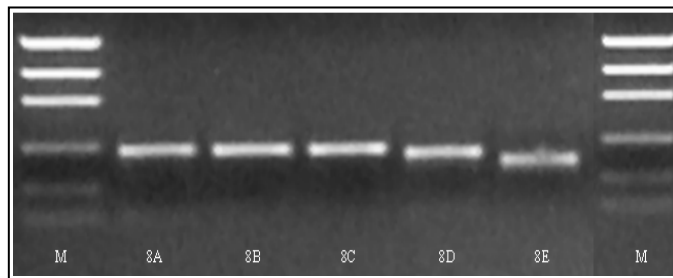


Figura 12 Gel de agarose 0,8% corado com brometo de etídio contendo bandas retiradas do gel de poliacrilamida e re-amplificadas por PCR. Os fragmentos possuem aproximadamente 300 Pb. “M” representa o marcador Low Mass Ladder (2 μ l, Invitrogen – a banda mais alta possui 100 ng de DNA). As amostras 8A, 8B, 8C, 8D, 8E referem-se a combinação de *primers* número 8 (NBS5 e M-ggc) acompanhada da letra da banda.

Seqüenciamento e análise das bandas polimórficas

As seqüências obtidas foram processadas com o auxílio do programa Staden Package (Staden et al., 2003). No total foram seqüenciadas 109 bandas polimórficas. Se forem somadas as repetições e os seqüenciamento nos dois sentidos, 403 amostras foram seqüenciadas neste trabalho. A Figura 13 é um exemplo do processamento das amostras seqüenciadas usando o Staden Package.

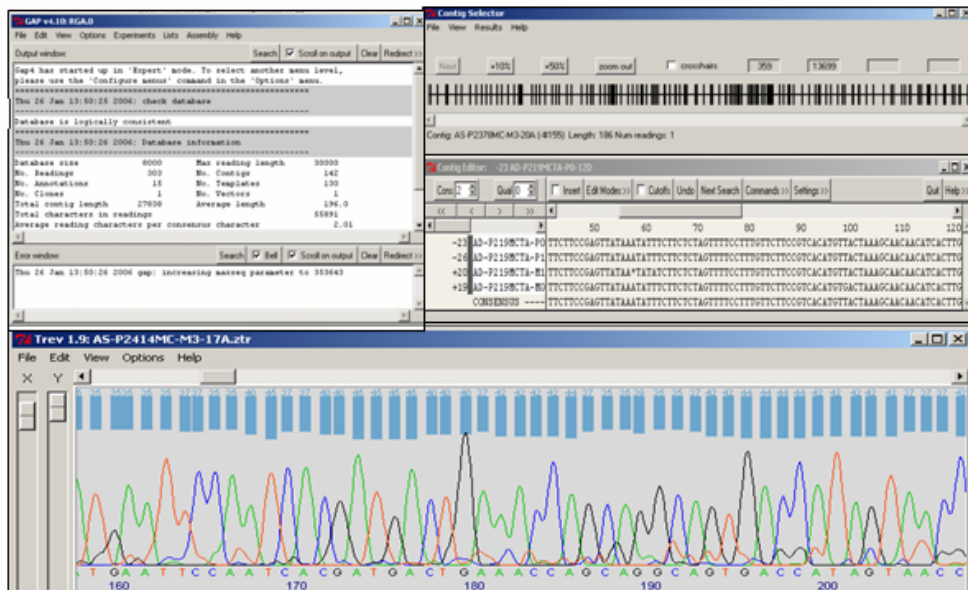


Figura 13 Processamento de seqüências utilizando a ferramenta Staden Package (Staden et al., 2003).

Após a junção de todas as leituras de seqüência de cada amostra, 100 seqüências completas foram utilizadas para as análises, ou seja, de 109 bandas seqüenciadas, 100 tiveram qualidade na seqüência suficiente para serem utilizadas. Primeiro, as seqüências de nucleotídeos foram comparadas com o GenBank utilizando o BLASTX (Altschul et al., 1997). Essas seqüências também foram comparadas com um banco local de genes de *Arabidopsis thaliana*. As bandas polimórficas seqüenciadas neste trabalho possuem entre 80 e 400 pb. A Tabela 8 é um resumo dos principais resultados encontrados do seqüenciamento das bandas polimórficas não-RGAs retiradas do gel de RGA-display com as 24 combinações de *primers*. As amostras não apresentadas não tiveram nenhum “hit” significativo com outras seqüências depositadas no GenBank.

Tabela 8 Bandas RGA-display seqüenciadas e suas principais homologias com seqüências conhecidas no GenBank.

Primer RGA	Marcador ^a	Hit ^b	E-value ^c
NBS 5	AD-NBSMGGC-8B	Disease resistance protein (NBS-LRR cla...	4,00E-22
NBS 5	AS-NBSMGGC-8C	Disease resistance protein (NBS-LRR cla...	7,00E-26
NBS 5	AD-NBSMGGC-8D	Disease resistance protein (CC-NBS clas...	3,00E-11
NBS 5	AS-NBSMGGC-8J	Disease resistance protein (CC-NBS clas...	0.003
NBS 5	AD-NBSMCTC-14A	Disease resistance protein (CC-NBS clas...	0.005
NBS 5	AS-NBSMCTC-14D	Disease resistance protein (CC-NBS-LRR ...	0.002
NBS 5	AS-NBSMCTT-24A	Disease resistance protein (NBS-LRR cla...	2,00E-23
NBS 5	AS-NBSMCTT-24E	Disease resistance protein (CC-NBS clas...	7,00E-14
NBS 5	AD-NBSMCTT-24F	Putative late blight resistance protein [Solanum...	0.029
NBS 5	AS-NBSMCTC-14B	At2g32290.1 68415.m03947 beta-amylase, putative / 1,4-alpha-D-gl...	5,00E-36
NBS 5	AD-NBSMCTC-14C	At2g32290.1 68415.m03947 beta-amylase, putative / 1,4-alpha-D-gl...	3,00E-41
NBS 5	AD-NBSMGGAG-22C	GI:4567295 putative retroelement pol polyprotein [Arabidopsis thaliana].	3,00E-18
NBS 5	AD-NBSMCTT-24B	gi 54644620 gb EAL33361.1 GA11129-PA [Drosophila pseudoobscura]	3,90E-01
NBS 5	AS-NBSMCTT-24H *	At1g52540.1 68414.m05931 protein kinase, putative contains prote...	4,00E-36
NBS 5	AD-NBSMCTT-24N	At3g16190.1 68416.m02044 isochorismatase hydrolase family protei...	0.020
NBS 5	AS-NBSMGGC-8E	gi 55770171 ref XP_550075.1 hypothetical protein [Oryza sati...	0.001
P2-12	AD-P212MCT-9A	Disease resistance protein (TIR-NBS-LRR...	0.003
P2-12	AD-P212MCT-9B	Disease resistance protein (TIR-NBS-LRR...	0.001
P2-12	AD-P212MCT-9C	Disease resistance protein (TIR-NBS-LRR...	0.001
P2-16	AS-P216MCAA-26A	Disease resistance protein (TIR-NBS-LRR...	4,00E-07
P2-16	AD-P216MCAA-26B	Disease resistance protein (TIR-NBS-LRR...	3,00E-10
P2-16	AS-P216MCTC-27A	Disease resistance protein (TIR-NBS-LRR...	6,00E-14
P2-16	AD-P216MCTC-27B	Disease resistance protein (TIR-NBS-LRR...	5,00E-22
P2-16	AD-P216MC-28G	Disease resistance protein (TIR-NBS-LRR...	2,00E-04
P2-16	AS-P216MCTC-27D	gi 15221080 ref NP_175246.1 hydrolase/ protein serine/threon...	4,00E-07
P2-19	AS-P219MCTA-12F	Disease resistance protein (TIR-NBS-LRR...	1,00E-06
P2-19	AD-P219MCTA-12B	At1g64260.1 68414.m07281 zinc finger protein-related contains Pf...	3,00E-10
P2-378	AS-P2378MCTC-11B	GI:49533772 putative polyprotein [Solanum demissum].	4,00E-57
P2-378	AS-P2378MCTC-11C	GI:7267154 putative transposon protein [Arabidopsis thaliana].	1,00E-09
P2-378	AD-P2378MCTC-11D	gi 2995405 emb CAA73042.1 polyprotein [Ananas comosus]	6,00E-31
P2-378	AD-P2378MC-20B *	At4g23160.1 68417.m03342 protein kinase family protein contains ...	5,00E-04
P2-378	AA-P2378MC-20D	GI:49533772 putative polyprotein [Solanum demissum].	2,00E-27
P2-378	AS-P2378MC-20H	gi 15724009 gb AAL06413.1 reverse transcriptase [Pisum sativum]	9,00E-11
P2-378	AD-P2378MCTA-23A*	At4g23160.1 68417.m03342 protein kinase family protein contains ...	3,00E-09
P2-378	AS-P2378MCTA-23E	gi 15724009 reverse transcriptase [Pisum sativum]	2,00E-11
P2-391	AS-P2391MCTA-13C *	At1g21230.1 68414.m02653 wall-associated kinase, putative simila...	6,00E-07
P2-391	AD-P2391MCTA-13D *	At1g21230.1 68414.m02653 wall-associated kinase, putative simila...	2,00E-11
P2-391	AD-P2391MGAG-16B	At1g72520.1 68414.m08386 lipoxygenase, putative similar to lipox...	2,00E-10
P2-414	AS-P2414MCTC-10A	Disease resistance protein, RPP13-like ...	3,00E-24
P2-414	AD-P2414MCTC-10C	GI:16118447 heat shock transcription factor [Phaseolus acutifolius].	3,00E-21
P2-414	AD-P2414MC-17B	gi 6691193 gb AAF24531.1 F7F22.17 [Arabidopsis thaliana]	4,00E-36
P2-414	AS-P2414MC-17D	GI:16118447 heat shock transcription factor [Phaseolus acutifolius].	4,00E-20
P2-414	AS-P2414MCTT-19A	GI:50930529 OSJNBa0014F04.3 [Oryza sativa (japonica cultivar-group)].	0.002

* Provável proteína do tipo Kinase. ** Prováveis regiões análogas a genes de resistência. ^a Nome dado a cada marcador identificado no trabalho. O nome possui as iniciais da planta de origem (AD – *A. duranensis*; AS – *A. stenosperma*), os dois primers utilizados na amplificação por PCR (por exemplo: P212MCT – P2-12 e Mse-ct) e o número da combinação de primers associado com a letra da banda extraída do gel. ^b Principal homologia encontrada no GenBank. O programa utilizado para a comparação com o banco de dados foi o BLASTX. ^c E-Value é o número correspondente ao nível de significância. Quanto menor o valor, melhor a qualidade da homologia. O nome do marcador pode ser resumido ao número e à letra final do nome, por exemplo: 10A, 24C.

Devido ao baixo valor de significância encontrado na comparação com outras seqüências no GenBank, outro tipo de análise também foi feito, onde as seqüências foram traduzidas utilizando a ferramenta “Translate – TrEMBL” no site www.expasy.ch/tools/dna.html. A escolha da melhor tradução foi baseada na presença da seqüência do *primer*. As traduções foram submetidas ao programa de domínio público “Motif-Scan” (Falquet et al., 2002) no site www.expasy.ch/prosite onde a ferramenta HAMAP procura no banco de dados do Pfam por domínios e motivos conhecidos dentro da seqüência estudada. A Figura 14 mostra como exemplo a amostra AS-P2414MCTC-10A traduzida processada pelo “Motif-Scan”. Na Figura é possível observar a significativa homologia com o domínio NB-ARC encontrada pelo Pfam (HMMS).

Para fins comparativos, a Tabela 9 mostra um resumo dos resultados de BLASTX e BLASTN para os marcadores seqüenciados neste trabalho, onde de 100 amostras com boa qualidade de seqüenciamento, 19 tiveram homologia com RGAs, 28 tiveram homologia com outros genes, como B-amilases, Kinases, Hidrolases, Retrotransposons e outros. 53 amostras não apresentaram homologias significativas com outras seqüências já descritas (Gráfico 1).

Tabela 9 Frequência de bandas RGAs e outros genes encontrados em RGA-Display.

<i>Primer</i> RGA	Nº de Bandas Seqüenciadas	Nº de RGAs	Não Identificadas	Outros Genes	Genes
NBS 5	38	9 (24%)	20	9	β -amilase, Retrotransposons, Kinase, Hidrolase
P2-12	3	3 (100%)	0	0	-
P2-16	11	5 (45%)	5	1	Hidrolase
P2-19	8	1 (13%)	6	1	Proteína “Zinc Finger”
P2-378	19	0 (0%)	10	9	Retrotransposons, Kinases e transcriptase reversa
P2-391	4	0 (0%)	1	3	Kinases e Lipoxygenase
P2-414	17	1 (6%)	11	5	Fator de transcrição de choque térmico e prováveis proteínas de Arroz e <i>Arabidopsis</i> .
Total	100	19 (18,7%)	53	28	

Gráfico 1 Representação das principais homologies encontradas no seqüenciamento de bandas polimórficas da técnica RGA-display.

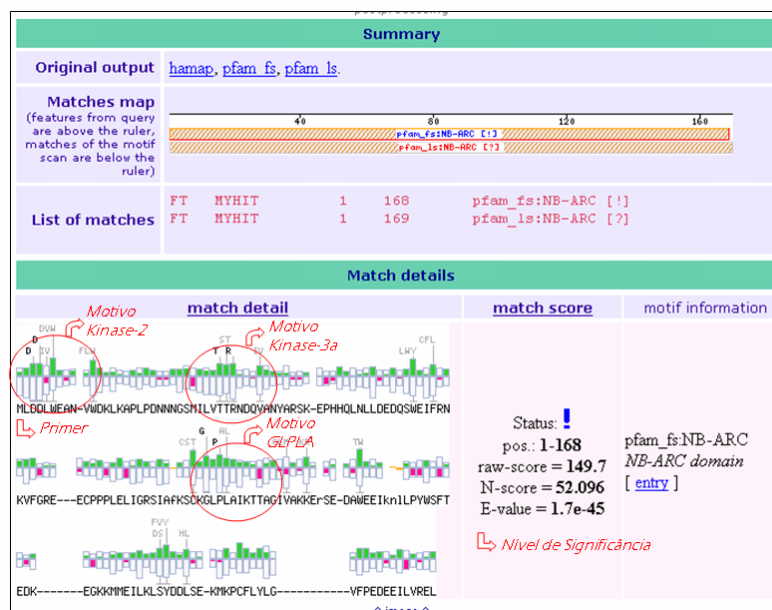
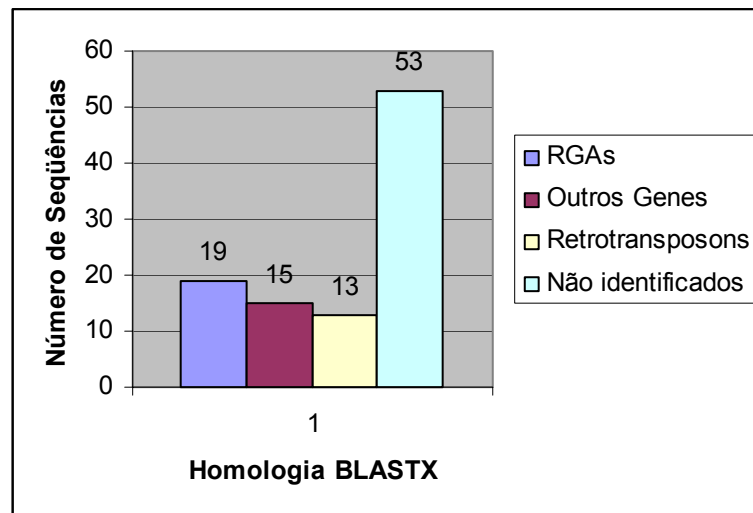


Figura 14 Imagem de processamento do provável RGA AS-P2414MCTC-10A obtida através do programa Motif-Scan, onde as ferramentas Hamap e Pfam encontraram as regiões conservadas que caracterizam o domínio NB-ARC. Os círculos realçam os motivos Kinase-2 (que representa a região amplificada pelo *primer*, ou seja, o início da seqüência), Kinase-3 e GLPLA (Site: www.expasy.ch/prosite).

O Nível de significância (E-Value) obtido pela comparação com o banco de dados do Pfam através do *Motif-Scan* de $1,7 E^{-45}$ é bem mais confiável do que o encontrado pelo BLASTX na tabela anterior (Tabela 8) que foi de $7,0 E^{-24}$. Esse tipo de análise foi feito para as 19 seqüências que tiveram homologia com RGAs ou genes de resistência no BLASTX. Algumas seqüências tiveram seus valores de significância aumentados e outras não apresentaram homologia com o domínio NB-ARC. A Tabela 10 é uma comparação das homologias encontradas no GenBank do NCBI e Pfam. Uma explicação seria que não foi conseguida a seqüência correta de aminoácido de algumas amostras, por isso nenhum motivo foi encontrado. Isso deve ter acontecido nas amostras: 26A, 26B, 27A e 28G. As amostras que obtiveram valor de significância muito baixo no BLAST podem ser falsos positivos. A semelhança encontrada deve ter sido apenas do *primer* utilizado que possui em sua seqüência a região conservada do motivo Kinase-2, como acontece nos casos das amostras 9A, 9B e 9C.

Tabela 10 Marcadores com homologia a RGAs e comparação entre os Hits encontrados pelo BLASTX e Motif-Scan (Pfam)

Primer RGA	Marcador	BLASTX ^a		Motif-Scan e EstWise ^b	
		Hits	E-Value	Hits	E-Value
P2-12	AD-P212MCT-9A	Disease resistance protein (TIR-NBS-LRR...	0.003	no hits	
P2-12	AD-P212MCT-9B	Disease resistance protein (TIR-NBS-LRR...	0.001	no hits	
P2-12	AD-P212MCT-9C	Disease resistance protein (TIR-NBS-LRR...	0.001	no hits	
P2-16	AS-P216MCAA-26A	Disease resistance protein (TIR-NBS-LRR...	4,00E-07	no hits	
P2-16	AD-P216MCAA-26B	Disease resistance protein (TIR-NBS-LRR...	3,00E-10	no hits	
P2-16	AS-P216MCTC-27A	Disease resistance protein (TIR-NBS-LRR...	6,00E-14	no hits	
P2-16	AD-P216MCTC-27B	Disease resistance protein (TIR-NBS-LRR...	5,00E-22	pfam_fs:NB-ARC Domain	3,20E-10
P2-16	AD-P216MC-28G	Disease resistance protein (TIR-NBS-LRR...	2,00E-04	no hits	
P2-19	AS-P219MCTA-12F	Disease resistance protein (TIR-NBS-LRR...	1,00E-06	pfam_fs:NB-ARC Domain	8,70E-06
P2-414	AS-P2414MCTC-10A	Disease resistance protein, RPP13-like ...	3,00E-24	pfam_fs:NB-ARC Domain	1,70E-45
NBS 5	AD-NBSMGGC-8B	Disease resistance protein (NBS-LRR cla...	4,00E-22	pfam_fs:NB-ARC Domain	1,80E-27
NBS 5	AS-NBSMGGC-8C	Disease resistance protein (NBS-LRR cla...	7,00E-26	pfam_fs:NB-ARC Domain	3,30E-34
NBS 5	AD-NBSMGGC-8D	Disease resistance protein (CC-NBS clas...	3,00E-11	pfam_fs:NB-ARC Domain	1,30E-11
NBS 5	AS-NBSMGGC-8J	Disease resistance protein (CC-NBS clas...	0.003	pfam_fs:NB-ARC Domain	1,40E-08
NBS 5	AD-NBSMCTC-14A	Disease resistance protein (CC-NBS clas...	0.005	pfam_fs:NB-ARC Domain	1,30E-08
NBS 5	AS-NBSMCTC-14D	Disease resistance protein (CC-NBS-LRR ...	0.002	pfam_fs:NB-ARC Domain	5,40E-07
NBS 5	AS-NBSMCTT-24A	Disease resistance protein (NBS-LRR cla...	2,00E-23	pfam_fs:NB-ARC Domain	4,50E-25
NBS 5	AS-NBSMCTT-24E	Disease resistance protein (CC-NBS clas...	7,00E-14	pfam_fs:NB-ARC Domain	5,80E-30
NBS 5	AD-NBSMCTT-24F	Putative late blight resistance protein [Solanum...	0.029	pfam_fs:NB-ARC Domain	3,70E-04

^a Homologias encontradas com seqüências de nucleotídeos traduzidas para aminoácidos pelo próprio BLASTX e comparadas com o GenBank. ^b seqüências traduzidas pelo programa "Translate Tool - TrEMBL" da Home Page: www.expasy.ch/tools/dna.html e processadas pela ferramenta "Motif-Scan" da Home Page: www.expasy.ch/prosite.

Outro aspecto analisado nas seqüências RGAs foi o tipo de estrutura (TIR ou NãoTIR) encontrado na porção N-terminal das amostras. Por meio de comparação com as seqüências completas já depositadas no GenBank, pôde-se prever qual estrutura encontra-se em cada amostra. Os RGAs 9A, 9B, 9C, 12F, 26A, 26B, 27A, 27B e 28G possuem a estrutura do tipo TIR (ou seja, esse grupo é composto por genes que codificam produtos contendo uma estrutura homóloga a da *Drosófila* “Toll-Like” e à humana “Interlekin-1” que são regiões receptoras). Os RGAs 8D, 8J, 14A 14D e 24E não possuem essa estrutura, sendo então denominadas NãoTIR, ou seja, compreendem genes que codificam o padrão conservado “Leucine-Zipper-Like”. Também é chamada de Coiled-Coil. As outras amostras RGAs não puderam ter o seu grupo definido por comparação com seqüências conhecidas. Alguns *primers* utilizados, como NBS5 e P2-414 possuem em sua porção 3’ a seqüência de aminoácidos DD(I/V)W conservada de Triptofano. Essa conservação nessa localidade está ligada a classe NãoTIR de genes de resistência (Meyers et al., 1999; Pan et al., 2000). *primers* com essa porção terminal com a seqüência DD(I/V)D amplificam regiões de genes do tipo TIR

Para que os RGAs encontrados nesse trabalho pudessem ser comparados entre si, foi feito um alinhamento das seqüências traduzidas para aminoácido através dos programas Jalview e ClustalW (www.ebi.ac.uk). Em uma primeira análise, apenas os RGAs encontrados foram alinhados, mostrando claramente as regiões conservadas na seqüência e a região de anelamento do *primer* (Figura 15). As amostras 9B, 9C, 14A e 28G não possuem a região de anelamento do *primer* nesse alinhamento, o que significa que a seqüência não está completa, foi usada apenas a leitura relativa ao *primer* reverso (complementar ao sítio de *MseI*), onde o mesmo não conseguiu alcançar a região Kinase-2 no começo da seqüência. Apesar dos baixos valores de significância para algumas amostras, no alinhamento é possível observar a presença das regiões conservadas que caracterizam o domínio NBS.

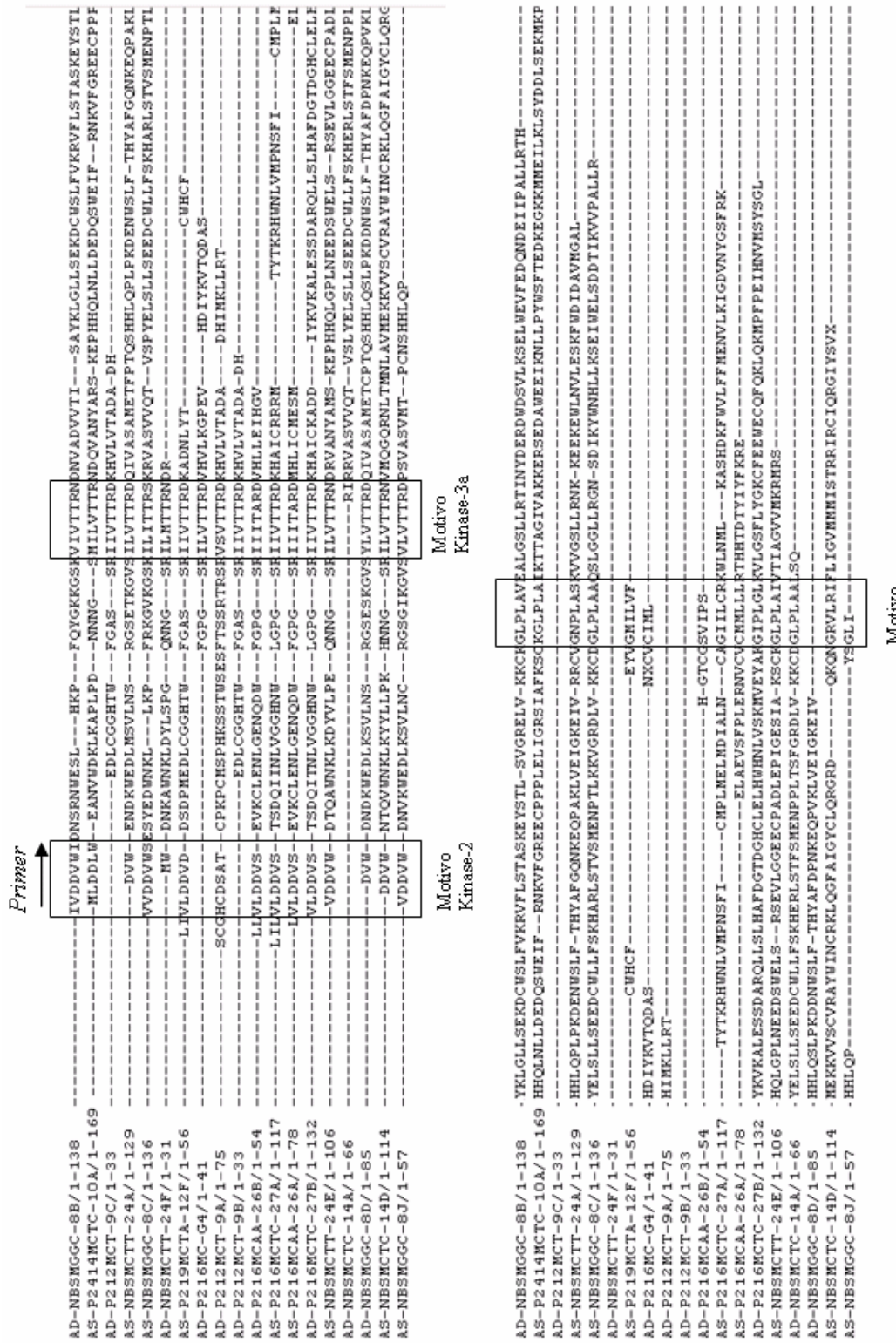


Figura 15 Imagem de alinhamento feito pelo programa Jalview com tipo de alinhamento ClustalW. As 19 seqüências traduzidas de RGAs encontrados neste trabalho foram alinhadas e tiveram as regiões conservadas destacadas. As caixas mostram os motivos Kinase-2, Kinase-3a e GLPLA do domínio NBS. A seta representa o sentido do primers NBS e a região de anelamento do mesmo. Os números após o nome dos RGAs representam o número de aminoácidos em cada seqüência.

Após o alinhamento múltiplo das seqüências foi feita uma árvore de similaridade com as dezenove seqüências RGAs encontradas neste trabalho. A ferramenta utilizada para a confecção da árvore foi a “Distância Média por Porcentagem de Identidade” do programa Jalview. As seqüências foram divididas em dois grupos que representam as estruturas TIR e NãoTIR. Aproximadamente o dobro de amostras apresentara a estrutura TIR em relação ao outro grupo.

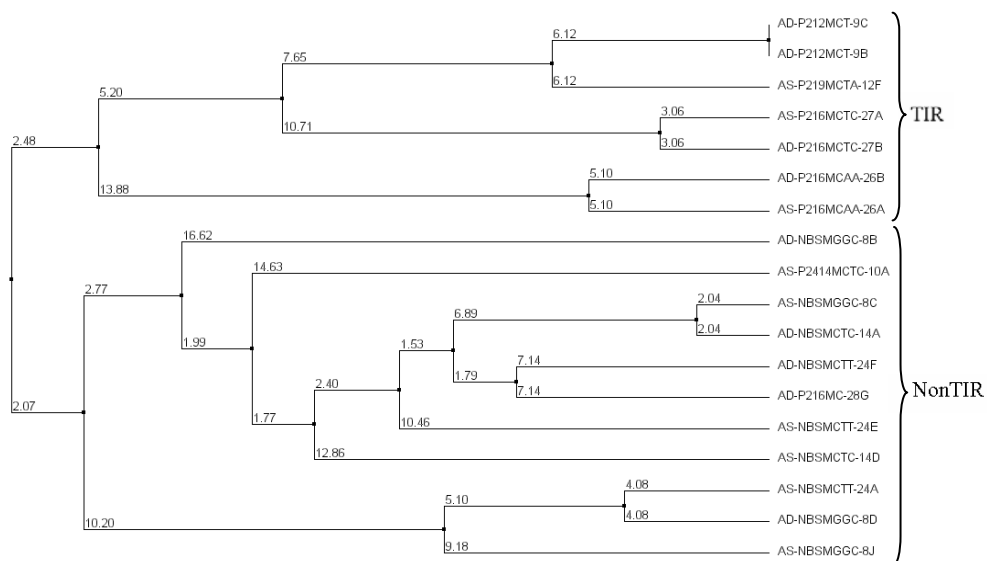


Figura 16 Árvore de similaridade baseada em um alinhamento ClustalW de seqüências de proteínas dos RGAs encontrados neste trabalho. Seqüências do motivo Kinase-2 até GLPLA foram usadas para o alinhamento. Os números indicam a distância média calculada através da similaridade das seqüências. As amostras foram divididas de acordo com sua estrutura na porção N-terminal TIR ou NãoTIR.

Em um segundo momento, as seqüências traduzidas alinhadas foram comparadas com seqüências disponíveis no Genbank. Para a comparação foram utilizados 67 RGAs descritos por Bertoli et al., 2003 (Figura 16). O motivo Kinase-2 (caracterizado pela região conservada LILDDVW) é extremamente conservado em quase todas as seqüências, só não aparecendo nas seqüências incompletas, como no caso das amostras 9B, 9C, 14A e 28G. O motivo Kinase-3a também está extremamente conservado em todas as amostras, menos na 14A. Este motivo está representado pelos aminoácidos (SRILMTTR – com variações).

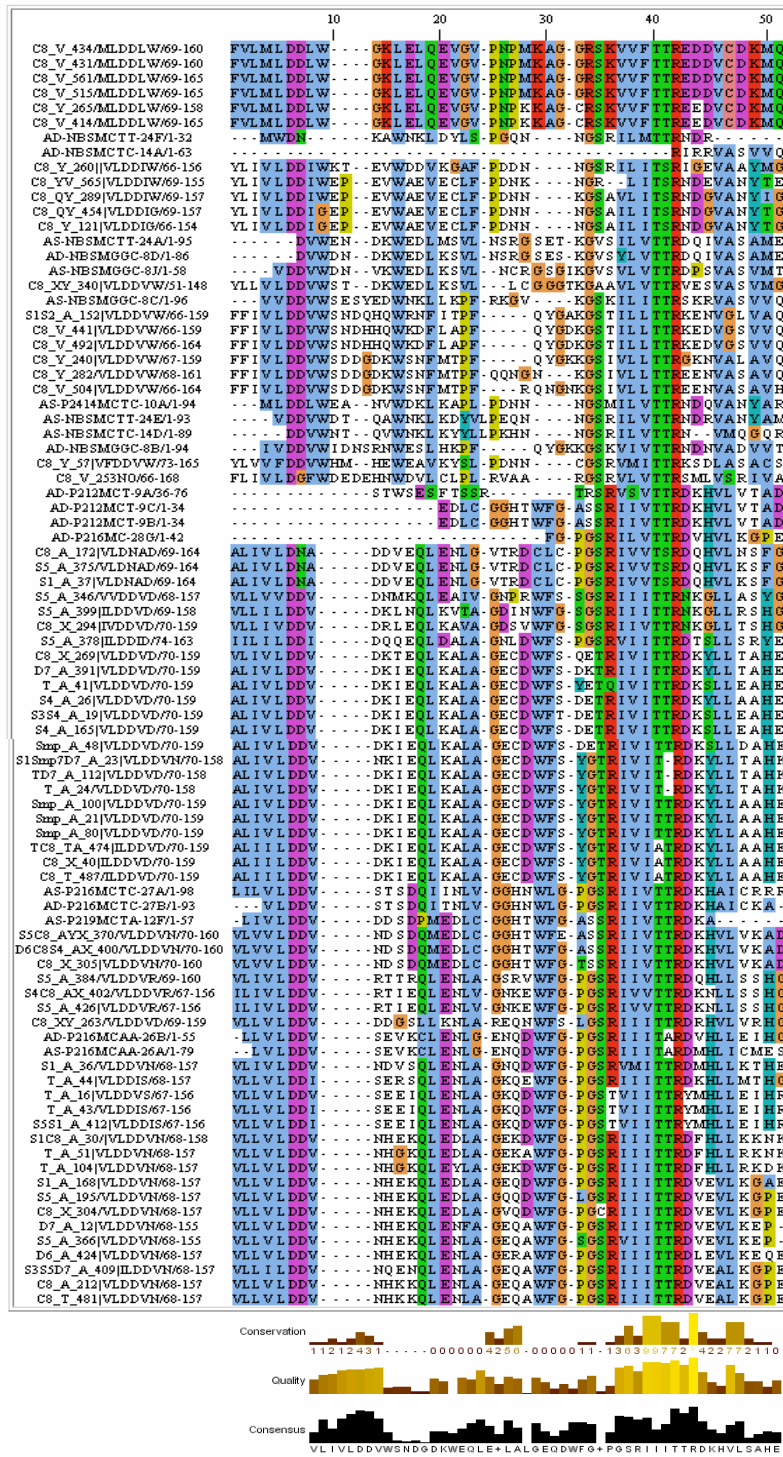


Figura 17 Imagem Jalview com alinhamento das seqüências RGAs, proteínas de resistência de *Arachis* spp. e domínio NBS. A parte inferior da figura mostra o grau de conservação da seqüência consenso e a qualidade do alinhamento. As amostras estão dispostas em grupos respeitando a ordem da árvore de similaridade.

Uma segunda árvore de similaridade foi construída para o alinhamento contendo também as seqüências RGAs obtidas por Bertoli et al., 2003. Nessa árvore, 67 seqüências RGAs depositadas no GenBank para *Arachis* foram utilizadas para uma comparação com as 19 seqüências RGAs encontradas aqui, onde através de distância média por porcentagem de identidade, as amostras foram separadas em grupos de similaridade de seqüência (Fig. 17).

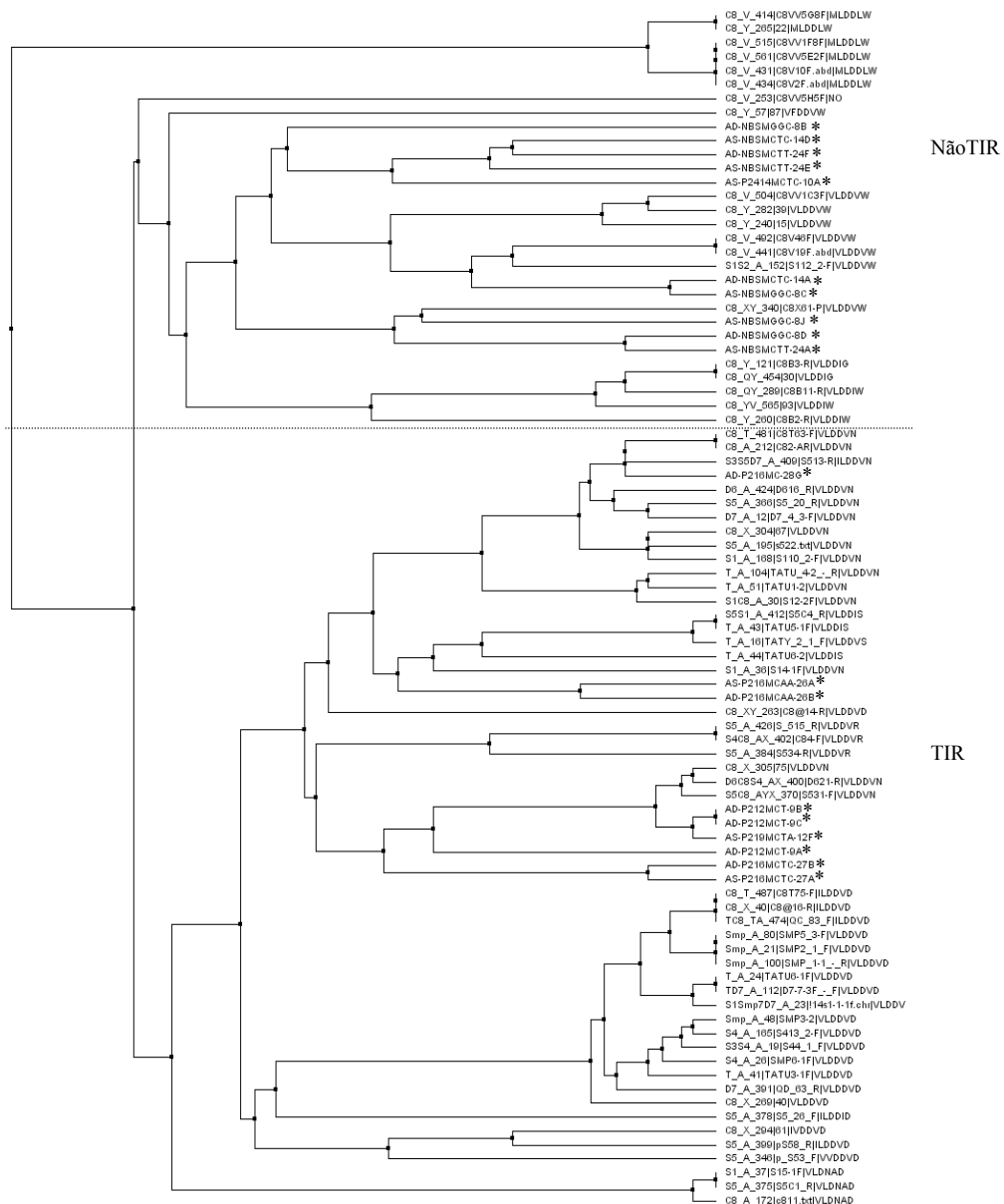


Figura 18 Árvore de similaridade das 19 amostras RGAs obtidas nesse trabalho e 69 obtidas em Bertoli et al., 2003. A árvore foi construída baseada em um alinhamento do tipo ClustalW de seqüências protéicas cortadas entre o motivo Kinase-2 e Kinase-3a. Os asteriscos representam as seqüências RGAs isoladas no presente trabalho.

Dois grandes grupos foram formados na árvore contendo os 19 RGAs isolados no presente trabalho e 67 RGAs descritos por Bertoli et al. (2003). Um grupo maior foi identificado por conter RGAs com estrutura do tipo TIR e o segundo grupo, um pouco menor, possui RGAs com estrutura do tipo NãoTIR. Um terceiro grupo, bem pequeno, contendo seis RGAs previamente isolados também foi identificado por não se encaixar nos dois grandes grupos.

Algumas amostras seqüenciadas mostraram homologia com proteínas do tipo Kinase-2 quando submetidas ao programa BLASTX e comparadas com o banco de dados não redundante. Essas amostras (AS-NBSMCTT-24H, AD-P2378MC-20B, AD-P2391MCTA-13D, AS-P2391MCTA-13C, AD-P2378MCTA-23A) foram, então, alinhadas juntamente com outras quatro seqüências com homologia para essa mesma proteína já depositadas no GenBank (gi42567050 kinase - *Arabidopsis thaliana*; gi4826399 wall-associated kinase 2 - *Arabidopsis thaliana*; gi15218170 WAK5 - Wall Associated Kinase e gi15218959 WAK2; ATP binding / calcium ion binding). Essas seqüências foram escolhidas por serem a primeira homologia mostrada no BLASTX. A Figura 18 representa um alinhamento do tipo ClustalW para essas nove seqüências traduzidas para aminoácidos.

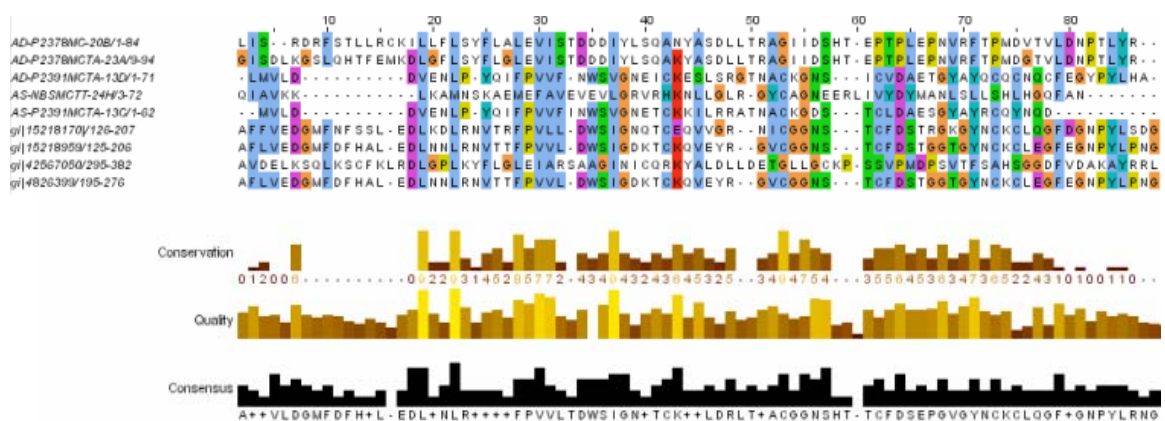


Figura 19 Alinhamento ClustalW utilizando o programa JalView de seqüências de proteínas do tipo kinase. As cinco primeiras amostras são seqüências obtidas do processamento de bandas polimórficas da técnica RGA-display e as quatro últimas são proteínas do tipo kinase de domínio público depositadas no GenBank (www.ncbi.nlm.nih.gov). A parte inferior da figura representa os gráficos de conservação das seqüências, a qualidade do alinhamento e a seqüência consenso obtida do alinhamento.

Observando-se o alinhamento, é possível a identificação de regiões de grande similaridade entre as amostras do presente trabalho e as já identificadas, corroborando com o esperado que as seqüências sejam proteínas do tipo Kinase-2. Esses genes que codificam as proteínas do tipo Kinase-2, têm função embriogênica ou meristemática (Haas et al., 2002). Segundo Yuksel et al. (2005) as seqüências contendo homologia com a proteína Kinase podem estar relacionadas com genes de resistência, embora as seqüências que apresentaram homologia com a proteína Kinase no presente trabalho não possuem as regiões conservadas do domínio NBS (NB-ARC).

Dois marcadores RGAs identificado na técnica RGA-Display (AD-NBSMGGC-8B e AD-NBSMGGC-8B) foram escolhidos para a hibridização com DNA genômico dos parentais em Southern-blot para a identificação do número de *loci* dos RGAs. Esses marcadores foram re-amplificados, resolvidos em gel de agarose, purificados com o Kit “QIAquick Gel Extraction” (Quiagen) e usados como sonda contra DNA dos parentais *A. duranensis* e *A. stenosperma*. O RGA AD-NBSMGGC-8B apresentou dez cópias no DNA digerido com *HindIII* e 11 cópias no DNA digerido com *EcoRI* enquanto que o RGA AD-NBSMGGC-8B apresentou 9 cópias tanto no DNA digerido com *HindIII* quanto no digerido com *EcoRI* (dados não apresentados). Nenhum polimorfismo foi observado entre os parentais com esses RGAs hibridizados.

Mapeamento Genético dos Marcadores

Os marcadores desenvolvidos no presente trabalho, foram inseridos em um mapa de ligação já existente, baseado em marcadores microssatélite (Moretzsohn et al., 2005). Neste mesmo mapa, foram incluídos marcadores AFLP desenvolvidos no Sainsbury laboratories na Inglaterra e marcadores âncora, desenvolvidos na Universidade de Aarhus, Dinamarca. No total, foram mapeados 335 marcadores, sendo 180 Microssatélite, 65 marcadores âncoras, 46 AFLPs, 3 RAPDs e 39 RGA-SAP (sendo 15 RGAs), usando-se um LOD mínimo de 5,0 e uma fração de recombinação máxima (θ) de 0,35 (Figura 20).

Os marcadores 14B e 14C, que têm homologia com o gene de β -amilase, mapearam com 0% de recombinação, ou seja, parecem estar ligados. Uma análise da presença e ausência de bandas destes dois marcadores na genotipagem da população F_2 revelou que eles poderiam ser co-dominantes. As seqüências nucleotídicas e protéicas dos marcadores em questão foram então comparadas para a busca de pequenas diferenças entre eles. Alguns nucleotídeos se apresentaram diferentes nas duas seqüências e essas diferenças causaram a troca do aminoácido resultante nessas regiões. A Figura 19 é um alinhamento dessas duas seqüências protéicas mostrando pequenas mutações em algumas regiões. A região central do alinhamento mostra estar altamente conservada nas duas seqüências e as diferenças encontram-se apenas nas extremidades, o que pode sugerir também que a troca de nucleotídeos seja em virtude de uma qualidade mais baixa do seqüenciamento nessas regiões de início e fim da amostra.

Os marcadores 9A e 9B também apareceram ligados no mapa, mas a análise da genotipagem revelou que os dois marcadores segregam exatamente juntos (dados não mostrados), e a sua seqüência nucleotídica também é igual. Com essa análise, é possível que sejam o mesmo marcador, pois as bandas no gel de poliacrilamida estavam próximas uma da outra.


```

Score = 5250
Length of alignment = 112
Sequence AD-NBSMCTC-14C : 8 - 118 (Sequence length = 119)
Sequence AS-NBSMCTC-14B : 1 - 112 (Sequence length = 112)

AD-NBSMCTC-14C VESKGTQEYDWSAYRDLFQLVQHCKLKLQAIMSFHQCGGNVGDVSVSIPKQWLEIG
               | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
AS-NBSMCTC-14B VASKGPEYDWSAYRDLFQLVQDCE LKLAIMSSHQCGGHVGDVSVSIPKQWLEIG

AD-NBSMCTC-14C VSDPDIFYYTNPNGIRNKECLSIGVDNRPLFHGRAAIEVDHVNLCFLTHILSI-C
               | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
AS-NBSMCTC-14B VSDPDIFYYTNPNGIRNKECLSIGVDNRPLFHGRTAIEVEHVQNFCALTHILSIYC

Percentage ID = 89.29

```

Figura 20 Alinhamento das seqüências protéicas dos marcadores 14B e 14C homólogos ao gene de β -amilase. O alinhamento é do tipo ClustalW e os pares foram comparados com a função “Pairwise Alignment”.

Dos nove marcadores obtidos através da técnica RFLP, quatro puderam ser mapeados. Embora esses marcadores possam ser bem mais informativos do que os RGA-Display se foram mapeados como co-dominantes, não apresentaram esse tipo de característica nesse caso. Devido a uma falta de resolução apropriada dos fragmentos hibridizados, os marcadores foram analisados como dominantes, embora tenham mapeado muito próximos de seus pares. Os marcadores S1-A-36B e S1-A-36D mapearam a uma distância de 5,3 cM no Grupo de Ligação 4, enquanto que os marcadores S1-A-37A e S1-A-36C mapearam a uma distância de 14,3 cM no Grupo de Ligação 3.

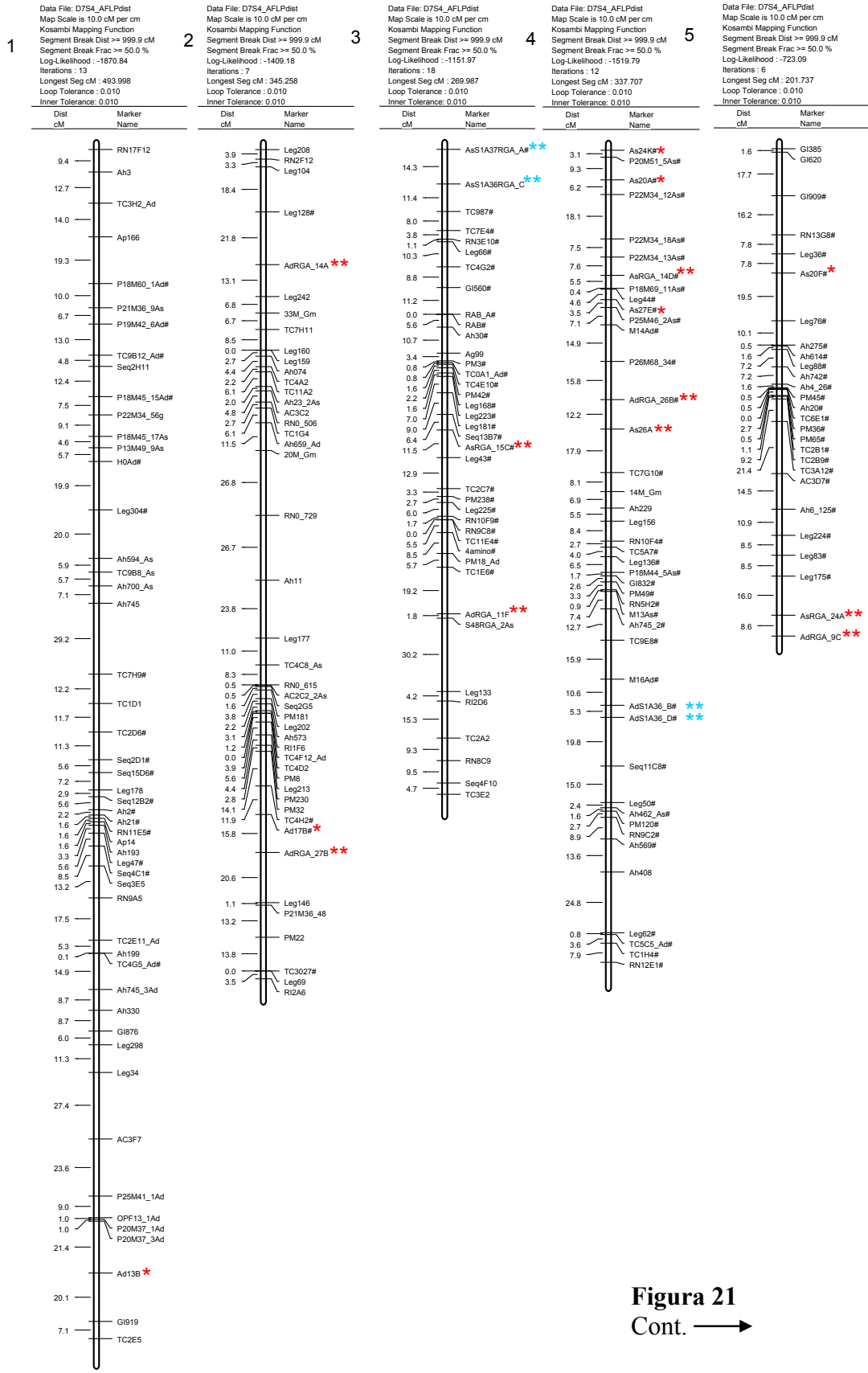


Figura 21
 Cont. →

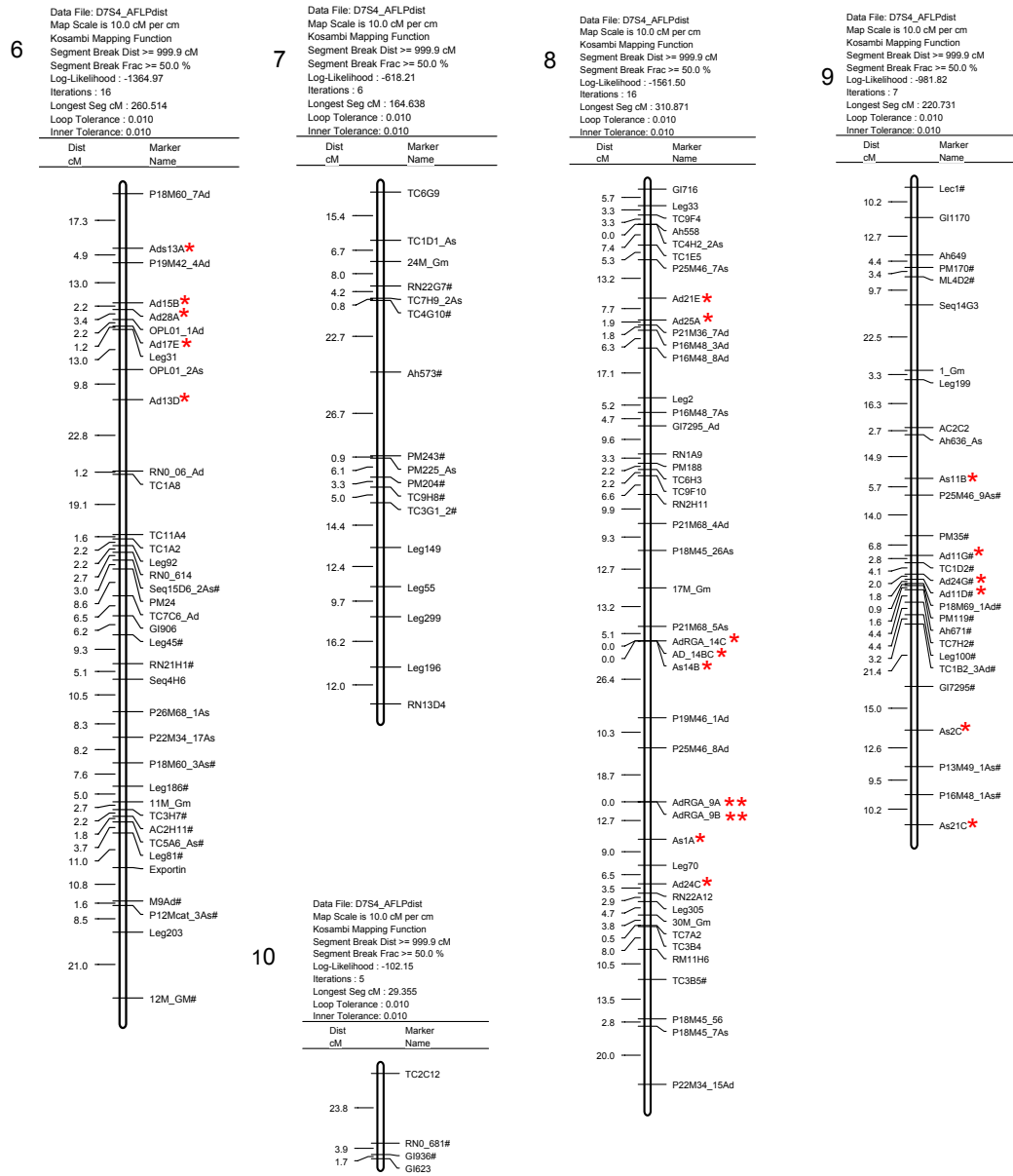


Figura 21 Mapa Genético de ligação para *Arachis*, baseado em marcadores microssatélite (SSR), obtido através de análise de 93 plantas de uma população F₂ gerada a partir do cruzamento de duas espécies silvestres diplóides com genoma AA, *A. duranensis* e *A. stenosperma*. O LOD mínimo utilizado foi de 5 e a máxima fração de recombinação (h) de 0,35. 335 marcadores foram mapeados em 10 Grupos de Ligação. Marcadores dominantes estão identificados por AD ou AS antes ou após o nome se amplificaram alelos para *A. duranensis* ou *A. stenosperma* respectivamente. As amostras RGA-Display estão identificadas com um asterisco preto e as que possuem homologia com RGAs possuem dois asteriscos pretos. Marcadores provenientes de RFLP possuem dois asteriscos cinza.

Dos 658 marcadores polimórficos incluídos na análise, 317 (48,2%) desviaram das proporções 1:2:1 ou 3:1 esperadas, ao nível de 5%. Essa proporção de marcadores distorcidos é maior do que a detectada em dois outros mapas de *Arachis*, baseados em marcadores RFLP, que foi de 25% em média (Halward et al., 1993; Burow et al., 2001). Com exceção do Grupo 8, os demais nove grupos de ligação apresentaram marcadores distorcidos.

5- Discussão

Nos últimos anos, um grande número de genes de resistência foi identificado e clonado. Na maioria dos casos, esses genes são da classe NBS-LRR que codificam proteínas receptoras que parecem reconhecer um fator de avirulência e dá início à resposta de defesa da planta contra patógenos. O conhecimento sobre a seqüência de um grande número desses genes de resistência presentes em genoma de plantas, assim como a organização genômica estão ajudando no entendimento do mecanismo básico de resistência ao patógeno, que serão de grande valor no melhoramento de culturas.

Neste trabalho, marcadores moleculares relacionados com genes de resistência e com outras proteínas importantes foram identificados e mapeados com sucesso utilizando uma população resultante do cruzamento de dois acessos silvestres da *Arachis* através de duas técnicas baseadas em RGAs: RFLP e RGA-Display.

Geração de marcadores através de RFLP

O mapeamento genético baseado em marcadores RFLP pode ser muito informativo e já foi feito para *Arachis* (Halward et al., 1993, Garcia et al., 1995, Burrow et al., 2001). No presente trabalho, nove marcadores foram identificados através da técnica e quatro puderam ser mapeados, no entanto, a técnica é muito dispendiosa financeiramente e demorada, além da dificuldade de se obter um número adequado de marcadores. Para a obtenção das membranas a serem hibridizadas, muito DNA precisa ser concentrado em um volume muito pequeno que, no caso de espécies de *Arachis* se torna um problema, pois o DNA dessas plantas é altamente contaminado por compostos que o deixa viscoso em altas concentrações e de fácil degradação. Conseqüentemente, esta não é uma ferramenta ideal para um trabalho de análise e

implementação de seleção assistida por marcadores (MAS) em programas de melhoramento do amendoim (Moretzsohn et al., 2005).

Os marcadores obtidos com a técnica podem ser gênicos, embora o número (quatro marcadores) é muito pequeno quando colocado em um mapa genético. Técnicas baseadas em PCR fornecem rapidamente um número muito mais significativo de marcadores. As sondas utilizadas nesse trabalho já haviam sido usadas para o desenvolvimento de marcadores por Guimarães e colaboradores em 2005, onde 33 marcadores RGA foram identificados para as espécies *A. duranensis* e *A. stenosperma*.

A técnica RGA-Display

A técnica descrita por Linden e colaboradores em 2004 já havia sido relatada por Hayes e Saghai Maroof em 2000 onde, estes, descreveram a utilização de uma técnica baseada em um AFLP direcionado para RGAs, encontrando marcadores ligados a um gene de resistência em soja. A modificação proposta por Linden et al. (2004) e utilizada no presente trabalho, foi a utilização de *primers* para os motivos do domínio NBS, com diversos graus de degeneração e o seqüenciamento dos fragmentos para a confirmação da sua homologia com genes de resistência.

A técnica RGA-Display possui muitas vantagens na geração de marcadores RGA em relação ao RFLP, como por exemplo, o custo moderado de material, o tempo e o grande número de marcadores produzidos. Por estes motivos, a técnica mostrou ser uma ferramenta muito útil na construção de marcadores moleculares, pois utiliza partes de genes de resistência conservadas para a amplificação de produtos, que, provavelmente são regiões de genes da mesma família (Soriano et al., 2005; Calenge et al., 2005). O presente trabalho é o primeiro

estudo deste tipo para o gênero *Arachis*, onde RGAs e outros genes foram mapeados com o auxílio dessa técnica. Anteriormente, apenas marcadores como RFLP, RAPD e SSR haviam sido incluídos em mapas genéticos de *Arachis* (Halward et al., 1993, Garcia et al., 1995, Burrow et al., 2001).

Para a otimização dos procedimentos, três tipos de polimerase foram testados com o objetivo de aumentar o rendimento da técnica. Assim como sugerido por Linden, foi detectado que o uso de uma polimerase do tipo Hot Start é fundamental para a obtenção de bons resultados. Através da técnica, 139 marcadores foram desenvolvidos e 19 mostraram ser RGAs após o seqüenciamento das bandas polimórficas. Quinze seqüências apresentaram forte homologia com outros tipos de proteínas, como Kinases, β -amilase, Hidrolase e outras. Semelhanças com Retrotransposons também foram encontradas em treze seqüências. Cinquenta e três marcadores não mostraram homologia com outras seqüência depositadas no GenBank, contudo, não podemos descartar a hipótese que essas seqüências sejam genes importantes ainda não descritos.

A técnica já foi anteriormente utilizada com sucesso para diversas espécies, como rosa, damasco, maçã, tomate e batata. Em uma análise feita para Damasco (*Prunus armeniaca* L.) 27 marcadores foram identificados, e destes, sete (26%) mostraram homologia para genes de resistência do tipo TIR (Soriano et al., 2005). Com a técnica, 16 marcadores puderam ser mapeados em um mapa previamente construído baseado em marcadores Microsatélite (SSR). Entre os marcadores identificados, seis mostraram homologia para proteínas depositadas no GenBank e quatro com Retrotransposons. Dez seqüências não tiveram nenhuma homologia com seqüências já depositas. Nenhum marcador do tipo RGA foi mapeado neste trabalho, enfatizando a importância do seqüenciamento das amostras obtidas com a técnica. Em outro trabalho (Calenge et al., 2005) foram mapeados 43 marcadores pela técnica RGA-Display em um mapa construído para maçã, onde 23 (53%) eram RGAs. Os marcadores foram mapeados

em 10 dos 17 Grupos de Ligação do mapa e muitos estavam organizados em pequenos agrupamentos. Vinte e cinco marcadores mapearam perto de genes ou QTLs para resistência à duas das maiores pragas da espécie.

Os resultados das análises das seqüências quando comparadas com seqüências depositadas no GenBank através de BLASTX (Altschul et al., 1997) causaram alguma dúvida se realmente tinham homologia com RGAs pelo discreto valor de Nível de Significância encontrado (*e-value*). O Nível de significância mínimo utilizado foi de 0,01, o que é um valor considerado insuficiente para outros tipos de análise. Visto que as bandas polimórficas produzidas e usadas para seqüenciamento eram de tamanho pequeno, as seqüências produzidas também eram pequenas. Quando foi feita uma comparação destas seqüências com outras já depositadas no GenBank de tamanhos maiores, os valores de *score* e *e-value* se apresentam pequenos em vista dessa diferença de tamanho entre as seqüências. A baixa qualidade das seqüências e o seu tamanho reduzido devem estar contribuindo para os valores referidos. Essa baixa qualidade deve-se ao fato das bandas serem retiradas do gel de poliacrilamida, re-amplificadas, retiradas do gel de agarose e seqüenciadas. Nesse processo, deve haver um pouco de degradação do DNA assim como a inibição de uma amplificação de qualidade por meio de substâncias não desejáveis nas reações de PCR. Para que a confiança fosse aumentada, outra estratégia também foi utilizada para a detecção de RGAs. As seqüências de DNA foram traduzidas para aminoácidos e a ferramenta Motif-Scan foi utilizada para encontrar regiões semelhantes a RGAs.

Após o alinhamento das seqüências protéicas, aumentou-se o nível de confiança que as mesmas são RGAs, pois os principais motivos que caracterizam o domínio NBS em genes de resistência mostraram-se claramente conservados em todas as seqüências previamente ditas como sendo RGAs. Esse argumento também foi utilizado como base de confirmação em outros trabalhos (Meyers et al., 2003; Van der Linden et al., 2004; Calenge et al., 2005).

Embora pareçam RGAs, não se pode excluir a possibilidade de que as seqüências com essa homologia possam ser pseudogenes (Soriano et al., 2005). De fato, já foi proposto que uma proporção significativa de genes do tipo NBS podem ser genes não funcionais (Meyers et al., 1999).

No total, 35 marcadores RGA-Display foram mapeados, onde os mesmos estavam bem distribuídos dentro dos grupos. Alguns marcadores mostraram uma leve tendência à formação de agrupamentos em alguns grupos de ligação (Figura 21). Segundo Meyers et al., 1999, é comum a formação desses pequenos agrupamentos dentro dos Grupos de Ligação, que pode ser chamado de região gênica. Onze marcadores que foram mapeados mostraram ser RGAs.

Embora algumas combinações de *primers* tenham rendido uma porcentagem alta de RGAs (100% das bandas polimórficas no caso da combinação P2-12 e Mct) outras não apresentaram RGAs no seqüenciamento de suas bandas polimórficas (como no caso da combinação P2-16 e Mcaa) diminuindo a média de RGAs em relação às bandas polimórficas obtidas. A porcentagem média de RGAs encontrada no presente trabalho (19%), assim como os resultados obtidos por outros autores, não corresponde ao descrito por Linden et al., 2004 que obteve um rendimento de 50 a 90% de RGAs. A proporção de seqüências não identificadas (53%) também é maior do que a encontrada por esse autor (21%), onde o mesmo diz que essas seqüências não devem ser ignoradas, e que uma grande proporção das mesmas devem ser seqüências ainda não descritas aumentando a possibilidade de que mais RGAS sejam isolados pela técnica. A escolhas dos *primers* a serem utilizados na genotipagem deve ser crucial para a manutenção de uma média alta de bandas RGAs na técnica. Desse modo, é sugerido que as bandas sejam seqüenciadas antes da genotipagem, para que somente aquelas que mostrarem homologia com os mesmos sejam utilizadas na caracterização da população de mapeamento.

Mapeamento Genético

Na análise dos dados para a inclusão dos marcadores no mapa foram usados 148 marcadores desenvolvidos no presente trabalho (139 marcadores obtidos pela técnica RGA-Display e 9 por RFLP). Trinta e nove marcadores mapearam em oito Grupos de Ligação (Grupos 1, 2, 3, 4, 5, 6, 8 e 9) e os demais marcadores incluídos não mapearam. Além disso, dos 39 marcadores RGA-Display mapeados, 20 apresentaram valores de BLASTX significativos. Dentre essas homologias destacam-se RGAs, Proteínas Kinase-2, β -amilases e Retrotransposons (Gráfico 1). Com isso, abre-se a perspectiva de usar essas informações para criar um mapa rico em marcadores desenvolvidos a partir de genes candidatos.

Como esperado, dez grupos de ligação foram obtidos. O número de marcadores por grupo de ligação variou de 4 (Grupo 10) a 51 (Grupo 1). A inserção de novos marcadores, em sua maioria dominantes, aumentou sobremaneira o tamanho dos grupos e do mapa total, em cM. Isso pode ter sido causado por alguns erros de genotipagem ou, mais provavelmente, por dificuldades do programa de mapeamento (Mapmaker) em trabalhar com marcadores co-dominantes e dominantes em conjunto. Uma evidência desse fato é que, dos 206 marcadores AFLP polimórficos incluídos na análise, apenas 46 mapearam e em somente cinco dos 10 grupos de ligação (Grupos 1, 4, 6, 8 e 9). Com isso, os novos marcadores devem ter sido inseridos na ordem correta, mas os valores das distâncias entre marcadores foram, possivelmente, superestimados. Devido a esse problema, a numeração dos grupos de ligação não obedeceu à ordem de tamanho dos mesmos. Dos marcadores RGA-Display, apenas 35 mapearam em oito grupos de ligação, mostrando-se bem distribuídos mas com a formação de pequenos grupos em alguns grupos de ligação.

Considerações Gerais

Os programas de melhoramento de plantas cultivadas têm sido amplamente auxiliados por Seleção Assistida por Marcadores Moleculares (MAS). A obtenção de mapas genéticos, possuindo marcadores gênicos ou ligados a Genes de Resistência, é um passo muito importante para esse processo. Sendo assim, o desenvolvimento de marcadores do tipo RGA é uma ferramenta promissora para a identificação de *loci* associados a esses genes, uma vez que já foi estabelecido que uma grande proporção de RGAs de várias espécies está ligado ou faz parte de genes de resistência. (Collins et al., 2001, Donald et al., 2002, Leal-Bertioli et al., 2004). Esses genes podem estar agrupados no cromossomo, possibilitando também a piramidização de genes de resistência (Aarts et al., 1998). Além disso, vários trabalhos demonstraram que genes de resistência podem manter sua atividade quando transferidos para outras espécies (Hammond-Kosack et al., 1998), o que também abre a perspectiva de que, uma vez identificados os genes de resistência em *Arachis* spp., estes possam ser transferidos para outras espécies, especialmente leguminosas.

Com as duas técnicas utilizadas, muitos marcadores gênicos foram incluídos no mapa genético recém construído de alta resolução. Para isso, foi utilizada uma população F₂ originada do cruzamento entre acessos (*A. duranensis* – K7988 e *A. stenosperma* – V10309) que contrastam para diversas pragas e patógenos como os nematóides das galhas *Meloidogyne arenaria* raça 1 e raça 2, *M. javanica* raça 4 e *M. hapla*. No momento, bioensaios estão sendo conduzidos para a obtenção de QTLs (*Quantitative Trait Loci*) e resistências ao fungo *M. arenaria* raça 1 que é uma das maiores pragas do amendoim cultivado nos Estados Unidos com grande potencial de infectar as culturas brasileiras.

Uma vez inserida esta resistência no mapa, será possível observar a associação desta com marcadores RGA. Estes marcadores poderão então ser utilizados em programas de seleção assistida, visando a introgressão dessa resistência no amendoim cultivado.

Após todas as análises, os resultados sugerem que o desenvolvimento e o mapeamento desses marcadores RGA-Display poderão contribuir para a identificação de regiões gênicas contendo prováveis genes ligados a resistência a patógenos do amendoim.

Estudos recentes têm demonstrado o potencial de uma outra técnica Single – Nucleotide Polymorphism (SNPs) como a próxima geração de marcadores genéticos em plantas (Hayashi et al., 2004) visto que esses marcadores são muito abundantes e ocorrem com frequência no genoma de plantas parecendo ser relativamente estáveis geneticamente. Esses poderão ser utilizados em espécies com baixo grau de polimorfismo como no caso do amendoim.

6- Conclusões

1. A escolha do método de extração de DNA para espécies de *Arachis* é um processo muito importante na geração de marcadores moleculares, visto que as folhas contêm muitos compostos que atuam na degradação e viscosidade do DNA dessas plantas, desfavorecendo o uso de algumas técnicas que necessitam o uso de grandes quantidades de DNA.
2. Com o uso de RFLP, nove marcadores gênicos RGA foram desenvolvidos e genotipados para uma população F₂ segregante de 93 indivíduos oriundos do cruzamento de duas espécies silvestres de *Arachis*. Destes, quatro puderam ser mapeados em mapa genético previamente construído a partir de marcadores Microsatélite.
3. Com a técnica RGA-Display foram isolados um total de 139 marcadores moleculares a partir de 24 combinações de *primers*. Quando seqüenciados, 19 marcadores mostraram homologias a genes de resistência, 15 a outras proteínas importantes, como β -amilase, Kinase e hidrolases e 13 mostraram forte homologia com Retrotransposons. 53 seqüências não apresentaram homologia com proteínas já descritas. Trinta e nove marcadores puderam ser mapeados, sendo que 11 são prováveis RGAs e nove são importantes proteínas já mencionadas.
4. A distorção da segregação encontrada na genotipagem da população F₂ foi considerada um pouco acima do já relatado para o gênero a partir de outras técnicas de marcadores moleculares.
5. As seqüências RGAs foram alinhadas e os principais motivos que caracterizam o domínio NBS se mostraram altamente conservados em todas as amostras.
6. Os marcadores foram mapeados em oito dos dez Grupos de Ligação existentes no mapa, com um LOD mínimo de 5,0 e uma fração de recombinação de 0,35. Os marcadores se apresentaram bem distribuídos no mapa, mostrando uma leve tendência a se agruparem em alguns GL (5, 6, 8 e 9 - Fig. 21).

7. O trabalho obedeceu ao objetivo proposto de desenvolvimento e mapeamento de marcadores RGA para as duas espécies silvestres de *Arachis*, confirmando que a técnica RGA-Display é uma boa ferramenta para o mesmo, fornecendo um bom número de marcadores com seqüência conhecida em um pequeno intervalo de tempo, e, pouco dispendiosa quando comparada ao RFLP.

8. Os resultados obtidos aqui serão comparados com QTLs oriundos de bioensaios que estão sendo produzidos pelo grupo para nematóides que causam grandes problemas na cultura. A partir disso, os mesmos poderão ser de grande auxílio no melhoramento do amendoim.

7- Referências Bibliográficas

- AARTS MGM, HEKKERT BL, HOLUB EB, BEYNON JL, STIEKEMA WJ, PEREIRA A. **Identification of R-gene homologous DNA fragments genetically linked to disease resistance loci in *Arabidopsis thaliana*.** Mol Plant–Microbe Interact, 11:251-258 (1998).
- AGRIOS G N. **Plant Pathology.** Fourth Edition, Academic Press, USA, 635 p. (1997).
- ALTSCHUL SF, THOMAS L, MADDEN A, SCHÄFFER A, ZHANG J, ZHANG Z, MILLER W, LIPMAN, DJ. **Gapped BLAST and PSIBLAST: a new generation of protein database search programs.** Nucleic Acids Res 25:3389–3402 (1997).
- BARZEN E, MECHELKE W, RITTER E, SCHULTE-KAPPERT E, SALAMINI F. **An extended map of the sugar beet genome containing RFLP and RFLP loci.** Theor Appl Genet 90:189–193 (1995).
- BENNET MD, SMITH JB. **Nuclear DNA amounts in angiosperms.** Philosophical Transactions of the Royal Society of London. B274 227-274 (1996).
- BERRY ST, LEON AJ, HANFREY CC, CHALLIS P, BURKHOLZ A, BARNES SJ, RUFENER GK, LEE M, CALIGARI PDS. **Molecular marker analysis of *Helianthus annuus* L. 2. Construction of an RFLP linkage map for cultivated sunflower.** Theor Appl Genet 91:195–199 (1995).
- BERTIOLI DJ, LEAL-BERTIOLI SC, LION MB, SANTOS VL, PAPPAS G. JR, CANNON SB, GUIMARÃES PM. **A large scale analysis of resistance gene homologues in *Arachis*.** Molecular Genetics and Genomics 270:34-45 (2003).
- BOTSTEIN D, SMITH KF, SKOLNICK M, DAVIS RW. **Construction of a genetic linkage map in man using restriction fragment length polymorphism.** Am. J. Hum. Genet. 32:314-331 (1980).
- BUROW MD, SIMPSON CE, STARR JL, PATERSON AH. **Transmission genetics of chromatin from a synthetic amphidiploid to cultivated peanut (*Arachis hypogaea* L.): Broadening the gene pool of a monophyletic polyploid species.** Genetics 159:823-837 (2001).
- CALENGE F, Van Der LINDEN CG, Van De WEG, SCHOUTEN HJ, Van ARKEL, DENACE EC, DUREL CE. **Resistance gene analogues identified through the NBS-profiling method map close to major genes and QTL for disease resistance in apple.** Theor Appl Genet 110: 660–668 (2005).
- COLLINS NC, WEBB CA, SEAH S, ELLIS JG, HULBERT SH, PRYOR A. **The isolation and mapping of disease resistance gene analogs in maize.** Mol Plant Microbe Interact 11:968–978 (1998)
- COLLINS NC, PARK R, SPIELMEYER W, ELLIS J, PRYOR T. **Resistance gene analogs in barley and their relationships to rust resistance genes.** Genome 44:375–381 (2001).
- CRESTE S, TULMANN A, FIGUEIRA A. **Detection of single sequence repeats polymorphisms in denaturing polyacrylamide gels by silver staining.** Plant Mol. Biol. Reporter 19:299-306 (2001).
- DANGL JL, JONES JD. **Plant pathogens and integrated defense responses to infection.** Nature 411:826–833 (2001).
- DESLANDES L, OLIVIER J, THEULIERES F, HIRSCH J, FENG DX, BITTNER-EDDY P, BEYNON J, MARCO Y. **Resistance to *Ralstonia solanacearum* in *Arabidopsis thaliana* is conferred by the recessive RRS1-R gene, a member of a novel family of resistance genes.** Proc Natl Acad Sci U S A. 19,99(4):2404-9 (2002).
- DONALD TM, PELLERONE F, ADAM-BLONDON AF, BOUQUET A, THOMAS MR, DRY IB **Identification of resistance gene analogs linked to a powdery mildew resistance locus in grapevine.** Theor Appl Genet 104:610–618 (2002).

- DOYLE JJ, DOYLE JL. **Isolation of plant DNA from fresh tissue.** Focus 12:13-15 (1991).
- FALQUET L, PAGNI M, BUCHER P, HULO N, SIGRIST CJ, HOFMANN K, BAIROCH A. **The PROSITE database, its status in 2002.** Nucleic Acids Res. 1,30(1):235-8. (2002).
- FÁVERO AP. **Cruzabilidade entre espécies silvestres de *Arachis* visando a introgressão de genes de resistência a doenças no amendoim cultivado.** Tese de Doutorado. ESALQ, Piracicaba (2004).
- FERNANDEZ A, KRAPOVICKAS A. **Cromosomas y evolución en *Arachis* (Leguminosae).** Bonplandia, 8:187-220 (1994)
- FERREIRA ME, GRATAPAGLIA D. **Introdução ao uso de marcadores RAPD e RFLP em análise genética.** Brasília: Embrapa Cenargen, 220p (1995).
- GALGARO L, LOPES CR, GIMENES M, VALLS JFM, KOCHERT G. **Genetic variation between species of sections *Extranervosae*, *Caulorrhizae*, *Heteranthae*, and *Triseminatae* (genus *Arachis*) estimated by DNA polymorphism.** Genome 41:445-454 (1997).
- GARCIA GM. **Evaluating efficiency of germplasm introgression from *Arachis* species to *A. hypogaea* L.** Ph.D. Dissertation, N.C. State Univ., Raleigh, 197 pp. (1995).
- GARCIA GM, STALKER HT, SHROEDER E, KOCHERT G. **Identification of RAPD, SCAR and RFLP markers tightly linked to nematode resistance genes introgressed from *Arachis cardenasii* to *A. hypogaea*.** Genome, 39: 836-845(1996).
- GREGORY WC, KRAPOVICKAS A, GREGORY MP. **Structure, variation, evolution, and classification in *Arachis*.** In: advances in Legume Science Edited by: Summerfield RJ, Bunting AH. Kew, England: Royal Botanical Gardens, 469-481 (1980)
- GUIMARÃES PM, JOSÉ ACVFI, PROITE K, BERTIOLI DJ, LEAL-BERTIOLI SCM. **Desenvolvimento de marcadores moleculares para análogos a genes de resistência em *Arachis* spp. silvestres.** Fitopatologia Brasileira 30:663-667 (2005).
- HALWARD TM, STALKER HT, KOCHERT G. **Development of an RFLP linkage map in diploid peanut species.** Theor Appl Genet 87:379-384 (1993).
- HALWARD TM, STALKER HT, LARUE EA, KOCHERT G. **Genetic variation detectable with markers among unadapted germplasm resources of cultivated peanut and related wild species.** Genome, 34: 1013-1020 (1991).
- HAMMOND-Kosack KE, TANG S, HARRISON K, JONES JD. **The tomato Cf-9 disease resistance product *avr 9* gene functions in tobacco and potato to confer responsiveness to the fungal avirulence gene.** Plant Cell 10:1251-1266 (1998).
- HASH CT, BHASKER Raj AG, LINDUP S, SHARMA A, BENIWAL CR, FOLKERTSMA RT, MAHALAKSHMI V, ZERBINI E, BLÜMMEL M. **Opportunities for marker-assisted selection (MAS) to improve the feed quality of crop residues in pearl millet and sorghum.** Field Crops Research 84: 79 – 88 (2003).
- HAYES AJ, SAGHAI MAROOF MA. **Targeted resistance gene mapping in soybean using modified AFLPs.** Theor Appl Genet 100:1279–1283 (2000).
- HERSELMAN L, THWAITES R, KIMMINS FM, COURTOIS B, Van Der MERWE PJA, Seal SE. **Identification and mapping of AFLP markers linked to peanut (*Arachis hypogaea* L.) resistance to the aphid vector of groundnut rosette disease.** Theor Appl Genet 109:1426–1433 (2004).
- HERSELMAN L. **Genetic variation among Southern African cultivated peanut (*A. hypogaea* L.) genotypes as revealed by AFLP analysis.** Euphytica, 133:319-327 (2003).

- HULBERT SH, ILOTT TW, LEGG EJ, LINCOLN SE, LANDER ES, MICHELMORE RW. **Genetic analysis of the fungus, *Bremia lactucae*, using Restriction Fragment Length Polymorphisms.** Genetics 120:947–958 (1988).
- JENCZEWSKI E, GHERARDI M, BONNIN I, PROSPERI JM, OLIVIERI I, HUGUET T. **Insight on segregation distortions in two intraspecific crosses between annual species of *Medicago* (Leguminosae).** Theor Appl Genet 94:682–691 (1997).
- JONES DA, JONES JDG. **The role of leucine-rich repeat proteins in plant defenses.** Adv Bot Res, 24:90-167 (1997).
- KRAPOVICKAS A, and GREGORY WC. **Taxonomia del género *Arachis* (Leguminosae).** Bonplandia, 8: 1-186 (1994).
- KOBE B. and DEISENHOFER J. **The leucine-rich repeat: a versatile binding motif.** TIBS 19: 425-430 (1994).
- KOCHERT G, HALWARD TM, BRANCH WD and SIMPSON CE. **RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species.** Theor. Appl. Genet., 81: 565-570 (1991).
- KOCHERT G, STALKER HT, GIMENES M, GALARO L, LOPES CR and MOORE K. **RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae).** Am. J. Bot., 83:1282-1291(1996).
- LANDER ES, GREEN P, ABRAHAMSON J, BARLOW A, DALY MJ, LINCOLN SE, ETOH J. **MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations.** Genomics 1:174–181 (1987).
- LAWRENCE GJ, FINNEGAN EJ, AYLIFFE MA, ELLIS JG. **The L6 gene for flax rust resistance is related to the *Arabidopsis* bacterial resistance gene RPS2 and the tobacco viral resistance gene N.** Plant Cell 7:1195–1206 (1995).
- LEAL-BERTIOLI SCM, GUIMARÃES PM, BRUZZI LION M, CARNEIRO, RMDG, VALLS, JFM, BERTIOLI, DJ. **Busca de resistência ao nematóide das galhas *Meloidogyne* sp. e seqüências análogas a genes de resistência em acessos silvestres de *Arachis*.** Boletim de pesquisa 20 pg 5-17 EMBRAPA (2000).
- LEAL-BERTIOLI SCM, GUIMARÃES PM, FÁVERO AP, MORETZSOHN MC, PROITE K & BERTIOLI DJ. **Amendoim selvagem: uma fonte de resistência a pragas.** Biotecnologia Ciência e Desenvolvimento. Site: www.biotecnologia.com.br (2004).
- LEISTER D, BALLVORA A, SALAMINI S, GEBHARDT C. **A PCRbased approach for isolating pathogen resistance genes from potato with potential for wide application in plants.** Nat Genet 14:421–429 (1996).
- MAUGHAN PJ, SAGHAI MAROOF MA, BUSS GR, HUESTIS GM. **Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near-isogenic line analysis.** Theor Appl Genet 93:392–401 (1996).
- MEYERS BC, CHIN DB, SHEN KA, SIVARAMAKRISHNAN S, LAVELLE DO, ZHANG Z, MICHELMORE RW. **The major resistance gene cluster in lettuce is highly duplicated and spans several megabases.** Plant Cell, 10:1817-1832 (1998).
- MEYERS BC, DICKERMAN AW, MICHELMORE RW, SIVARAMAKRISHNAN S, SOBRAL BW, YOUNG ND. **Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily.** Plant J 20:317–332 (1999).
- MEYERS BC, KOZIK A, GRIEGO A, KUANG H, MICHELMORE RW. **Genome-wide analysis of NBS–LRR-encoding genes in *Arabidopsis*.** Plant Cell 15:809–834 (2003).

MEYERS BC, DICKERMAN AW, MICHELMORE RW, SIVARAMAKRISHNAN S, SOBRAL B, YOUNG ND. **Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily.** *Plant J* 20:317–332 (1999).

MILLIGAN SB, BODEAU J, YAGHOUBI J, KALOSHIAN I, ZABEL P, WILLIAMSON VM. **The root knot nematode resistance gene Mi from tomato is a member of the leucine zipper, nucleotide-binding, leucine-rich repeat family of plant genes.** *Plant Cell* 10:1307–1319 (1998).

MORETZSOHN MC, HOPKINS MS, MITCHELL SE, KRESOVICH S, VALLS JFM, FERREIRA ME. **Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hypervariable regions of the genome.** *BMC Plant Biol* 4:11. <http://www.biomedcentral.com/1471-2229/4/11> (2004).

MORETZSOHN MC, LEOI L, PROITE K, GUIMARÃES PM, LEAL-BERTIOLI SCM, GIMENES MA, MARTINS WS, VALLS JFM, GRATTAPAGLIA D, BERTIOLI DL. **A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (Fabaceae)** *Theor Appl Genet* (2005)

MUELLER UG, WOLFENBARGER LL. **AFLP genotyping and fingerprinting.** *Trends Ecol Evol* 14:389–394 (1999).

MULLIS KB, FALOONA FA. **Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction.** *Methods Enzymol.* 155:335–350 (1987).

NELSON S C, STARR J L, SIMPSON C E. **Expression of resistance to *Meloidogyne arenaria* in *Arachis batizocoi* and *A. cardenasii*.** *Journal of Nematology*, College Park, US, V. 22, p. 242-244 (1990).

PROITE K, DIAS JGO, GUIMARÃES PM, BERTIOLI DJ, CARNEIRO RG, VALLS JFM, LEAL-BERTIOLI SCM. **Avaliação de resistência a fitonematoides e fungos em diversas espécies de *Arachis* Silvestre.** VI encontro Talento Estudantil, Embrapa (2000)

QUILLET MC, MADJIDIAN N, GRIVEAU Y, SERIEYS H, TERSAC M, LORIEUX M, BERVILLÉ A. **Mapping genetic factors controlling pollen viability in an interspecific cross in *Helianthus* sect. *Helianthus*.** *Theor Appl Genet* 91:1195-1202 (1995).

SAIKI RK, GELFAND DH, STOFFEL S, SCHARF SJ, HIGUCHI R, HORU GT, MULLIS KB, ERLICH HA. **Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase.** *Science* 239:487-491 (1988).

SEIJO GJ, LAVIA GI, FERNANDEZ A, KRAPOVICKAS A, DUCASSE E, MOSCONE DEA. **Physical mapping of the 5s and 18s-25s rRNA genes by fish as evidence that *Arachis duranensis* and *A. ipaensis* are the wild diploid progenitors of *A. hypogaea* (leguminosa)** *American Journal of Botany* 91(9): 1294-1303 (2004).

SIMPSON CE, KRAPOVICKAS A and VALLS JFM. **History of *Arachis* including evidence of *A. hypogaea* L. progenitors peanut.** *Science* 28: 78-80 (2001).

SORIANO J M, VILANOVA E S, ROMEIRO C, LLÁCER G, BADENES M L. **Characterization and mapping of NBS-LRR resistance gene analogs in apricot (*Prunus armeniaca* L.)** *Theor Appl Genet* 110: 980–989 (2005).

SOUTHERN EM **Detection of specific sequences among DNA fragments separated by gel electrophoresis.** *Journal of Mol. Biol.* 98:503-57 (1975).

STADEN R, JUDGE DP, BONFIELD JK. **Managing sequencing projects in the GAP4 environment.** In: Krawetz SA, Womble DD. *Introduction to Bioinformatics. A Theoretical and practical approach.* Human Press Inc, Totawa (2003).

STALKER HT. **A new species in section *Arachis* of peanuts with D genome.** *Am J Bot*, 78:630-637 (1991).

STALKER HT. **Peanut (*Arachis hypogaea* L.)** *Field Crops Research* 53 205-217 (1997).

SUBRAHMANYAM P, NAIDU DH, SIMPSON CE. **Resistance to *Didymella arachidicola* in wild *Arachis* species.** *Oleagineux*, V. 40 n. 11 p. 553-556 (1985).

TAMELING WIL, ELZINGA SDJ, DARMIN PS, VOSSEN JH, TAKKEN FLW, HARLING MA, CORNELISSEN B. **The tomato R gene products I-2 and Mi-1 are functional ATP binding proteins with ATPase activity.** *Plant Cell* 14:2929-2939 (2002).

THOMAS C M, JONES DA, PARNISKE M, HARRISON K, BALINT- Kurti PJ. **Characterization of the tomato Cf-4 gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognition specificity in Cf-4 and Cf-9.** *Plant Cell* 9: 2209-2224 (1997).

UPENDER M, RAJ L, WEIR M. **Rapid method for the elution and analysis of PCR products separated on high-resolution acrylamide gels.** *Biotechniques* 18:33-34 (1995).

VALLS JFM, SIMPSON CE. **Taxonomy, natural distribution and attributes of *Arachis*.** In: Kerridge, PC, Hardy, B, *Biology and agronomy of forage *Arachis**. Centro Internacional de Agricultura Tropical. p. 1-18 (1994).

Van der BIEZEN EA, JONES JDG. **Plant disease resistance proteins and the gene-for-gene concept.** *Trends Biochem Sci*, 23:454-456 (1998).

LINDEN CG Van Der, WOUTERS DC, MIHALKA V, KOCHIEVA EZ, SMULDERS MJM, VOSMAN B. **Efficient targeting of plant disease resistance loci using NBS profiling** *Theor Appl Genet* 109:384-393 (2004).

VOS P, HOGERS R, BLEEKER M, REIJANS M, VAN DE LEE T, HORNES M, FRIJTERS A, POT J, PELEMAN J, KUIPER M, ZABEAU M. **AFLP: a new technique for DNA fingerprinting.** *Nucleic Acids Res* 23:4407-4414 (1995).

WANG ZX, YANO M, YAMANOUCHI U, IWAMOTO M, MONNA L, HAYASAKA H, KATAYOSE Y, SASAKI T. **The Pib gene for rice blast resistance belongs to the nucleotide binding and leucinerich repeat class of plant disease resistance genes.** *Plant J* 19:55-64 (1999).

WAUGH R, McLEAN K, FLAVELL AJ, PEARCE SR, KUMAR A, THOMAS BBT, POWELL W. **Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (SSAP).** *Mol Gen Genet* 253:687-694 (1997).

WHITHAM S, MCCORMICK S, BAKER B. **The N gene of tobacco confers resistance to tobacco mosaic virus in transgenic tomato.** *Proc Natl Acad Sci USA* 93:8776-8781 (1996).

XU Y, ZHU L, XIAO J, HUANG N, MCMCOUCH SR. **Chromosomal regions associated with segregation distortion of molecular markers in F2 backcross doubled haploid and recombinant inbred populations in rice (*Oryza sativa* L.).** *Mol Gen Genet* 253:535-545. (1997).

YOUNG ND. **The genetic architecture of resistance.** *Curr Opin Plant Biol* 3:285-290 (2000).

YU YG, BUSS GR, MAROOF MA. **Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site.** *Proc Natl Acad Sci USA*, 93:11751-11756 (1996).

YUKSEL B, ESTILL JC, SCHULZE SR, PATERSON AH. **Organization and evolution of resistance gene analogs in peanut.** *Mol Gen Genomics* (2005).

Documentos e Fontes em Meio Eletrônico

Home Page EXPASY (Ferramenta TRASLATE e MotiScan) Disponível em:
<<http://www.expasy.ch/prosite>>. Acesso em: 11Dez. 2005

Programa JalView 2.07 Disponível em:
<<http://www.jalview.org>> Acesso em: 20 Jan. 2006

Home Page FAO (Food and Agriculture Organization) Disponível em:
<<http://www.Fao.org>> Acesso em: 25 Fev. 2006

Home Page EBI (European Bioinformatics Institute) – Ferramenta ClustalW. Disponível em:
<<http://www.ebi.ac.uk>> Acesso em: 15 Jan. 2006

Imagens *Arachis hypogaea*:

<<http://plantatlas.usf.edu>> Acesso em: 28 Dez. 2005

Anexo 1

Protocolos, Soluções e Reagentes utilizados no trabalho.

Extração de DNA

Tampão CTAB 2X

2% de CTAB
100 mM de Tris-HCl (pH 8,0)
20 mM de EDTA (pH 8,0)
1,4 M de NaCl
1% de PVP 40

Tampão de precipitação CTAB (1X)

1% de CTAB
50 mM de Tris-HCl (pH 8,0)
20 mM de EDTA (pH 8,0)

Tampão de corrida TBE (10X)

108 g de Tris-base
55 g de ácido bórico
40 ml de EDTA 0,5 M (pH 8,0)

Southern-blot

Tampão SSC 20X (1L)

175,35 g de NaCl (ou 3M)
88,23 g de Citrato de Sódio (ou 0,3M)

Solução de Denhardt's 50X (100 ml)

1 g de BSA
1 g de Ficoll 400
1 g de PVP 40 T
Dissolver em Água e estocar a -20°C.

Solução de Esperma de Salmão 4 mg/ml (10 ml)

40 mg de esperimidina
H₂O para completar 10 ml

Solução de Pré-hibridização (200 ml)

50 ml de 20X SSC
20 ml de 50X Denhardt's
10 ml de 10% SDS
960 µl de Esperma de Salmão (4mg/ml) Desnaturar e adicionar somente na hora do uso.

Solução I de Lavagem (500ml) (1X SSC; 0,1X SDS)

25 ml de SSC 20X

5 ml de 10% SDS

Solução II de Lavagem (500ml) (0,1X SSC; 0,1X SDS)

2,5 ml de SSC 20X

5 ml de 10% SDS

RGA-Display**5 X Restriction/Ligation Buffer (1 ml)**

500 µl de 10X All-Phor-one Buffer

25 µl de 1M DTT

25 µl de 10 µg/µl BSA

450 µl de H₂O**Gel de Poliacrilamida****Gel 4 % Desnaturante (250 ml)**

105 g de Uréia

121 ml de H₂O destilada

25 ml de 10X TBE

25 ml de Acrilamida:bis (19:1) 40% (Solução comercial)

Misturar em um béquer de vidro os reagentes na ordem que aparecem, e agitar com um agitador magnético com temperatura de 55°C até a total dissolução da uréia. Filtrar a solução em filtro Milli-Pore (20) e acrescentar o 166,5 µl de TEMED. Guardar a 4°C.

Preparação do gel na hora do uso (30 ml)199,9 µl de 10 % APS (1 g de Persulfato de Amônio em 10 ml de H₂O)

30 ml de Gel 4% previamente preparado

Tampão de Amostra para gel desnaturante (Formamida Deionizada) (Xu *et al.*, 1997).

Formamida Deionizada 95% (v/v)

EDTA 10 mM (pH 8,0)

Xileno Cianol 0,09%

Azul de Bromofenol 0,09%

Placa de Bind (placa menor)

995 µl de Etanol Absoluto

5 µl de Ácido Acético Glacial

3 µl de Bind Silane

Lavar a placa com água e sabão e depois três vezes com álcool. Secar e espalhar a solução de Bind. Secar por 5 minutos e lavar três vezes com álcool.

Placa de Repel (placa maior)

1 ml de Repel Silane

Lavar muito bem a placa com água e sabão, secar e lavar três vezes com água destilada. Secar e espalhar 1 ml de Repel pela placa. Secar por 10 minutos e lavar uma vez com água destilada. Secar bem.

Montagem do Gel Desnaturante

Colocar espaçadores apropriados na placa maior e colocar a placa menor por cima. Prender bem com grampos de metal (ou colocar fita) e inclinar o conjunto de placas um pouco. Verter o gel (já com o catalisador APS) lentamente e voltar as placas à posição horizontal. Colocar o pente e prender com um grampo. Esperar 2 horas para a total polimerização do gel.

Pré-corrída

Após a polimerização, colocar as placas sem os grampos na cuba de eletroforese e deixar correndo por 40 minutos (ou até que a temperatura atinja 50°C) a 45W. O tampão de corrida deve ser o mesmo utilizado na preparação do gel (TBE 1X).

Preparação das amostras

Adicionar 50 % do volume do produto de PCR de Tampão de amostra com Formamida (se a PCR foi amplificada em 10 µl adicionar 5 µl de tampão de amostra). Desnaturar as amostras por 5 minutos a 95°C e colocar no gelo imediatamente. Aplicar as amostras no gel aquecido o mais rápido possível para evitar a re-naturação do DNA.

Coloração com Nitrato de Prata (Creste et al)**Solução de Fixação (1L)**

890 ml de H₂O destilada

50 ml de Etanol absoluto

10 ml de Ácido Acético glacial

Colocar o gel após a corrida e agitar lentamente por 10 minutos.

Lavagem

Lavar o gel por um minuto em H₂O destilada agitando lentamente

Solução de Oxidação

985 ml de H₂O destilada

15 ml de Ácido Nítrico 37%

Agitar o gel lentamente por 3 minutos

Lavar o gel por um minuto em H₂O destilada agitando lentamente

Solução de Nitrato de Prata

1000 ml de H₂O destilada

2 g de Nitrato de Prata

Agitar o gel lentamente por 20 minutos no escuro

Lavar o gel 2 vezes de um minuto em H₂O destilada agitando lentamente

Revelador

990 ml de H₂O

30 g de Carbonato de Sódio

540 ul de Formaldeído 37% (adicionar somente na hora do uso)

Agitar o gel lentamente na solução gelada até que as bandas apareçam

Solução Stop

950 ml de H₂O destilada

50 ml de Ácido Acético

Agitar lentamente o gel por 5 minutos. Lavar o gel por um minuto em H₂O destilada agitando lentamente e deixar o gel secar completamente (aproximadamente 24 horas).

Anexo 2

Seqüências FASTA de nucleotídeos de todos os marcadores RGA-Display seqüenciados no trabalho

>AD-P212MCT-9B RGA.0.17
GGAAGATTTATGTGGAGACATACATGTTTGGGGCAAGTAGCAGAATCATAGTGACAACAAGAGATAAGCATGTCCTTGTACGGCGGACGCA
GATCATAT

>AD-NBSMCTC-14C RGA.0.6
TGGATGATGTGTGGTGGGGCATTGTGGAATCAAAGGGGACTCAGGAGTATGATTGGTCAGCATATAGAGACTTGTTCACCTTGTTCACATG
CATGCTCAAGTTACAGGCCATAATGTCATTTACCAGTGTGGTGGCAATGTGGGAGATTCCGTTTCCATCCCAATCCCTAAATGGGTCCCTTGAA
ATCGGAGTATCAGACCTGATATCTTCTACACCAACCCCAATGGCATAAGGAACAAGGAATGTTTGTCCATGGAGTTGACAATAGGCCTTTAT
TCCATGGCAGAGCTGCTATTGAGGTTGACCATGTTCAATAATTTGTGTTTTCTTACTCATATTTCTTCTATTGTCGCTAACCTTGGAGAACATG
AGTTACTCA

>AD-NBSMGGC-8B RGA.0.9
TCGTGGATGATGTGTGGATTGATAATTCAGAACTGGGAAAGTCTTTCATAAACCTTTTCAATATGGGAAAAAGGGAAGCAAGGTTATTGTAA
AACTCGAAATGATAATGTTCCCGATGTAGTGTAACTATTTCCAGCATACAAGCTTGGTTTATTGTGAGAGAAAGATGTTGGTCACTTTTTGTG
AAACGTGTATTTCTTCTACTGCATCGAAGGAGTATTCACCTCTATAATCGGTTGGTCGAGAACTTGTCAAAAAGTGCAAAGGGCTACCTTTGG
CGGTAGAGGCCCTTGGTTCCTTATTACGTACAATAAATAATGATGAAAGGATGGGATAGTGTATGAAAAGTGAACCTTTGGGAAGTTTTTGA
AGATCAGAATGACGAGATTATTCCTGCCTTACTCAGGACTCATC

>AD-NBSMGGC-8G RGA.0.15
TCCAATACTGTAACCTTTGTGTCTTACTCAAGA

>AD-P212MCT-9B RGA.0.16
GGTGGTCATGAGGATTTGGTTATTGGCCCCAAACCGG

>AD-P219MCTA-12B RGA.0.21
TGAGTCCCTGAGTAACTAAAACTAATGAAAATATACTAAAACTAATATATCATACTAAAAACATACTAAAAACAATGCCAAAAAGCGTACAA
ATTATCCGCTCATCAGGCAACCTCCTCATACGTATCAAGCATTGTTGATTTCTGCCTATTGCCTTTGCACTTGTGGAGGGAGAAAATACAGAGTC
TTGGTCATACTTTCTGACCAACCTCAGAAGACATGTGACGCCACAATGGGCATTCTGGTGATTTCTGACAGGCACAACGGCATCAAGACAGGA
CTAGAGACACCTGATAGTGGGTGGCAACCTCCTCATACGTATCAAGCATTGTTGATTTGGCATGTTGACGCAATTTGCTCTCAGCTTCAAGG
AGCAGGATGTTAGGAGGAGGCTGGTGAATGCTGCTTATGCTAAAACGGAGGCCGATTTTGAGTATCGGTTGACATCATCCACAACAATA

>AD-P219MCTA-12D RGA.0.23
TGAGTCCCTGAGTAACTACTTGTGACCAAGTATACATCATATTTCTCCGAGTTATAAATATTTCTTCTCTAGTTTTCTTCTTGTCTTCCGTCACA
TGTTACTAAAGCAACAACATCACTTGATATTTCTTTCTGAAAGAAACAGAAACAAGTCAAATAAGAATACCTATGGGCCTATTGGTTCGTGAT
TGTGGCCCATAAATCAAATAACCACATCCACATAGCCATCGTGGTCGACATCATCAAGAACAATAAG

>AD-P219MCTA-12C RGA.0.22
TTGCTGATGATAGTGGCGATGATATTGGACCGAGTGTGATGAATGGATTTTTGACGGTTTAGAATTTCTCAAATAAAATCTTGTGCAAGTAT
AGTATCTAAACCAAGCAATAATCCTTTTCATACAAAAAGTTGTTTGTCACTAAAACAAAACCCCTAAATATATAAACCGAAGTATTCAAACCTCGG
GTCATTTCTCCTAGGAATTGTAATGAAGTGTGTTGTAATGGTTGAGTTATTTTTGGGGTTTGATAAGAGGCATGAAAGTAAATGGCAATGAA
GATAAACTAACTATAAAAGGCTCTTGGCAAGGTATGAAAATTAGAAGTCTATCCTAGTTATCTTCTCAATTTGTGATGAGAATTTGTTTCAAT
TGCTACCCTTAGTTACTCAGGACTCATCAANCAATTGAGAAGGATAACTAGGATAGGACTTCTAAT

>AD-P2378NCTC-11D RGA.0.38
TGAGTCCCTGATTAACCTCTCGGCAACAGGTACTTATCTTTTATTTGTGACTTTGTTTAGCTGCCTGTAATGCACACAGAGCCGCATACTCCTGTCC
TTTTCTTTACTAGTAGACACCGGGCACCCTATGGAAAAACACTTGGTCGGATAAAAATTTTACCTAATAGATCCTCCAAGTGAGACTTTAGCT
CGGCCATTTCCAGCGGTGACATCCTGTAAGGAGCACTCGAGATTGGTCCAGCCCCAGGCACCAACTCAATAGCAAACTCAACCTCTCAGTTAGG
TGAAAATTCATCAATGTCATCTAGAATCAAATAAATTT

>AD-P2378NCTC-11E RGA.0.35
TTATTTTGAATCTAGATGACATTGAGGGACTTGTGCAAAAATCTGATTCAGAGGCTGAAAAGGACTGCAGATGCTGTGGATTCTGACCTCCTT
GCACTCGAAGTGGATTTTCTAGAATCTACAGAAACCAATTTGGCAGCAGTCAATTTGCGTTGGAAAAGTAAACATCCTTGGCTTTCCAGCAATAT
ATAATAGTCCATACTTTGCCCGAGATTTGATGGCCAAACAGGCGTTCATGTGAGCTCAAGAATTTCTGGCGTAAAAACCCGGAAGTGGCAAAA
GAATGGGAGTTACTCAAG

>AD-P2378NCTC-11F RGA.0.40
GAGTCTGTAGTAACTCTTGTGTTGTGTTTCAGACACCTCCGAGTTCAATATTTATGACCTCTTTTTCTTTTGTACAACCGGCTAGGTCCAACTTT
CATTATAACACCTCCTTGCCAATTTCTACTCTCCTTCTACAGTACTATTCTATCCGAAAGTCAAGAAATTTTATGCAAGTAAAGTAAAGC
GACTATGGCAAGTCCGTTCAATGTCATCTAGAATCAAATAAATTTA

>AD-P2391MCTA-13D RGA.0.43
CAGCATGAAGGTATGGATATCCTTCAAACATTTGGTGTGACTTGGTAAGCATATCCAGTCTCTGCATCCACACAAAATACTATTTCCCTT
GCAAGCGTTGGTACTCTACTCAAAGATTTCTTACATATCTCGTTTCCAACACTCCAATTTGAACACCACAGGGAATATCTGATATGGTAGATTC
TCCACGTCGTCAGCACCATCAGCG

>AD-P2391MCTA-13E RGA.0.45
GCCACATTACTTCAGCCTGTAT

>AD-P2414MC-17B RGA.0.49

GTCCTGAGTAACACACGTTTCTAAGACTTTGGCCAAAGTTCTCTAAACAACATCAAATGAAGTCTCATAACACTGAAGTCATCCATAAAGACT
TCCAAACAACCTCTCAATAAGATCGAAAAGACACTAGTCTATGCACCGCTGAAAAGTGGCGGGTGCATTGCATAGCCCAAATGGCATCTTTTGT
AAGCAAAGGTGCCAAAGGGACAAGTAAATGTGCTCTTTTCTTAATCTCATGAGCAATATGTATCTGGAAGTAAACCAGTATAACCATCAAGAAA
GCAATAATGGGATTTACCC

>AD-P2414MCTA-15B RGA.0.55

TGAGTCCTGAGTAACTAGTACCAAGTGACATGGATGGCTACCAAGAGGTCTACGTGAGCGAGAAAAGTCTGGATTTGACTGGGACTTGATCCTC
CAAGTCATTTGCTGAACCAGAAGCATTTTGGACCCATGGTCAACTCCGGATGCGACCCAAGTGCATTGACACACATTTTTTCACTATAAAGGCTA
GAGCTTGGGCCAGATCCTATCTCACTATGTCTACCTAGCACCCACAGGTCGTCCAGCATCAGCAC

>AD-P2414MCTA-15A RGA.0.52

CTGATGCTGGACGACCTGTGGGTGACACCGTGCCTCTATGGAACCTTGTCTGCACTTTGGGATCATTTGGCAACTGGAATGAGGACACTCTTAT
ATTGGCTCTATTATTCATCAAGAACAGACCCGCATGTTCAACACGTCCGGTGCCGGAGGTTCTCTTTCAGGTGCCACAGCAGCTGCCCGGTC
CACTCCTCTACATCCCTCTGCAAATGAGGGGTAAAGTCAAGGTAATTTATGTTTACCGCATTTCTATGAACCTGTTTGAACACTCATTTGTTGT
AGTTACTCAGGACTC

>AS-NBSMGGC-8E RGA.0.68

ATGAGTCCTGAGTAAAGGCTGAAACTGGTTATCTCCTTTGGTAGCATTTTCTCCTGCCATACCTGCACAAACGATTTAGTTGAAAAAATACCTTG
CCCATCATATTTCCACACCACTCTATCCTCTATGTCTGTACTAGCCTCACAGGTCTCAAATGTATGCAACTGGTTCAGTAGTTCTAATTC
CATTTGGAATAGCTCACGCTTCCATTTGAAAGTTCAAATCCACTCTAACCCATCCCAAAACCCACAAGCCCAATAACAGAACCCTCTGGTTTG
AAACAGAGCAGAGCCTCGGAAATCGATCTTTTATAGCTCCACCAAGAAGCCACACATCATCCACGA

>AS-NBSMGGC-8H RGA.0.74

ATGAGTCCTGAGTAAAGGCAACAAAGGAACAGTATTGGGAGGCCGAGCCGAAAGCTGAAGCTGGTTTGGTTGAGGTGCCTAGGTTAGGT
GAGGGTAAGAGGTTTAGGTTAGTGTGGGAAATGAAGGAATGAATGTAGATTGAAGCTTGGGCTGTGTGTACTTCTAGATTGTTGTGTA
TATGTTTATCCACACATCATCCACGACAACCA

>AS-NBSMGGC-8I RGA.0.70

GGAGTGATCATGATTACAACAGGAGAGAGGGGTGTAGCTGATCTCTTGCCTGGTATCAGGCAAGGTTTCATAGCTA

>AS-NBSMGGC-8I RGA.0.75

GTGATCATGATTATAACAAGAGAGCGGGTGTGCTGATTCTTGTGTGGTATCAGTCAAGGTTCTTAGCTACC

>AS-P219MCTA-12A RGA.0.84

TGAGTAACTATCATCTTAATCACAACGTTGTACATAGTAGGACTACACGTAGGAATTTGGAGGCCGCGACCCAGTAACTCTTAGCCTAGTGGGG
ACGATCCCAACTCCGATACCTCAGTCGTAGATCCTGATAGGGTTTGGGGTGAAGCCCTCTGAGCCCTACAAGAATCGCGTCCACAGGTTGGG
GTTGCTCTTCCGACGAGCCTCCGCACTTCCATATTTGGCGGCTTCTATCTGCTCTGCCTCTGCCACCAGTCTGCCAATTTCCGAGGAAGTTGTC
GATTTGAGGGAGGAGATGCATAAGCTCACACAGGAGCTTGACCAACAAGTTGAGCAGTCAGAGGAGAGGTACAATGAACCTCTTACACACATGG
GAGACACCGTTGCCATAACGGCGGAGCTGACAGAGAAGCTGGAGCGGTTAGAACCTTTGTGAGACTAGATAGTGGTATACAACGAGTAGATGCA
CGCTGGAGGCAGCAACGCTACTGGTACCAGCGACGGTTCTGCTAGTGGATCATCGACATCATCAAGAACAATAA

>AD-P219MCTA-12G RGA.0.24

TGAGTCCTGAGTAACTTAAGACTAATAACAATATACTAAAACTAACTAGATTATAACAAAAATATACTAAAAACAATGCCAAAAAGCGTATAA
ATTATCCGCTCATCATTTGGCGGTGATGACAAACGCAAAAGAATCAATTGATTCATTTTCGACATCATCAAGAACAATAA

>AS-P219MCTA-12F RGA.0.88

TCCTGAGTAACTAAAACTAATAAAAAATATACTAAAACTAACTAGATCATACCTTAAACATACTCAAAACAATGCCAAACAAGTGTACAAATTA
TCCGCTCATCACTTATCTCTTGTGCTACTATGATTTCTGCTACTTGCCTCAACAGGATGTCTCCACATAAATCTTCCATGGGGTCTGAAT
CATCGACATCATCAAGAACAATAAAG

>AS-P219MCTA-12E RGA.0.87

TGAGTCCTGAGTAACTTACTACATAGGGTTAGAAATTTCTCCAATGGTACCATCTCAGGGCGAAGGAGGTCGAGTCTGAAGGGGCACTGCATCTACAT
GTTCCATCATCTCAGTTTCAAAATTTGAATATCGGGATCTGACTCCGGTTGAGCGATTTTCAGCCGAAGTGAAGGGGGGCGAAATATTAGAGGTT
TTTGAGAAGAAACAACAGAAAGGTCGACATCATCAAGAACAATAAGGACGNNNTCATCAAGAACAATAA

>AS-P2378MCTC-11C RGA.0.90

TGAGTCCTGAGTAACTCTTGGAGCAATGCATTCGAGTGTGCGAGAGGACAGACCCGAGCAGTACGCTTTTTTTCAGGGCATATGGAGCTCGAGATT
GGGTTGAAGTTTCTCAACCGGGAACAGCAATGCTGGCAGTCAAAAACACTACAACATCCGTAGGAGTGCAGAAATATAAGGTGGTAGAGTCAGACC
AAAGTAGGTATGTATGTCGATGCAAGCAGTTCGAGGATCAATGTCGTTGGATGGTACGGGTCGCGAAGACAAGGTCCTCAAGATTTTGGGAAAT
TCGAAAATACGAAGGCCTCATAGTTGCTTGGCAAGTTCGATGTCTCAAGATCACGCTCAACTTGACATCAATGTATCTAGAATCAAAAT

>AS-P2378MCTC-11H RGA.0.91

TATATGCAAGAGATG

>AS-P2391MCTA-13C RGA.0.92

AATCTTGGTTGATTTGACACCGGTAAGCATATCCACTCTCTGCATCCAGACAAGTACTATCTCCCTTGAAGCGTTGGTAGCTCTTCTCAAAAT
TTTTTTACATGTCTCATTTTCAACACTCCAATTGATGAACACCACAGGAAATATCTGGTATGGTAGATTCTCCACGTCGTCAGCACCAT

>AS-P2414MC-17A RGA.0.93

TTCCGATGCGGCTCGTGGGTAG

>AS-P2414MC-17H RGA.0.98
TGAGTCTGAGTAACGAAGGGTTCTGTCTGGTTTTCCGAATTTACAAAATTTCTCTCCCATCGTTCCTCTGAATGCTGAAACTGGTTTGCCTG
ATCCATAGTGCCTTACGCCGTTTACCCTGACTAGATGGG

>AS-P2414MCTC-10B RGA.0.104
TGCTGATGCTGGACGACCTGTGGTTGCCGGAAGACACTATTGTTGTCTACTGCTGAAAGAGGGGAAAAGCTGTGGCATTGTTCTGGCCGCCAGGAG
AAGTTATGAGCCCTCTTGAACACCCGCTGGAGCTTCTGGCTACTTCTGTCTGTTGTCAAAAAAGTTGCCAGTAAGGGTTTTATTTGAGGTTTTTG
TTCTTTTTGAATTTTGGAGTTACTATGGTCACTGCCTGCTGGTTTTCAGTCATCGTGATTGGAATTCATCGCCGCTGCCGCTTGGAGTGGCTGTC
GGGCTGCCGCCGAGCCGGTTCGAAGACCCGCGGTGTCGGTTACGCTTTCTCTCTCGATTTCGTTAAGCGTTTTTCGATTTGAAAGCCCTTAA
GTTACTGTTCTGTTATACATACGCTGAGGTTTGTGGAGTTACTCA
GGACTCA

>AS-P2414MCTA-15C RGA.0.101
AGTCTGAGTAACAAAACTAATAAAAAATACTAAAACTCACTAAATATACTAAAAACAATGCCAAAAAGCGTACAAATATCCGCTCAT
CAGCTTACAACCAACAAGTTCAACAAAAGTCCACAGGTCGTCCAGCATCAGCACGAACCAAGAATCTATATCATCTCTTTCATCATCAAAATAT
AAATTCATATATAGTTACTCAGGACTCA

>AS-P2414MCTC-10A RGA.0.105
TGAGTAACTCTCTCACCAGAATCTCTTTCATCTTTCAGGAAACACCCCAAGATATAGGAAGCATGGCTTCATTTTCTCAGACAGATCATCGTAGCT
AAGCTTCAATATCTCCATCATCTTCTTACCTTCTTGTCTTTCAGTAAAAGACCAGTAAGGGAGCAGGTTCTTGATTTCTTCCCATGCATCTCT
GATCTCTCCTTCTTCGCGACAATCCCGGCTGTGGTTTTGATAGCCAATGGTAAACCTTTGCAGCTTTTGAAGCAATGATCTTCCAATAAGCT
CTAAAGGAGGAGGACTCTTCTTGCCTAACACCTTATTTGCGGAATATTTCCCACTTTGATCTTCAACAAGTTTAGTTGGTGGTGGAGG
CTCCTTTGACCTTGCATAAATTCGCCACTTGATCATACGAGTAGTTACTAGTATCATGCTGCCATTGTTGTTATCCGGTAAAGGAGCCTTTAGC
TTGTCCTACATCGTTTGTCTCCACAGGTCGTCCAGCATC

>AS-P2414MC-17D RGA.0.97
GAGTAACTGACAGAAAACATGGAAAAGAAAACGAAGGTTGCCAAAAGGCAGCAGTCACTTTACTGATGAAGCTAGCATTGAAGATCCTATGGA
AATATCCCAAGTGTACCTAGTGAATAATGCAGAGAGTAGTGACTTCGTCACATCATGCATAGAACGAATGAATCAGCTCGAGTCAACCCCTGGTA
TTTTGGGAATATATTTGCACAAGATGTGAATGAAACCCCTTGGTCAAAGTCATTCAACTTGGATGTCGATGAATCCACAGGTCGTCCAGCATC

>AD-P2414MCTC-10D RGA.0.58
GAGTAACTGACAGAAAACATGGAAAAGAAAACGAAGGTTGCCAAAAGGCAGCAGTCACTTTACTGATGAAGCTAGCATTGAAGATCCTATGGA
AATATCCCAAGTGTACCTAGTGAATAATGCAGAGAGTAGTGACTTCGTCACATCATGCATAGAACGAATGAATCAGACTCGAGTCAACCCCTGGT
ATTTTGGGAATATATTTGCACAAGATGTGAAT

>AS-NBSMCTC-14C RGA.0.60
GTGAGGTTCCCTTTGAGCCTGCATAACATTTCTGTAGTTAGCAATATCCTGCTCCGTTATTTGTGTTTTGGGAGAAGATAATACTTCAGCTTGT
TCCACACCTGAGTG

>AS-NBSMCTC-14B RGA.0.59
GTGGCATCAAAGGGGCCCTCAGGAGTATGATGGTCAAGTATAGAGACTGTTTCAACTAGTTCAAGATTGGCAGCTCAAGTTACAGGCCATAA
TGTCATCCCAACAGTGTGGAGGCCATGTGGGAGATTCCGTTTCCATCCCAATCCCTAAATGGGTCCTTGAATCGGAGTATCAGACCCCTGATAT
CTTCTACCAATCCCAATGGCATAAGGAACAAGGAATGTTTGTCTATTGGGTTGACAATAGGCCTTTATTCATGGCAGACTGCTATTGAG
GTTGAACATGTCAAAATTTCTGTGCACTGACTCATATTTCTTATTTATTTGCTAACTATTTTCGGAGAACATGAGTTACTCAGG

>AD-P2414MC-17E RGA.0.51
TGAGTCTGAGTAACCAAGTACAGTGGCTACCAAGAGGCTACGTCAGCGAGAAAAGCTGGATTTGACTGGGACTTGTATCCGT
CCAAGTCATTGCTGAACCAAGAAGCATTGAGCCCATGGTCAACTCCGGATGCGACCCAAGTGCATTGACACACATTTTTTCACTATAAAGGCT
AGAGCTTGGGCCAGATCCTATCTCACTATGTCCTACCTAGCACCCACAG

>AD-P2414MC-17C RGA.0.50
CTGGAATGAGGACACTCTTATATAGGGCTCTATAATTCATCAAGAACAGACCCGCATGTTCAACACGTCGCGTGGCAGGTTTCTCTTACAGG
GCCACAGCAGCTGCCCGGTCCTCTCTACATCCCTCTGCAAGTGAAGGGTAAAGTCAGGGTAATTATGTTTACCGCATTTCTATGAACC
TGTTTGAACACTATTTGTTGTTGTTACTCA

>AS-P2414MC-17I RGA.0.99
GAGTCTGAGTAACAAAACTAATAAAAAATACTAAAACTCACTAAATATACTAAAAACAATGCCAAAAAGCGTACAAATATCGGCTCA
TCAGCTTACAAA

>AA-P2378MCT-20D RGA.0.107
CCGACCAAGTGTTCCTCCGTTGGGTT

>AD-NBS8MC-21E RGA.0.108
AATGCATTATAGTAAGGAGTTGTCGATGGATA

>AD-NBS8MC-21F RGA.0.109
GGGTCATCAGTGAATAAACTTTTTGCAAGTCATCAGATGGTGAAGGCTTTTCTCGGAAGTCGGCTCTTGGATCTAGGTCTGCCAGTGTGAGT
TGTTGTTGATGTTACTCAG

>AD-NBS8MGAG-22C RGA.0.110
ATATGCGCCTTGATCATCTCTGTAGACACCGTCGGCCTTCCCGCATTGGTCTCGGGATGGTAACGTTGTTGCTATCTTGTGAAGGATGCCAT
GTTCTTTCAGTAGGTTCTTACT

>AD-P2378MCT-20C RGA.0.113
TCAGACCCTTGCCTCGAAGTGGATTATCTAGAATACAGAGACCCAATTTGGCGGCACTCAATTTGGCTTGGAAAGTAAACATCCTTGGCTT
TCCAGCAATATATAATAGTCCATACTTTGCCCGAGATTTGATGGCCAAAACAGGCTTCCATGTCAGCTCAAGAATCTGGCGTAAAACAGCGG
AAATGGCACAAGAATGGGAGTTACTCAGGA

>AS-NBS8MC-21A RGA.0.118
ATGATGTGGTATAATCAGGGACCAAGCATATCTCTCTTCATGAGCGACTAGACCAAGAAATTGGCTATTGATCAAGATCTGAGAGATTGAGT
TACCAAGGGATTGGGGCTCAATCAATCATGATTGCCAAGAGGTCAATGAGTTGCATGATTGAAGAGGATATGAGCTGGATTCAATCCAAAGAGA
CAATATCTCTGATCTCAATGAATTCTCCCATTTTATAGCTTTTATATACATTTCTGCAAATTCATTCCCCTTACATTCCTGCTATTTTC
CATTTTCAGCACTTTATTGCTTTCTTTATTTCTTATGCCATTCATGTTTCTGTAATTTATAATTTCTGCATTTATAACTTACTCTGTACTCAGG

>AS-NBS8MC-21C RGA.0.119
TTCATGAGCAATTAGACAAGGATTGGCTATTGATCAAGATCTGAGAGATTGAGTCACCAAGGGATTGGGGCTCAATCGATCATGATTGCCAAGA
GGTCAATGAGTAGCATGATTGAAGAGGATATAAGCTGTAATTTGATCCAAAGAGACAACATCTCCTGATCTCAATGAATCTCCCAATTCCTATCT
ACCATTTTCTTTATAGCTTGCTAGAACATTCTGCAATACCCCATTTCTCATATACAGTTACTCA

>AS-NBS8MC-21H RGA.0.120
TGAAAGAGATGCGAGATCAAGGAAGTAAAGTAAAGAAATAGTGGTTGAAAAAGTCTGGACTCTGAGTGAAGTGAATGGATCCATTGTATAC
TAAAAATACTTTTGAAAACTAATGTAATGTTTACTGTTTACTGAGGATATAAGACGTTATGCGCCAGCAGGGACGGTGGTTACTCAGGACTCATC
ACCACACATCATCCACGA

>AS-P219P15-18A RGA.0.124
ATCTCTACAACAGCCACGACAAAAAGTTACATAGAGGAAGAAAAAGGAAATAAGATAGAAAACAAATCAGGCAATAGACAAAACAAATACCCA
AAGCAGTCCGAACCAAGGAAGGATCATTCAAAAATCTTTTTGTATCAAGATGGGGTGGTTTCCACCAGAGATCCTCTAGAACAACACTACAAAAGC
GCGTTCGACATCATCAAGAACAAT

>AS-P219P15-18A RGA.0.125
GAATCTTACTCCTCCCTAATAGTTCTTGTGTTGTCGACAGGGTGTCTTCTCAAAAAACTCG

>AS-P2378MCT-20E RGA.0.126
CTAAGAGATAAGAAGGAATCTTATCTCTGGGTGACATCATGACAACCTTATCTCACATCACTCCTCTTTATGATTATTTAAGGACAACTTACT
TTATGTACACCATATTGTACATGATATCAAAGACAAAATATCATGATTACTTATGGGCATTTATGATTGTTGTGGGTCCTTCTCACGGAC
TGTATTTCTGTAATGGGCTCATGTTTCATCAATGTCTAGAAAT

>AS-P2378MCT-20F RGA.0.127
GATTCTAGATGACATTGATGCACATGAGCTCATTACGACGATACCATCTGTGAAGAATGGGACCCAGCAAGAATCATGATGTCCAGAAGTCATA
ATGGTATTAATGCTTTAATAATCATGTAAATCATGGTGTGACATAAAGTAAAGTTTGTCTTTAGTAATCATAAAAAGGTGTGATGTGAGATAA
AGTTGTATGATGTAGAGATAAGATTCGCTTCTTATCTACTAGTAGTGCGCACCCAGTGTCTGTAGTTACTC

>AS-P2378MCT-20H RGA.0.129
GATTCTAGATGACATTGATCCAGTATGAAAAGCATGTGCTTAGTTCCTACAAGTACCATCACAGGATGGAGAATGTTTATAGACTACATGAGGC
TCCCACTGCTACAAGGAAGGATCATT

>AS-P2378MCT-20A RGA.0.130
TTTCCAACCTTGGTATGCCTTTGGGGGGGCTTGGGGGGGGGGGGCTTGGGTTTTAAGTTTGGGTTGG

>AS-P2414MCTT-19B RGA.0.135
GATTGAAAATGAGGCTTGATAGAAAATTGCAAAACTTTGGTTTTGGGCCAAACTTTAGCGGGCTATAACTTGGTTTTCGGAACTCCAAGTTACT
CAGGA

>AA-P2378MC-20D RGA.0.137
TTGATTCTAGAAGACTTTTTATGAGTCCCTGAGTAACTGTGACCTTGTTCAGCTGCCTGCGATCCACACAGAGCCGCATCTCCATCTTTTTTCT
TTACCAGTTACACTGGGGCACCCACGGGGAAACACTTGGTCCGATAAAGTTCTTACCCAAAAAATCCTCTAACTGAGACTTTAGGTCGGACAT
CTCCAACGGTGACATTTCTATAAGGAGCACTTGAGATTGGTCCCAGGACCAATGCGATAGCAAACTCGACCTCTCGGTTAGGTGGAAT
TCATCAATGTCTATAGAA
ATCAAAAATAATT

>AD-NBSMC-21G RGA.0.138
CAAGTACTATTATTAGCCGAGGATCGTAAACCAT

>AD-NBSMC-21G RGA.0.139
CTGCACCACACATGGAGTGGGAAGTGCCATGGGT

>AD-NBSMGAG-22C RGA.0.140
CGTGGATGATGTGGTAAGCCAGAGTGCATCAACAAGTCTAGTGTCCAGTCCCTTTTCGATGAGACTTGACCGTCTTCTACAATATGCGCTTG
ATCTCTCTGTAGACACCTCGGCTTGGCCATTGGTCTGGGGATGGTAAGTTGTTGCTATCTTGTGAAGGATGCCATGTTTCTCAGTAGGCCCTC
TTACTCAGGAC

>AD-P2378MC-20B RGA.0.229
CCTGAGTAACTAGACGAGACCGCCAACTAGCTATCGATATAGAGTAGGATTATCCAGAACAGTGACATCCATAGGAGTAAATCGAACATTAGGC
TCGAGAGGAGTAGGTTTCAAGTGTGACTATCTATAAATCCGGCTCGAGTAAGGAGATCTGAAGCATAAATAGCCTGAGAGAGATAGATGTCATCAT
CGGTGGATATGACTTCTAGAGCAAGAAAGTAGCTGAGAAAAGAGCAAGATCTTGATCTCAAAGTGTGCTGAAGCGATCCCTTGAGATCAGAGA
T

>AS-NBSMC-21B RGA.0.148
TCAATGTCTCTTCTTGCCATTGTGGCTCCTCATCTCATGATTCTTTGAT

>AS-P2378MC-20A RGA.0.155
AACAAAAGTTGATCATGGTTGTAAAAACTTTTCGACATGTAGCAAAATGTCTATGCGGGACAAGAAAAATAGGGCACGACAGAAATTTAGGCAT
CGAAATGGACCAACAATTTTGTAAAGAAATACGCGCAAGATTCGTACAACCTTTTCAAACCTTGTGATGCATTTGTGTGCTGTTGTGTGGTGT

>AS-P2378MC-20H RGA.0.156
 AAATATATTTGATTTAGATGACATTTGATCCAGAATGAAAAGCATGAGCTTATTCCTACAAGAACAATCACAGGATGGAGAATGTGTATAGACT
 ACATGAGGCTCAACACTGCTACAAGGAAGGATCATTCCCCCTGCCTCTCATTGATCAAATGCTTGAGAGGTTAGCTGGTCATGCTTTTTATTGT
 TTTCTAGATGGATATTTCTGGATATAGTTACTCAGGACTCATCAATGTCNTCTAGAATCAAAAATAATTT

>AS-P2378MC-20A RGA.0.157
 TTGGTTTTACAAAATTAATTTGGACTGGACTTACCAGCTACTTCTAAAAAAAATTTGTGGAATTC

>AS-P2378MC-20I RGA.0.159
 TCTCGAGTGTCCGTATAGGATGTCAACACTGGAAATGTCTGAGTTACTCAGGACT

>AS-P2414MC-17A RGA.0.164
 GATGCTGGACGACCTGTGGTTGCCGGAAGACACTATTGTTGTCTACTGCTGAAAGAGGGAAAAAGCTGTGGCATTGTTCTGGCCGCCAGGAGAAGT
 TATGAGGCCCTTTGAACCACCGCTGGAGCTTCTGGCTACTTCTGTGCTGTGCAAAAAAGTTGCCAGTAAGGTTTTATTGAGGTTTTTTGTTCT
 TTTTGAATTTTGGAGTTACTATGGTCACTGCCTGCTGGTTTCAGTCATCGTGATTGGAATTCATCGCCGCTGCCGCTTGAGGTGGCTGTCCGGC
 TGCCGCCGAGCCGGTTTCAAGACCGCCGCGGTGTCGGTTACGCTTTCCTCCTTCGATTCGGTAAGCGTTTTTCGATTTGAAAGCCCTTTAGTTA
 CTGTTCTGTTATCATACGCTGAGGTTTGTGGCGTTACTCAGGACTCA

>AS-P2378MCTC-11H RGA.0.162
 GGTTCGTGCTGATGCTGGACGACCTGTGGATTCATCAACATCCAAGTTTGAATGACTTTGAACAAGGTTTTCAATTCACATCTTGTGCAATATAT
 TCCAAAAATACCAGGG

>AA-NBSMCTT-24I RGA.0.169
 GATGATTGGTGGTGGTATGATGAGTCTCCTGAGTAACTTGAGGTAGGACAAGCTAAAAGGTGATGATGAGAAAAAGGAAGCGCGCTAGTAGTGGA
 GGTGGAGGAATGATACCAAGACAGTTGCTTGAATTTAGTTGCAAGTGGCACAGCCGAAGTAGACGATCAAGTATCTAATCTTCTTCCAGAAGAAA
 GAACTCGATCCAGCACACCTCAAAAACAGCAATGGCAAAATGGAGAATACTGGAGAGGATTTAGAATCATCACAAACAAGTTNGGGA
 TAATCCCAACAAGTTTCAAGACTGAACCTGCCACACATCATCCAC

>AD-NBSMCTT-24B RGA.0.170
 CAGTAGCGATAGGGGAACCTGTTTGAATCCTTAGGAGGGTTATCCGCTCTCAAGAACATCAATGGGGCCCGTGGTTTTCTGCAAGAGTGGCTTGA
 CCGTTTGCCTTGGTGGCTCTGTTGGGAATAGGCTTCTTGGGAGCTTTCCTTTTGTCTTCTTGGATAACCATCCTATGAGAGTAAAGAAGGGA
 AGTACATCAAAATCCAAGTAAGCGGAAGTGAATATGATGATGATGATGAAAAATGAAGCCATAGAATATGGGTATCGTGGTCACATGACAGCTGCA
 ACATGTGACTAAGATAACATGAAAAGAGCATGGCAAAATCATCTCCATATGGTGAAGGGCGATGATGCATGCAAGTAGGGTAGCATGAGAAGCA
 TGTA

>AD-NBSMCTT-24F RGA.0.172
 CCTATCATTTTCGCGTAGTCATCAATATCCTG

>AD-NBSMCTT-24J RGA.0.174
 AAATGTTGAGTACGATCATTGGACAATGTT

>AD-NBSMCTT-24C RGA.0.177
 GAGTAACTTCAAGCACAATTCAAAGGAGAATTACGAAACCATGCTATTTTCATGAAATAAATGTGGGAAAAGGTGAATAAATACCTCAAAAATAAG
 CACAAAATAAACCAGAAATCGGGTTTATCACTCATCAAGGTGGTTTGTGCTTGTACTCATCTTGGAACTTCCACCAATGCTTGGACTTCCAA
 TAAGTCCATCATTCAGATCAATATCTCCAAGCTTTGATGGATTTCTTCAAAACCATGGGCTCCAAATGTTGATCCTCATCTTGAATCCGG
 GACCCACACTTATTTTACACCCATCTCCAAGTTGATTATATTAGTCCATGTGGGTGGTGGAGCAAAGTGAATTCAAAGAAATGACCACAC
 ATCATCC

>AD-NBSMCTT-24G RGA.0.179
 CACAATTGCAATTTACCTCTTGGCAACAAAATCTTCTCAACCTCTCTTGAATCTTCTTTCATCATCACTCGAGATTCAATATCGGAAGTTGTGT
 GAAGTCCACATCAGCATCTCCTTCCAAATTAATAATGGAGACTCATGAAGGTGATTGAAGGAATTAACCAAGTCCAGACAAAAAATCTTGAATG
 ATGAAATCCTCTTGGCTGATTCCTACCAACTCTTCAATTTATCTCTTTAGGCACTTCATATTTGTGACTTGACATCTTCTTCTGATTTGTCAATG
 CTTCTTGCACTCCACACTTTTCACTTGGTTCC

>AD-NBSMCTT-24N RGA.0.180
 CAATTCAGACTCAAAAAGTTGCTTATGTCTCTGTCAAGTTACTCAGGACTCA

>AD-P216MCAC-E4 RGA.0.181
 TCATCCTCATCATCACCGGCTGGTAATCAACCCCATACCTCTTATTCAAGTAATCCACAGCAGCAGGTGTAGCACAAGCTTATCAGCATTCA
 AAATATCTAACAATTCAGCGTCTCGGCGACAAAACCTTACCGTCCCAATGTTCCGGCTCGAAAGCATCACGCTCCTCCGGCACTTACATC

>AD-P216MCAC-E4 RGA.0.182
 AGAAGAAGATTATGAAGATGGAGGAGTCGCAGAAGAAGGTTACTTGAATTTTGTGTTCTTGGGCAATTTTTACTTGGCAGAGTTACTAGTTACATG
 TGTACTC

>AD-P216MCAG-F2 RGA.0.183
 AGTGCAGCTTTATTGGCAGCAATTTCTTCTTCAACACCCCTACAGTATTTTGAATTTGGACCTTCAATTTCTTTCGAGATCTTGTCAATCAGAA
 GGCTTTCTATATTGTTTACACCCCTCTGTTACTCAGGA

>AD-P216MC-G4 RGA.0.184
 GGTTTGGTCTGGGAGCAGAATCCTAGTGACAACCTAGAGATGTGCACGTGCTAAAGGGACCAGAGGTGCATGATATTTATAAGGTTACTCAGGA
 TGCATCAAAATGNTGTGTCTGTATCATGTTGT

>AD-P216MCTT-C5 RGA.0.185
 ACATACATGAAGTTACTCAGG

>AS-NBSMCTT-24D RGA.0.187
GATGAAGTTTGGGTTGGAACTTTTAAACTTGTAGGGCCCTTTTGTCTGTATGTGTAGCAGGCCCTAAAATTATATTATTTTTTTATCCAAA
GTTTATGAGTA

>AS-NBSMCTT-24L RGA.0.189
AGCCACGAAGCATCCAGCGGGGAACCACCTCTGTATTGCGGTGTGATCCAATCGATATGATAGTGGCCGAGGTGCATCTAGCAAAGCGAGATGAG
A

>AS-NBSMCTT-24E RGA.0.192
TGATCTCATCCTCTTTCATGACAACCCTGTATGGTACGATAGCCAAGGGCAAACCCTTGCAACTTTTGGCAATTGATTCACCAATAGGCTCT
AAATCAGCAGGACACTCTTCTCCACCAAGCACCTCGCTGCGAGACAATTCCCAACTGTCTTCTTCATTCAAAGGGCCAAAGTTGATGGTGAGGTT
CCTTTGACATTGCATAGTTTGCACCCTATCATTTTCGCGTAGTGACCAATATCTGCTCCCATTATCTGTCTGGTAGAACATAATCCTTCAG
CTTGTTCACGCTTGAGTGTCCACACATCATCCACGA

>AS-NBSMCTT-24K RGA.0.193
GGATTGAGCAAATTTAGTTTGTGGTTCGATTTAGCTCAAGCGAATCAAGTATTTGTGTGAGTGATTTTATATCTAACAGTAAGTAAATAGCAGGA
AATGCTGACATGTGCTGAATCATATC

>AS-NBSMCTT-24M RGA.0.194
ATATTTGGGGCGTTGGGCGGTGTCTACGGGCGGATGGCACATAAGCTGGAATGTGGCTTGCCTC
>AS-NBSMCTT-24O RGA.0.195
ATGCTTGGTCCCTGTTTTTGTCCCACC

>AS-P216MCAA-D2 RGA.0.196
AGGAAGTAGAATAATAATCACAGCCAGAGATATGCACTTGATATGCATGGAGTCCATGGAAGTTGCGGAAGTGCATTCCTCTTGAAAGAAAT
GTATGTGTCTGATGATGTTGTTACTCAG

>AS-P216MCAC-E5 RGA.0.197
TCACAGTTCAAAAATATCAAGAAGACATTCCATATATGCCAAACTAACAAACAAATCACCAGGAAGAACCCTCATAAACTCATGCATTGGCAA
ACTTATTCACAGCAGCCGAATATGTGTGTTACTCAGGA

>AS-P216MCTT-C4 RGA.0.198
TCATAGTGAAGCTATTGCCCTGGTCTTGGTACTCGTGGTTTTTACTTCTCTCATTGAGGAGTTTTCAAGTTACTCA

>AD-NBSMCTC-14A RGA.0.201
CTGGGATAGAGCTGCCAAGGGCAGTCCATCACACTTCTTTACGAGATCTCTGCCAAATGATGTCAAGGGTGGATTCTCCATAGAAAAAGTTGAC
AGGGCTCATGTTTTGAAAAACAGTAACCAACAATCTTCTCAGACAATAAGCTCAGTTCATAGAGTGAACAGTTTGCACCACAGAAGCCACTC
TTCTAATTCT

>AD-NBSMCTT-24F RGA.0.203
CCTATCATTTCGCGTAGTCATCAAGATCCTGCTCCCATTATCTGTCTGCTGGTGCAGATAGTCCAAGTTGTTCCACGCCTTATTGTCCCACAT

>AD-NBSMCTT-24J RGA.0.204
AATGAGCCATGTGAATGAATGAGCCAGTGATAGGCTACGATAACCTAGATTTCGCGCATGATGGAGAGAGGGTGAAGGGTTTTACGCATAGTTG
AGGAGAGGTAAAACGTTGACGACGGTGATTGGACAATGTTTACAGCCACACATCATCCACGA

>AD-NBSMCTT-24L RGA.0.205
CAGTGTCTATCATTTTTCCAAACATCATC

>AD-NBSMCTT-24N RGA.0.206
GGCCATCTTGTGCGGTTGGGGTTGCGACCCCGATGTTGCTCATGTCAATCACATTGGCTGGAACCACACATCATCCACGA

>AD-NBSMCTT-24F RGA.0.207
AGTTGCTTGGGTTTTGTCATCGGTGATCGTAGCCATAGCAGGGGTTGTCATTAAGAGGAAGAGATCTAAAACAGAGTGT

>AD-NBSMCTT-24L RGA.0.208
ATCCAACAGTAAATAAATAGCAACTAATGTAGTAGCGGAGAAGGATGA

>AD-NBSMCTT-24O RGA.0.210
CCCCGTTAATATCCCACATCATC

>AD-NBSMGCT-25A RGA.0.213
ATGCCTCGTAGTAAGGAGTTGTTGC

>AD-NBSMGGC-8D RGA.0.14
TGAGTCTCTGAGTAAGGCATTTCTTTTTCCTCTTTTTTATTACGCAAAGGCTTCCGACAACCTTTGGATGCAAGGGGATAAGCCACACATCTTCAG
ACGATTTCTTGGCAATCTCCACCAGCTTTACCGGTTGCTCCTTGTGGATCAAATGCATAGTGTGTGAACAATGACCAATTGTCATCCTTAG
GTAATGACTGTAAGTGATGAGATTGAGTAGGGCATGTTTCCATGGCAGATGCAACAATCTGATCGCGTGTGTTGTGACCAAAATAAGAAAACACCTTT
GGATTCACTACCTCTAGAATTCAGCACAGATTTCAAGTCTCCCATTTGTCAATTGTCCCACACATCAT

>AD-P212MCT-9C RGA.0.216
TATTCCTGATCAATACGGAGACTGTGATTCTCTTGAATATGTGAGAAACAGTCTCCGGTCAAGTACCACCGAAGTTCTGGAGC
CTCCCGAAGCTTCAGCTTCTCAAATGGCTAACAAAGGTTTGAAGGTTCTGTCTTCTTCTGTATCCAGCGCCAGGGGAATCACTGTTCTTC
TTTTATCCGGTAGCAGCTTCTCA

>AD-P212MCT-9A RGA.0.217
 TGAGTCTCGAGTAACCTCATTATATGATCTGCGTCAGCTGTAACAAGAACATGCTTATCTCTTGTGGTCACTGTGATTTCTGCTACTTGCCCCAA
 ACCATGTATGTCCACATAAAATCTCCACTTGGTCTGAATCATTACATCATCAAGAACAAGAAGCAGAGTCTCNGTGACAACAAGAGATAAG
 CATGTTCTTGTACAGCTGACGCAGATCATATAATGAAGTTACTCAGGACTCA

>AD-P212MCT-9B RGA.0.218
 GATTATGTGAAGGACATACATGGTTTGGGGCAAGTAGCAGAATCATAGTGACAACAAGAGATAAGCATGTTCTTGTACAGCTGACGCAGATCA
 TAT

>AD-P216MCAA-26B RGA.0.220
 TCTTCTTGTCTTGTGATGATGAAGTGAGGTGAAGTGTGTTAGAGAATTTAGGCGAGAATCAAGATTGGTTTGGACCAGGAAGTAGAATAATAATC
 ACAGCCAGAGATGTGCCTTGTAGAAATACATGGAGTCCATGGAAGTTGCGGAAGTGTCAATCCCTCTTGAAGAAATGTATGATGTTGTTAC
 TCAGGACTCA

>AS-P216MCTC-27A RGA.0.272
 TTCTTGTCTTGTGATGATGAAGTACTTCTGATCAAATCATAACTTAGTTGGAGGACATAAATGGTTGGGTCTGGTAGTAGAATCATTGTTAC
 AACTAGGGATAAACATGCAATTTGTAGAAGACGGATGACATATACAAAGTAAAGGCATTGGAATCTAGTGATGCCAACAGCTTCTAAATTTGC
 ATGCCTTTGATGGAAGTGTGACATTTGCCTTGAATTGCGCTGGCATAATCTTGTGTCGAAAATGGTTGAATATGCTAAAGGCATCCCATAGG
 ATTTGAAAGTTTTGGGTTCTTTCTTTATGGAAAATGTTTTGAAGATTGGGGATGTCAATTACGGAAGCTTCAGAAAATGCCCTTCAAGAGATT
 CATAAATGTGTTGAGTTACTCAGGACTC

>AD-P216MCTC-27C RGA.0.224
 CTTCTTGTCTTGTGATGATGAAGTTGAGAAGCGCAGAAAATCTGAGTTTTGCAATGAAATGTTGAGAATGATGAATAAGACATGATTGAGGATG
 ATGAAAAGTTGATGATTGACTTGAATGAATTTGAGATTGAATGAATTTGAAATGAGATACCTAGACAGTAGCAAAGATTGGTTTGGTCCACTTGTCCGGG
 TCAGTGATTGAGATACCTGGGTAGTCGCAAGGATTGTGGTTTGGTCCCGCTTGTCCAGGTCAGTGATTTATGAAGTCTGGGTAGTAGCAGTAGTA
 GTAGATTATTTCACTTACTCCAGGTTGAGCTTGTAAACACCCGCTGGGTAGTAGCAGTAGTAGTTACTCTACTTGTCTGAGTTACTCAG
 GACTCATC

>AD-P2378MCTA-23A RGA.0.231
 CCTGAGTAAC TAGACGAGTCCGCCAAGTAGCTATCGATAAAGAGTAGGATTATCCAGAACAGTGCCATCCATAGGAGTAAATCGAACATTAGGC
 TCGAGAGGAGTAGGTTTCACTGTGACTATCTATAATTTCCGGCTCGAGTAAGGAGATCTGAAGCATATTTAGCCTGAGAGAGATAGATGTCATCAT
 CGGTGATATGACTTCTAGACCAAGAAAGTAGCTGAGAAAACCAAGATCTTTCATCTCAAAGTATGCTGAAGGGATCCCTTGAGATCAGAGAT
 ACCATCAATGTCTCTAG
 AATCAAAAATAAT

>AD-P2378MCTC-11G RGA.0.29
 AATTATTTTGTATTCTAGATGACATTTGACAACATCCAAGCCAAAAAGAAGAGCTCGAGCTACTCCCAGAGGTCGAGAACAAGCGCATATAAGA
 CAAGCAGCACTAAGCAAAGAAATGGCTATAGAGGTATAACCATGAAGTCAT
 >AD-P2378MCTC-11G RGA.0.234

GGTATAACCAAAAAGTCAATTCGAATAAGTTTTGCTACATACCATTATTTTTGAAAATCAGATCGGACTGACCGGTTCAACTGAGTTACTCA

>AD-P2391MCTA-13E RGA.0.235
 GCCACATTACGGCAGCCTGTATGGA

>AD-P2391MCTA-13B RGA.0.236
 ATGGTGTGGACGACGTGGACCCTATAGAGTGGAGTATGCTTCCGGGATACTATTCACTGAAATGAGTAAAGAGAGAGATGAAGCAATTAGAG
 CTACTAATGCAATAAGACTGTCAAAGCCATCCTCCGTATTTAGTCCGTTTTGTGAGTAAATTTTACTGATATAGAAAATGAGTAATCCAA
 CTACTAGCTAGTTTGTAAATGAACAAATGCTGTAAAAAATGCCATTTATAAACAACCTTATTTCAATGAAATTTTTTCCATAGTTACTCAG
 GA

>AD-P2391MGAG-16B RGA.0.238
 TATCGTTGTCTGGATCTCCAAGGTCGTTAGACATCGTAATCGTATATGCGGTGCAAGGCTTGTGCTTCCGCTTCCCTCTGAGAAG
 CTGTAGCTCTCTCCCTTAGCCTCTCAATCCTTCTGGTGTCTGATGGCAAATGACTGTGTAGTGGTGTATGAGTAAGAAAACAAGTATC
 AGAAAAATAACAACCTCTAAAAGTAAGGTAATAATAAGAAGGTACCTTGTGTTGAAAAGATCCTCTTACTC

>AD-P2414MCTC-10C RGA.0.239
 TTCGTGCTGATGCTGGACGACCTGTGGATTTCATCAACATCCAAGTTTGAATGACTTTGAACAAGGGTTTCATTCACATCTTGTGCAATATATTC
 CCAAATACCAGGGTTGACTCGAGCTGATTCATTCGTTCTATGCATGATGTGACGAAGTCACTACTCTCTGCATTTTCACTAGGTAACACTTGG
 GATATTTCCATAGGATCTTCAATGCTAGCTTCATCAGTAAAGTACTGCTGCCTTTTGGCAACCTTCGTTTTCTTTCCATGTTTTCTGTGCTAGTG
 TCAACAGAGTTACTCAGGACTCATCAACAGTNTAAGTTGGGAATGGCGGTGTGTGACATCGGGTTGTGCACGTGTACCCTT

>AD-P2414MCTT-19C RGA.0.242
 TTCGTGCTGATGCTGGACGACCTGTGGGTGCTAGGTAGGATATAGTGGGATAGGATATGGGCCCTGATGACAAGTCACTCATATACCCATTTTTC
 AAGCTAATTTCACTTGTGTTTGTAGCATTATGCACCTCTCTGCTCCTAAGTAAGTATTTGGAGTGAAAATGCGTACTTCTCTAAATCAAGC
 AACCACCATGAAGTCACTCAGGAC

>AD-P2414MCTT-19B RGA.0.243
 GTAGGTTTGAAGTGAATAAAATGAGACTTTTGGATTGAAAATGAGGCTTGATAGAAAATGCAAACTTTGGTTTTGGCCAAACTTTAGCGG
 GCTATAACTTGGTTTTCCGAACTCCAAGTTACTCAGGACTC

>AS-NBSMC-21A RGA.0.245
 TCGTGGATGATGTGTGGTATAATCAGGGACCAAGCATATCTCTTTCATGAGCAACTAGACCAAGAAAATGGCTATTGATCAAGATCTGAGAGA
 TTGAGTTACCAAGGATTTGGGGCTCAATCAATCATGATTGCCAAGAGGCTATGAGTTGCATGATTGAAGAGGATATGAGCTGGATTCAATCCA
 AAGAGACAATATCTCTGATCTCAATGAATTTCTCCATTTCTTTATAGCTTTTATATACATTTCTGCAAAATTTCTATTCCCTTTTACATTTCTGTC
 TATTTCCATTTTCACTTTTATGCTTTCTTTTATCTTTTGCATTTCTGTTTCTGATTTATAATTTCTGCATTTATAACTTACTCTGTTTAC
 TCAGGACTCATATCCCA

>AS-NBSMC-21B RGA.0.246
AATGTGGTTGGTGATCCACCCTTA

>AS-NBSMC-21C RGA.0.247
TGGATGATGTGTGGTATAATCAGGGTCCAAGCATATCTCTTTCATGAACAATTAGACCAAGGAATTGGCTAATGATCAAGATCTGAGAGATTG
AGTACCAAGGGATTGGGGCTCAATCGATCATGATTGCCAAGAGGTCATGATTTTCATGATGAAGAGGATATAAGCTGGAATTGATCCAAAG
AGACAACATCTCCTGATCTCAATGAATCTCCCCATCTTATCTACCACCTTCTTTATAGCTTGCTTCAACATCTGCAATACCCCATCTCATA
TACAGTTACTCAGGACT

>AS-NBSMC-21B RGA.0.249
GAGATCAGTAACAGAAGGGAGGATAGAT

>AS-NBSMCTC-14D RGA.0.253
CTGAGTAAATCCCTCACCTTTGTATTCAACATCTTATTCGTCTGGTCGATATCATCATCACACCAATGAGGAATAAATCTCAATACTCT
TCCATTTCTGTTCTGATCTCTTCCCTTTGCAAACAATAGCCAATGGCAAATCCTTGCAACTTTCGGCAATTGATTTACCAATAGGCTCTAAT
CAACAGGACACAACCTTTTTCTCCATTCAAACGGCTAGGTTTCATGGTGAGGTTCTTTGACCCTGCATAACATTTCTGTAGTCACCAATATCC
TGCTCCCGTTATTGTGTTTTGGGAGAAGATAAATCTCAGCTTGTTCACACCTGAGTGTTCACACATCATCC

>AS-NBSMCTT-24D RGA.0.258
CTAATTTTTTACCATTTTTTCTCATGTCTAACTTACCCAATAGAAAATGGAATTAGTGTCTTCTAAGCACGTGATTTTATGATTTTATCCTTT
TTACCACCTTTCATCGAAACAGAGATACTATATAATATGTCATTATTTATAAATAGACTCTCTTATATACTCTTTAATATAATAGTGGAGTAG
TATATTTGATTTTCTCATATTTATGCACATGATTTTTTGTGGTTTTTCTTCTTCTTACAAAAAATATTCATAAACTTTGGATAAAAAAAT
TATTTAATTTTAGGGCTGCTACACTTACAAGCAAAGGGCCCTACAAGTTTTAAAAGTTTCCAACCAAACTT

>AS-NBSMCTT-24M RGA.0.259
CTCAGGACTCATCATACATTATTCATG

>AS-NBSMCTT-24A RGA.0.260
TTGTCCATTGGCCATCCAAAGATGAATTAGCTGTTCTTTGAGATTCGAAAACTTTCAGGATAGAGAGCACAAAAACAAAAGCATTGCCGTGAT
GATGGCTTCAAATGAAAATAACTTAGTCAAAAAGCACCCATGACAGCATCAATGTCCAAAACTTACTTTTCCAACACATTGAGCCATTTCTTTT
CCTCTTTTTTATTACGCAAAAAGGCTTCCAACAACCTTGGATGCAAGGGGATTACCCACACATCTTCTGACGATTTCTTCCCAATCTCCACAG
CTTTGACAGGCTGCTCCTTGTGTTTTGGCCAAATGCATAGTGTGTAACAATGACCAATCTCATCTTTAGGTAATGGCTGTAAGTGATGAGATTGA
GTAGGAATGTTCCATGGCAGATGCAACAATCTGATCGCGTGTGTGACCAAAATGAAACACCTTTGGTTTCACTACCTCTTGAATTCAGCA
CAGACATCAAGTCTCCCATTTGTCAATTTCCACACATCAT

>AS-NBSMCTT-24H RGA.0.262
CGTGGATGATGTGTGGCTCGTGATGATGGTGCCGATGGCACCACGGTGCCAGAACAGTGATGAGTCTGAGTAACTTGAATGAAAGTGTGGCA
GATAGCGGTGAAGAACTGAAGGCAATGAACTCGAAGGCAGAGATGGAATTTGCGGTGAGAAGTAGAAGTACTGGGAAGGGTTAGGCCACAAGAA
TTGTGGGACTGAGAGGCTACTGTGCAGGGAATGAAGAAAGGCTTATAGTGTACGATTACATGGCGAATCTAAGCCTCTTTCTCATCTGCATG
GGCAGTTTGCAAAATGAAGTGAACCTGGATTGGAAAAGGAGAATGAAGATTGCAATTGGATCTGCAGAAGGAGTGTGTATTTGCACCATGACGT
CACACCACACATCATCCACGA

>AS-NBSMGCT-25B RGA.0.263
GGATCCATTCCATACTAAAAATACTTTTGGAAACTCAATGTAGCTACTGTTGATACTGAGGATATAAGACGTTATGCGCCAGCAGTCACGATG
CTTACTCAGGACTCATCATCAACATCATCCATG

>AS-NBSMGGC-8C RGA.0.267
TCCTGAGTAAGGCAGGAACAACCTTTATCGTATCATCGGAGAGTTCACAGATCTCACTCTTCAATAAATGATTCCAATACTTGATATCAGAATT
TCCACGCAATAAGCCTCCAAGGGATTGAGCTGCCAAGGGCAATCCATCACACTTCTTACGAGATCTCTGCCAACTTTTTTCAAGGTTGGATT
TCCATAGAAACAGTTGAGAGGCGCGCATGTTTTGAAAACAGTAACCAACAATCTTCTCAGACAATAAGCTCAGTTTATAAGGTGAAACAGTTT
GCACCACAGAAGCCACTCTTTACTTCTAGTTGTTATGAGAATTTTACTTCCCTTGACCCCTTTTTCGAAAAGGTTTTAGAAAGTTTATTTCCAATC
TTCATAACTTTCACTCCACACATCATCCACGAC

>AS-NBSMGGC-8J RGA.0.266
TCGTGGATGATGTGTGGACAATGTCAAATGGGAGATCTGAAGTCTGTGTGAATTGTAGAGGTAGTGAATCAAAGGAGTTTCAGTTTTGGT
CACCACAGAGATCCGAGTGTGGCATCTGTCTATGACACCATGCAATAGTCTACACTTGCAGCCCTTACTCAGGACTCATC

>AS-P216MCAA-26A RGA.0.269
TCTTGTCTTGATGATGTAAGTGAAGTGTGTTAGAGAATTTAGCGGAGAATCAAGATTGGTTTGGACCAGGAAGTAGAATAAATATCACA
GCCAGAGATGCACTTGATATGCATGGAGTCCATGGAACCTGCGGAAGTGCATTTCCCTCTTGAAGAAATGTATGTCTGTATGATGTTGT
TACTCAGGACTCATCATACAGATACATACATTTATTTCAAGAGGGAATGACACTTCCGCAA

>AS-P216MC-28G RGA.0.271
GGTTTGGTCTCGGAGCAGAATCATAGTGACAACCTAGAG

>AS-P216MCTC-27D RGA.0.277
GTTCTTGATGATATAAGTCGTAACCAGTAACATCCAATATTTTTATTTTTCAGAAGAATAATTTTTCTAATCGGCCCTCTTCTGTAACATAATGTGCAG
TTCATATGGACGCCGTATATGGGAGTGGGGTTCTTGATGTCTCGCCACCCATTTGGTTATCTGCTCCACTTAGTCGCCGCTAGTGTATTGCG
AATGCATAGAATGGCAGCAGCAGACCGAGTTAGACGACAGTTCGGGATCTAACAGCTTCCACCAAGCCCGGTTCCGACCTTGGTCATGATCA
TTGCAAGCGGTTGACAGGACACAGAACCATGACTGGGGACAGATTTACAGACAATGAGTAACTCAGGACTCAACAAGAAGAGATGANTTACC
ATCATAAAGAACAGGAGGA

>AS-P216MCTC-27E RGA.0.274
ACTTACATCATCTAGAACAAGA


```

>AS-P216MCTC-27G RGA.0.276
AATTCTCAATTCAGCCAGCTAGATAGATACAGCGTGCAAAACGAAATTCATGCTCTAAATCAAGGGCATGCCAATAATACGAGGTTAGGACAG
TTACTTTCATCATCAAGAACAAGA

>AS-P216MCTC-27F RGA.0.280
AGTCCTGAGTAACTCAAGTTATCAACTGATAAGGCACCCCTGACGAGCTTATTAATCCAAGCATATCCCTTGACAGGAGCACCACAGACACAT
GCCTCAAGATTCAAACCCCTGGTGCTAGCCTTATTTCTTATAAAATTTCTTTCTCCTTCAAACCTCATTGTTCTTTTTTTCTTCTTATTAG
GATCTTGTGTTTGTCTAGTCTCAATAGGATGTGCTTCAAGCATGTGACTTAGAGTATATAGTTGCCTATATACCTTGTGGTGAACCAACTTA
GCTAATCAATTACCATACCATATACATTGAGAACTTACATCATCAAGAACA

>AS-P216MCTC-27D RGA.0.281
GTAACCTCAGGGCTCAACACTTACATCATCAAGAAC

>AS-P2378MC-20G RGA.0.286
TCAGATGCTGATAAATATGCAAAGGTTTCATAATATG

>AS-P2378MC-20I RGA.0.289
ACTCGAGATTGTCCCAACCCAGGCACCAACTCGATAGCAAACCTCAATCTCT

>AS-P2378MCTA-23B RGA.0.290
TTTTGATTCAGATGACACTGATGAACATGAGCCCATACGAAATACAGTCCGTGAAGAATGGGACCCAGCAACAATTATAATGTCCATAAGTA
ATTATGATATTTTGTCTTTCATAATCATGTGAAATATGGTGTGACATAAAGTAAGTTTGTCTTTTGTAAATCATAAAAAGGAGTGATGTGAGAT
AAAGTTGTCAATGATGTATCCAGAGGTAAGATTCTTCTTATCTCTTAGTAGAGCGGTCCAGGTGTCTGTTGTTACTCAAGAATCACACAAGTGC
TCTAGAATCAAAATAATTT

>AS-P2378MCTC-11B RGA.0.295
TGACCATAACGAGTCTGAAAGCGGTCTTAGGGATATCTTACCCCTCACCTTATCTGGTGATAATCGGATCGCAAATCGATCTTGAGAAAA
TCTCAGCTCCTTGTAACATGATCCATGAGATCATCAATCCTCGGCAACGGGTACTTATCTTTATCGTGACCTTGTTCAGCTGCCTGTAATCTAC
GCAGAGCCGCATACCTCCATCTTTCTTCTTACCAGTAACATTGGAGCACCCACGGGAAATACCTGGTGGATAAAGTTCTTACCAACAAA
TCCTCTAACTGAGACTTTAGCTCGACCATCTTAACGGTACATTTCTATAAGGAGCACTTGAGATTGGTCCCGCCAGGCACCAATTCATAG
CAAAATCAACCTCTCGATTAGGTGGAATTCATCAATGTCACT
AGAATCAAAAT

>AS-P2378MCTC-11H RGA.0.296
TTTGGAGAGTTACTCAGGACGCATCAATGTCTATAGAAATCAAAATAATTTT

>AS-P2414MCTT-19A RGA.0.302
CTGGACGACCTGTGGATGTGGACGGGTCAAGGGAGAGGTGCAAGTGGCAGTGGAGGACTCGAAACGACTGAGAATAATCCTATAAACTTCATG
ACTACCCTGAAAAACATGGCTGCTGCTATGCAGGCCACGACTGTGGTGTAGGGAAATCAAGTAGGTAATGGGAACGGCAGTAATGGAAGAAACA
GGCCAAATGACCTTAGCGACCTTCTCTGAAAGTAAACCCACCAAACTTTAGAGGGGCGACGAAACCCATGAGGAGCAGATAAATGGTTTTCAGGCTAT
GGAAAGAGCCTTGAAGCGCAATAGGTGCCTAAGGAGAGATAAGCCTCAAAGTAGTCCCTGAACTTGCCTCGAGTCTCAAAGTAATCCCTGAA
GTTACTCAAGA

>AD-P212MCT-9C RGA.0.219
ATATGATCTGCGTCCGCCGTAACAAGAATGCTTATCTCTTGTGTGCTATGATCTGCTACTTGCCCCAACCATGTATGTCTCCACATA
AATCTTCC

>AS-P2378MCTA-23E RGA.0.292
ATTTTGATTCAGATGACATTGATCCAGAATGAAAAGAATGAGCTTATTCCTACAAGAACAATCACAGGATGGAGAATGTGTATAGACTACATG
AGGCTCAACACTGCTACAAGGAAGGATCATTTCCCTGCTCTCATTTGATCAAAATGCTTGAGAGGTTAGCTGGTCAATGCTTTTATTGTTTTCT
AGATGGATATTTGGATATAGTTACTCAGGACAATCCAATGTCTATAGAAATCAAAAT

>AS-NBSMGGC-8F RGA.0.265
GTGACAGTTATGAAGATTGAATAAAC

>AD-P216MCTC-27B RGA.0.222
GTTCTTGATGATGTAAGTACTTCTGATCAAAATCACAACTTAGTTGGAGGACATAATGGTTGGTCCCTGGTAGTAGAATCATTGTTACAAC
GGGATAAACATGCAATTTGTAAAGCGGATGACATATACAAAGTAAAGGCATTGGAATCTAGTGATGCCCCACAGCTGCTAAGTTTGCATGCCTT
TGATGGAACATGATGGCATTGCCTTGAATTGCACTGGCATAATCTTGTGTCGAAAATGGTTGAATATGCTAAAGGCATCCCATTAGGATTGAAA
GTTCTGGGTTCTTTTCTTTATGGCAAATGCTTTGAAAGTGGAAATGTCAATCCAGAAGCTTCAGAAAATGCCCTTCCAGAGATTACATAATG
TGATGAGTTACTCAGGACTCA
    
```

Seqüências FASTA traduzidas para aminoácidos de amostras RGA-Display com homologia para RGA usadas para Alinhamentos e Árvore Filogenética

```

> AD-NBSMGGC-8B
IVDDVWIDNSRNWESLHKPFQYQKKGSKVIVTTRNDNVADVVTISAYKLGLLSEKDCWSLFLVVKRVFLSTASKEYSTLSVGRELVKKCKGLPLAV
EALGSLRLTINYDERDWDVSLKSELWEVFEDQNDIIPALRTH

> AS-P2414MCTC-10A
MLDDLWEANVWDKLPDNNNGSMILVTTTRNDQVANYARSKEPHQLNLLDDEQSWEIFRNKVFGRECPPLLELIGRSIAFKSCKGLPLAI
KTTAGIVAKKERSDAWEEIKNLLPYWSTFEDKEGKMMELKLSYDDLSEKMKPCFLYLGVFPEDEEILVRELL
    
```

```

> AD-P212MCT-9C
EDLCGGHTWFGASSRIIVTTRDKHVLVTADADH

> AS-NBSMCTT-24A
DVWENDKWEDLMSVLNSRGSETKGVSILVTRDQIVASAMETFPQSHHLQPLPKDENWSLFTHYAFGQNKQPAKLV EIGKEIVRRCVGNPLA
SKVVGSLLRNKKEEKEWLNVLSEKFWIDAVMGAL

> AS-NBSMGGC-8C
VVDDVWSESYEDWNKLLKPFKRGVKGSKILITTRSKRVASVVQTVSPYELSLSEEDCWLLFSKHARLSTVSMENPTLKKVGRDLVKKCDGLPL
AAQSLGGLLRGNSDIKYWNHLLKSEIWELSDDTIKVVPALLR

> AD-NBSMCTT-24F
MWDNKAWNKLDYLSPGQNGSRILMTRNDR

> AS-P219MCTA-12F
LIVLDDVDDSDPMEDLCGGHTWFGASSRIIVTTRDKADNLYTCWHCFEYVGMILVF

> AD-P216MC-G4
FGPGSRILVTRDRVHVLKGPVHDIYKVTQDASXCVCI ML

> AD-P212MCT-9A
LHYMICVSCNKNMLISCGHCD SATCPKPCMSPHKSSWSESFTSSRTRSRVSVTTRDKHVLVTADADHIMKLLRT

> AD-P212MCT-9B
EDLCGGHTWFGASSRIIVTTRDKHVLVTADADH

> AD-P216CAA-26B
LLVLDDVSEVKLENLGENQDWFPGSRIITARDVHLL EIHGVHGTGCGSVIPS

> AS-P216MCTC-27A
LILVLDDVSTSDQIINLVGGHNWLGPGSRIIVTTRDKHAICRRRMTYTKRHWNLVMPNSFCMPLMELMDIALNCAGIILCRKWL NMLKASHDK
FWLFFMENVLKIGDVNYGSFRK

> AS-P216CAA-26A
LVLDDVSEVKLENLGENQDWFPGSRIITARDMHLICMESMELAEVSPFLERNVCVMMLLLRTHHTDTYIYFKRE

> AD-P216MCTC-27B
VLDDVSTSDQITNLVGGHNWLGPGSRIIVTTRDKHAIC KADDIYKVALESSDARQLLSLHAFDGTGDGHCLELHWHNLVSKMVEYAKGIPLGLK
VLGSFLYGKCFEEWECQFQKLQKMPFPEIHNVMYSGL

> AS-NBSMCTT-24E
VDDVWDTQAWNKLKDYVLP EQNNGSRILVTRNDRVANYAMSKEPHHQLGPLNEEDSWELSRSEVLGGEPCADLEPIGESIAK SCKGLPLAIV
TIAGVVMKRMRS

> AD-NBSMCTC-14A
RIRRVASVQTVSLEYLSLSEEDCWLLFSKHERLSTFSMENPPLTSFGRDLVKKCDGLPLAALSQ

> AD-NBSMGGC-8D
DVWENDKWEDLMSVLNSRGSESKGVSILVTRDQIVASAMETCPTQSHHLQSLPKDDNWSLFTHYAFDPNKQPVKLV EIGKEIV

> AS-NBSMCTC-14D
DDVWNTQVWNKLYLLPKHNNGSRILVTRNVMQQRNLTMNLAVMEKKVSVRAYWINCRKLQGF AIGYCLQRGRDQKQNGRVLRI FLIGV
MMISTRRIRCIQRGIYSVX

> AS-NBSMGGC-8J
VDDVWDNVWEDLKSVLNCRGSGIKGVSVLVTTRDPSVASVMTPCNSHHLQPYSGLI

```

Seqüências FASTA traduzidas para aminoácidos de amostras RGA-Display com homologia para Proteína Kinase-2 usadas para Alinhamentos e Árvore Filogenética

```

> AD-P2391MCTA-13D
LMVLDDVENLPYQIFPVVFNWSVGNEICKESLSRGTNACKGNSICVDAETGYAYQCQCQCFEGYPYLHA

> AS-P2391MCTA-13C
MVLDDVENLPYQIFPVVFNWSVGNETCKKILRRATNACKGDSTCLDAESGYAYRCQYNQD

> AD-P2378MC-20B
SLISRDRFSTLLRCKILLFLSYFLALEVISTDDDIYLSQANYASDLLTRAGIIDSHTEPTPLEPNVRFTPM DVTVLDNPTLYR

> AD-P2378MCTA-23A
IILILDDIDIGISDLKGLQHTFEMKDLGFLSYFLGLEVISTDDDIYLSQAKYASDLLTRAGIIDSHTEPTPLEPNVRFTPM DGTVLDNPTLYR

> AS-NBSMCTT-24H
KCWQI AVKKLKAMNSKAEMEF AVEVEVLGRVVRHKNLLGLRGYAGNEERLIVYDYMANLSLLSHLHGQFANEVQLDWKRRMKIAIGSAEGLYL
HHDVTPHIIH

```

Anexo 3

BLASTX Seqüências de bandas genotipadas por RGA-display

BLASTX 2.2.13 [Nov-27-2005]

Reference:

Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402.
 RID: 1137179812-6478-173421752491.BLASTQ4

Database: All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples
 3,210,408 sequences; 1,102,455,957 total letters

AD-P212MCT-9B (102 letters)

Sequences producing significant alignments:	Score (Bits)	E-Value
gi 37221969 gb AAN85376.1 resistance protein [Arachis carden...	71.6	8e-12
gi 37221989 gb AAN85386.1 resistance protein [Arachis cardenasi	70.1	2e-11

Alignments

>gi|37221969|gb|AAN85376.1| resistance protein [Arachis cardenasi] Length=159

Score = 71.6 bits (174), Expect = 8e-12
 Identities = 32/33 (96%), Positives = 32/33 (96%), Gaps = 0/33 (0%)
 Frame = +2
 Query 2 EDLCGGHTWFGASSRIIVTTRDKHVLVTADADH 100
 EDLCGGHTWFGASSRIIVTTRDKHVLV ADADH
 Sbjct 84 EDLCGGHTWFGASSRIIVTTRDKHVLVKADADH 116

AD-NBSMCTC-14C (385 letters)

Sequences producing significant alignments:	Score (Bits)	E Value
gi 2559006 gb AAD04188.1 beta-amylase [Medicago sativa] >gi ...	179	2e-44
gi 2935474 gb AAD04259.1 beta-amylase [Trifolium repens] >gi ...	175	4e-43

Alignments

>gi|2559006|gb|AAD04188.1| beta-amylase [Medicago sativa]

gi|3913031|sp|O22585|AMYB_MEDSA Beta-amylase (1,4-alpha-D-glucan maltohydrolase) Length=496
 Score = 179 bits (455), Expect = 2e-44
 Identities = 80/101 (79%), Positives = 91/101 (90%), Gaps = 0/101 (0%)
 Frame = +3
 Query 6 DVVWVGIVESKGTQEYDWSAYRDLFQLVQHCLMLKLQAIMSFHQCGGNVGDVSVIPIKWVL 185
 DVVWGI+E KG +EYDWSAY+ LFQLVQ C LKLQAIMSFHQCGGNVGD V+IP+PKWVL
 Sbjct 54 DVVWGIIEQKGPKEYDWSAYKSLFQLVQKCGLKLQAIMSFHQCGGNVGDVVIPLPKWVL 113
 Query 186 EIGVSDPDIFYTNPNGIRNKECLSIGVDNRPLFHGRAAIEV 308
 +IG SDPDIFYTN +GIRN+E LSIGVDN+P+FHGR AIE+
 Sbjct 114 DIGESDPDIFYTNRSGIRNQEYLSIGVDNKPIFHGRTAIEI 154

AD-NBSMGGC-8B (420 letters)

Sequences producing significant alignments:	Score (Bits)	E Value
gi 62632823 gb AAX89382.1 NBS-LRR type disease resistance prote	123	2e-27
gi 33341543 gb AAQ15192.1 resistance protein [Vitis vinifera]	122	4e-27

Alignments

>gi|62632823|gb|AAX89382.1| NBS-LRR type disease resistance protein Rps1-k-1 [Glycine max] Length=1229

Score = 123 bits (309), Expect = 2e-27
 Identities = 62/136 (45%), Positives = 91/136 (66%), Gaps = 3/136 (2%)
 Frame = +3
 Query 3 VDDVWIDNSRNWESLHKPFQYQK-KGSKVIVTTRNDNVADV*⁺TISAYKLGLLSEKDCWS 179
 +DDVW ++ +W L KPFQ G + SK+++TTR++ A VV T+ Y L LS +DCWS
 Sbjct 268 LDDVWTEYVDWSSLKPFQCGIIRRSKILLTTRSEKTASVVQTVQTYHLNQLSNEDCWS 327

Query 180 LFKVRFVFLSTASKEYSTL*SVGRELKVKCKGLPLAVEALGSLRTINYDERDWDSVLKSE 359
 +F LS SE +TL +G+E+ VKKC GLPLA ++LG +LR +D DW ++L S+
 Sbjct 328 VFANHACL SLESNENTTLEKIGKEIVKCC DGLPLAAQSLGGMLRR-KHDIGDWYINLNSD 386

Query 360 LWEVFEDQNDEIIPAL 407
 +WE+ E + ++IPAL
 Sbjct 387 IWELSESE-CKVIPAL 401

AD-P219MCTA-12B (466 letters)

Sequences producing significant alignments:	Score (Bits)	E Value
gi 77555600 gb ABA98396.1 Transposable element protein, puta...	68.2	9e-11
gi 12320952 gb AAG50597.1 hypothetical protein [Arabidopsis tha	67.8	1e-10

Alignments

>gi|77555600|gb|ABA98396.1| Transposable element protein, putative [Oryza sativa (japonica cultivar-group)]

Length=1219

Score = 68.2 bits (165), Expect = 9e-11
 Identities = 37/97 (38%), Positives = 56/97 (57%), Gaps = 11/97 (11%)
 Frame = +1

Query 142 ILPIAFALVEGENTESWSYFLTNLRRHV-TPQMGLVISDRHNGIKTGLETDPDSGWQP-- 312
 +LP+AFAVE ENTESW +FL +RR V + ++I DRH G+ ++ +GW+
 Sbjct 295 VLPMAFAFVESENTESWYWF LDRVRRKVMCMRPNVCLIHDRHAGMLRSIDYLQNGWEEKG 354

Query 313 -----PHTYQAF CIRHVAANFALS FKEQ---DVRRRRL 399
 P + +C+ RH+ ANF FK + D+ +RL
 Sbjct 355 IPAKWPDVWSRWCMRHMGANFYKQFKNKHLMDL FKRL 391

AD-P2378NCTC-11D (320 letters)

Sequences producing significant alignments:	Score (Bits)	E Value
gi 2995405 emb CAA73042.1 polyprotein [Ananas comosus]	135	6e-31
gi 47900698 gb AAT39297.1 putative gag-pol protein [Solanum dem	132	5e-30

Alignments

>gi|2995405|emb|CAA73042.1| polyprotein [Ananas comosus] Length=871

Score = 135 bits (339), Expect = 6e-31
 Identities = 63/88 (71%), Positives = 74/88 (84%), Gaps = 0/88 (0%)
 Frame = -3

Query 273 EVEFAIELVPGAGPISSAPYRMSPLEMAELKSHLEDLLGKNFIRPSVFPWGAPVLLVKKK 94
 E+EF I+LVPG PIS APYRM+P E+ EL++ L+ DLL K FIRPSV PWGAPVL VKKK
 Sbjct 4 EIEFVIDLVP GTTPI SKAPYRMAPAELRELRAQLQDLLDKGFIRPSVSPWGAPVLFVKKK 63

Query 93 DRSMRLCVHYRQLNKVTIKNKYLLPRVN 10
 D S+ RLCV YR+ LNKVTIKNKY LPR++
 Sbjct 64 DGSLRLCVDYRELNKVTIKNKYPLPRID 91

AD-P2391MCTA-13D (213 letters)

Sequences producing significant alignments:	Score (Bits)	E Value
gi 4826399 emb CAB42872.1 wall-associated kinase 2 [Arabidop...	63.5	2e-09
gi 15218170 ref NP_173546.1 WAK5 (WALL ASSOCIATED KINASE 5);...	63.5	2e-09

Alignments

>gi|4826399|emb|CAB42872.1| wall-associated kinase 2 [Arabidopsis thaliana]

gi|24417340|gb|AAN60280.1| unknown [Arabidopsis thaliana] Length=732
 Score = 63.5 bits (153), Expect = 2e-09
 Identities = 31/66 (46%), Positives = 41/66 (62%), Gaps = 3/66 (4%)
 Frame = -3

Query 202 LDDVENLP-YQIFPVFNWSVGNEICKESLSRGTNACKGNSICVDAETGYAYQCQCNCQCF 26
 L+D+ NL FVVV + WS+G++ CK+ RG C GNS C D+ G Y C+C +F
 Sbjct 209 LEDLNLRNVTTFPVVDWSIGDKTCKQVEYRGV--CGGNSTCFDSTGGTGYNCKCLEGF 266

Query 25 EGYPYL 8
 EG PYL
 Sbjct 267 EGNPYL 272

AD-P2414MC-17B (301 letters)

Score E

Sequences producing significant alignments:	(Bits)	Value
gi 6691193 gb AAF24531.1 F7F22.17 [Arabidopsis thaliana]	152	4e-36
gi 15724047 gb AAL06421.1 reverse transcriptase [Arabidopsis th]	151	8e-36

Alignments

>gi|6691193|gb|AAF24531.1| F7F22.17 [Arabidopsis thaliana] Length=1799

Score = 152 bits (384), Expect = 4e-36
 Identities = 70/93 (75%), Positives = 77/93 (82%), Gaps = 0/93 (0%)
 Frame = -2

Query 291 HYCFLDGYTYFYQIHAHED*EKSTFTCPFGTFAYKRMPFGLCNAPATFQRCMTSVFFDL 112
 +YCFLDGY G+FQI I D EK+TFTCP+GTFAYKRMPFGLCNAPATFQRCMTS+F DL
 Sbjct 1047 YYCFLDGYNGFFQIPIHPNDQEKTTFTCPYGTFAFKRMPFGLCNAPATFQRCMTSIFSDL 1106

Query 111 IESCLEVMDDFSVYETSFDSCLENLAKVLETC 13
 IE +EVFMDDFSVY SF SCL NL +VL C

Sbjct 1107 IEEMVEVMDDFSVYGPSFSSCLLNLGRVLTRC 1139

AD-P2414MCTA-15B (255 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 68129555 emb CAJ08098.1 actin-like protein, putative [Leishm]	34.3	1.5
gi 48788696 ref ZP_00284675.1 hypothetical protein Bcep02001175	32.7	4.3

Alignments

>gi|68129555|emb|CAJ08098.1| actin-like protein, putative [Leishmania major] Length=423

Score = 34.3 bits (77), Expect = 1.5
 Identities = 21/69 (30%), Positives = 36/69 (52%), Gaps = 5/69 (7%)
 Frame = +3

Query 21 PSDMDGYQEYVYSEKSLDLTGT*SSKSLLNQKHFGP-----MVNSGCDPSALTHIFSL*R 185
 P++ D + Y+ KSLD SS+ N++ GP SG P+AL ++ SL+
 Sbjct 34 PEEVDLISLLYAAGKSLDTAPLPSSSAFANEEAGGPADSLLSQGSAPPNALLPLWSLRQ 93

Query 186 LELGPRSYL 212
 +E P+++

Sbjct 94 VERHPKQHI 102

AD-P2414MCTA-15A (297 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 33331722 gb AAQ11030.1 hypothetical protein [Mamestra con...]	34.7	1.1
gi 20069889 ref NP_613093.1 unknown [Mamestra configurata NP...]	34.7	1.1

Alignments

>gi|33331722|gb|AAQ11030.1| hypothetical protein [Mamestra configurata nucleopolyhedrovirus A] Length=319

Score = 34.7 bits (78), Expect = 1.1
 Identities = 15/34 (44%), Positives = 19/34 (55%), Gaps = 0/34 (0%)
 Frame = +1

Query 34 PLWNPCCTLGSLATGMRTLLYWALLFIKLRPACS 135
 P+++ CC LA + LY LLF NRP CS
 Sbjct 186 PMYDVCCAYNGLAHSVALLNYIPLLFYLNRPVCS 219

AS-NBSMGGC-8E (348 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 55770171 ref XP_550075.1 hypothetical protein [Oryza sati...]	44.7	0.001
gi 62734238 gb AAX96347.1 retrotransposon protein, putative,...	43.1	0.003

Alignments

>gi|55770171|ref|XP_550075.1| hypothetical protein [Oryza sativa (japonica cultivar-group)]

gi|54290304|dbj|BAD61304.1| hypothetical protein [Oryza sativa (japonica cultivar-group)] Length=385

Score = 44.7 bits (104), Expect = 0.001
 Identities = 29/117 (24%), Positives = 58/117 (49%), Gaps = 19/117 (16%)
 Frame = -3

Query 343 DDVWLLGGAIKDRFPRLCSVSNQEGSVIGACGFWDGLEWIWNFQWKREL-----FQWELE 179
 D WL G +KD P++ S++ S+ D QW ++ L W + E
 Sbjct 123 DSGWLQGRRLKDVAPLVYAASKKKTSTLQQASLSD-----QWMQDLDPENTGWSIE 174

Query 178 LLNQLHDI---LRPVLVQNIEDRVVWKYDGGQIFSTKSFVQVWQEEMLPKEITSFS 17
 L++QL ++ +++ L+++ ED++ WK G++ TS +++L T+F+

Sbjct 175 LIDQLIEVSVVQNLHLIEHEEDKITWKLTSHEGYTTTS---AYKAQLLGTATNFN 228

AS-P2378MCT-20H (121 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 15724009 gb AAL06413.1 reverse transcriptase [Pisum sativum]	47.4	2e-04
gi 15724013 gb AAL06415.1 reverse transcriptase [Pisum sativum]	47.0	2e-04

Alignments
>gi|15724009|gb|AAL06413.1| reverse transcriptase [Pisum sativum] Length=254
 Score = 47.4 bits (111), Expect = 2e-04
 Identities = 19/27 (70%), Positives = 24/27 (88%), Gaps = 0/27 (0%)
 Frame = +1
 Query 40 LVPTSTITGWRMFIDYMLRHTATRKDH 120
 L+ PT T TGWR+ ID+ RL+TATRKDH
 Sbjct 45 LIPTRTVTGWRVCIDHRLNTATRKDH 71

AA-P2378MC-20D (312 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 49533772 gb AAT66771.1 putative polyprotein [Solanum demissu]	123	2e-27
gi 47900698 gb AAT39297.1 putative gag-pol protein [Solanum dem	122	4e-27

Alignments
>gi|49533772|gb|AAT66771.1| putative polyprotein [Solanum demissum] Length=1769
 Score = 123 bits (308), Expect = 2e-27
 Identities = 57/87 (65%), Positives = 69/87 (79%), Gaps = 0/87 (0%)
 Frame = -2
 Query 293 DIDEFPPNREVEFAIALVPGAGPISSAPYRMSPLEMSDLKSQLEDFLGKNFIRPSVSPWG 114
 D+ PP+R+++ FAI L P PIS PYRM+P E+ + L +QLED LGK FIRPSVSPWG
 Sbjct 826 DLPGLPPDRDIDFAIDLEPDTRPISIPPYRMAPAELRELSAQLEDLLGKGFIRPSVSPWG 885

 Query 113 APV*LVKKKDGSMRLCVDRLQNLKQVTV 33
 APV VKKKDG+MR+C+D RQLNKVTV
 Sbjct 886 APVLFVKKKDGTMRCIDYRQLNKVTV 912

AD-NBSMGAG-22C (199 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 4567295 gb AAD23707.1 putative retroelement pol polyprotein	92.8	3e-18
gi 9884624 dbj BAB02630.1 retroelement pol polyprotein-like [Ar	89.4	4e-17

Alignments
>gi|4567295|gb|AAD23707.1| putative retroelement pol polyprotein [Arabidopsis thaliana] Length=466
 Score = 92.8 bits (229), Expect = 3e-18
 Identities = 43/59 (72%), Positives = 49/59 (83%), Gaps = 0/59 (0%)
 Frame = -2
 Query 183 LLKKHGILHKIATTYHPQTNGQAEVSNREIKRIL*KTVKSHRKDWSTRLVDALWAYHTS 7
 LLKKHG+ HK+AT YHPQT+GQ E+SNREIK IL KTV RKDWST+L DALWAY T+
 Sbjct 246 LLKKHGVKHKVATPYHPQTSGQVEISNREIKTILEKTVGITRKDWSTKLDDALWAYKTA 304

AD-P2378MC-20B (283 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 37534632 ref NP_921618.1 putative pol polyprotein [Oryza ...	65.5	6e-10
gi 78708691 gb ABB47666.1 retrotransposon protein, putative,...	65.5	6e-10

Alignments
>gi|37534632|ref|NP_921618.1| putative pol polyprotein [Oryza sativa (japonica cultivar-group)] Length=1688
 Score = 65.5 bits (158), Expect = 6e-10
 Identities = 33/63 (52%), Positives = 41/63 (65%), Gaps = 0/63 (0%)
 Frame = -2
 Query 225 LSYFLALEVISTDDDIYLSQANYASDLLTRAGIIDSHTTEPTLEPNVRFPMVDVLDNP 46
 LSYFL+EV ST D YLSQ Y DLL ++ G+ DS T TP+E + VR D T LD+P
 Sbjct 983 LSYFLGIEVTSTVDGYLSQHRYIEDLLAQSGLTDSRTTTTPMELHVRLRSTDGTPPLDDP 1042

 Query 45 TLY 37
 + Y
 Sbjct 1043 SRY 1045

AS-P2378MC-20H (256 letters)

	Score	E
Sequences producing significant alignments:	(Bits)	Value
gi 15724009 gb AAL06413.1 reverse transcriptase [Pisum sativum]	68.2	9e-11
gi 15724013 gb AAL06415.1 reverse transcriptase [Pisum sativum]	67.4	2e-10

Alignments

>gi|15724009|gb|AAL06413.1| reverse transcriptase [Pisum sativum] Length=254

Score = 68.2 bits (165), Expect = 9e-11
 Identities = 29/38 (76%), Positives = 35/38 (92%), Gaps = 0/38 (0%)
 Frame = +1

Query 22 TLIQNEKHELIPTRTITGWRMCIDYMLNLTATRKDHSP 135
 ++I NEK+ELIPTRT+TGWR+CID+ RLNTATRKDH P
 Sbjct 36 SVILNEKNELIPTRTVTGWRVCIDHRLNLTATRKDHFP 73

AD-P216MC-28G (126 letters)

	Score	E
Sequences producing significant alignments:	(Bits)	Value
gi 37221993 gb AAN85388.1 resistance protein [Arachis cardenasi]	52.0	7e-06
gi 37222003 gb AAN85393.1 resistance protein [Arachis cardenasi]	50.8	2e-05

Alignments

>gi|37221993|gb|AAN85388.1| resistance protein [Arachis cardenasi] Length=156

Score = 52.0 bits (123), Expect = 7e-06
 Identities = 22/28 (78%), Positives = 25/28 (89%), Gaps = 0/28 (0%)
 Frame = +3

Query 3 FGPGSRILVTTTRDVHVLKGPEVHDIYKV 86
 FGPG RI++ TTRDV VLKGPEVH+ YKV
 Sbjct 91 FGPGCRIITTRDVEVLKGPEVHETYKV 118

AS-NBSMCTT-24E (320 letters)

	Score	E
Sequences producing significant alignments:	(Bits)	Value
gi 48057663 gb AAT39962.1 putative late blight resistance prote	108	6e-23
gi 48057628 gb AAT39951.1 putative disease resistance protein P	95.9	4e-19

Alignments

>gi|48057663|gb|AAT39962.1| putative late blight resistance protein [Solanum demissum] Length=1159

Score = 108 bits (270), Expect = 6e-23
 Identities = 54/102 (52%), Positives = 69/102 (67%), Gaps = 1/102 (0%)
 Frame = -3

Query 138 SRSEVLGGEECPADLEPIGESIAKSCKGLPLAIVTIAGVVMK 13
 +EV G E C LE +G+ IA+ C GLPL+IV +AG++ K
 Sbjct 675 LCTEVFGNESCFFLLEKVGQEIARKCGGLPLSIVLVAGILSK 716

AD-NBSMCTC-14A (198 letters)

	Score	E
Sequences producing significant alignments:	(Bits)	Value
gi 62632823 gb AAX89382.1 NBS-LRR type disease resistance prote	63.2	3e-09
gi 62632825 gb AAX89383.1 NBS-LRR type disease resistance prote	58.5	7e-08

Alignments

>gi|62632823|gb|AAX89382.1| NBS-LRR type disease resistance protein Rps1-k-1 [Glycine max] Length=1229

Score = 63.2 bits (152), Expect = 3e-09
 Identities = 37/65 (56%), Positives = 44/65 (67%), Gaps = 0/65 (0%)
 Frame = -1

Query 198 RIRRVASVVQTVXXXXXXXXXXEDCWLLFSKHERLSTFSMENPPLTSFGRDLVKKCDGLP 19
 R + ASVVQTV YL+ LS EDCW +F+ H LS S EN L G+++VKKCDGLP
 Sbjct 301 RSEKTASVVQTVQTYHLNQLSNEDCWSVFANHACLSLESNENTTLEKIGKEIVKKCDGLP 360

Query 18 LAALS 4

LAA S

Sbjct 361 LAAQS 365

AD-NBSMCTT-24F (93 letters)

	Score	E
Sequences producing significant alignments:	(Bits)	Value
gi 48057663 gb AAT39962.1 putative late blight resistance prote	40.8	0.016
gi 53854439 gb AAU95620.1 R1 [Solanum tuberosum]	37.7	0.13

Alignments

>gi|48057663|gb|AAT39962.1| putative late blight resistance protein [Solanum demissum] Length=1159
 Score = 40.8 bits (94), Expect = 0.016
 Identities = 17/29 (58%), Positives = 21/29 (72%), Gaps = 0/29 (0%)
 Frame = -1
 Query 93 MWDNKAWNKLKDYLSPGQNNNGSRILMTTRN 7
 +WD K W+ L P +NNGSRIL+ TTRN
 Sbjct 619 VWDCKVWDDLKCCFPDRNNGSRILLTTRN 647

AD-NBSMGGC-8D (350 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 37221985 gb AAN85384.1 resistance protein [Arachis cardenasi	118	6e-26
gi 50948145 ref XP_483600.1 putative NBS-LRR resistance prot...	84.7	1e-15

Alignments

>gi|37221985|gb|AAN85384.1| resistance protein [Arachis cardenasii] Length=339
 Score = 118 bits (296), Expect = 6e-26
 Identities = 63/108 (58%), Positives = 76/108 (70%), Gaps = 3/108 (2%)
 Frame = -3
 Query 348 DVWDNDKWEDLKSVLNSRGSESKGVSYLVTTTRDQIVASAMETCPTQSHHLQSLPKDDNWS 169
 DVW DKWEDLKSVL G + KG+ LVTTTR+ VAS M TCP + HHL L + DDNW
 Sbjct 57 DVWSTDKWEDLKSVLLCGGG-TKGAALVTTTRVESVASVMGTCP--AHHLSPLSEDDNWL 113

Query 168 LFTHYAFDPNKEQPVKLVEIGKEIV*RCVAYPLASKVVGSLLRNKKEE 25
 LF ++ AF + K + +L IGK+IV + C PLASK + GSLLRNKKEE
 Sbjct 114 LFKYHAFGSDKVERTELEAIGKKIVKKCGGSPLASKALGSLLRNKKEE 161

AD-P212MCT-9C (211 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 2827714 emb CAA16687.1 receptor protein kinase - like pro...	65.5	6e-10
gi 30698151 ref NP_201372.2 ATP binding / kinase/ protein se...	65.5	6e-10

Alignments

>gi|2827714|emb|CAA16687.1| receptor protein kinase - like protein [Arabidopsis thaliana]
 gi|10177329|dbj|BAB10678.1| receptor protein kinase-like protein [Arabidopsis thaliana] Length=976
 Score = 65.5 bits (158), Expect = 6e-10
 Identities = 32/70 (45%), Positives = 49/70 (70%), Gaps = 0/70 (0%)
 Frame = +2
 Query 2 IPDQYGDCDSLEYVRIENNQFSGQVPPKFWSLPKLQLQMANNRFEGXXXXXXXXXARGIT 181
 IP+ YGDC SL Y+R+ +N+ SG+VP +FW LP +L NN+ + GS+ S+S AR ++
 Sbjct 403 IPESYGDCHSLNYIRMADNKLSGEVPARFWELPLTRLELANNNQLQGSIPPSISKARHLS 462

Query 182 VLLLSGSSFS 211
 L +S ++FS
 Sbjct 463 QLEISANNFS 472

AD-P212MCT-9A (241 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 37221969 gb AAN85376.1 resistance protein [Arachis carden...	59.7	3e-08
gi 37221989 gb AAN85386.1 resistance protein [Arachis cardenasi	58.2	9e-08

Alignments

>gi|37221969|gb|AAN85376.1| resistance protein [Arachis cardenasii]
 gi|37221911|gb|AAO20356.1| resistance protein [Arachis stenoperma] Length=159
 Score = 59.7 bits (143), Expect = 3e-08
 Identities = 27/29 (93%), Positives = 27/29 (93%), Gaps = 0/29 (0%)
 Frame = -3
 Query 107 GHTWFGASSRITVTTTRDKHVLVTADADHI 21
 GHTWFGASSRI VTTTRDKHVLV ADADHI
 Sbjct 89 GHTWFGASSRIIVTTTRDKHVLVKADADHI 117

AD-P212MCT-9B (97 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 37221969 gb AAN85376.1 resistance protein [Arachis carden...	64.3	1e-09
gi 37221989 gb AAN85386.1 resistance protein [Arachis cardenasi	62.8	4e-09

Alignments

>gi|37221969|gb|AAN85376.1| **resistance protein [Arachis cardenasii]**
 gi|37221911|gb|AAO20356.1| resistance protein [Arachis stenosperma] Length=159
 Score = 64.3 bits (155), Expect = 1e-09
 Identities = 29/31 (93%), Positives = 29/31 (93%), Gaps = 0/31 (0%)
 Frame = +3
 Query 3 LCEGHTWFGASSRIIVTTRDKHVLVTADADH 95
 LC GHTWFGASSRIIVTTRDKHVLV ADADH
 Sbjct 86 LCGGHTWFGASSRIIVTTRDKHVLVKADADH 116

AD-P216MCAA-26B (198 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 1663547 gb AAC49508.1 disease resistance protein homolog	75.9	4e-13
gi 37221937 gb AAO20369.1 resistance protein [Arachis hypogaea]	73.9	2e-12

Alignments

>gi|1663547|gb|AAC49508.1| **disease resistance protein homolog Length=191**
 Score = 75.9 bits (185), Expect = 4e-13
 Identities = 35/48 (72%), Positives = 39/48 (81%), Gaps = 0/48 (0%)
 Frame = +2
 Query 2 LLVLDDVSEVKLENLGENQDWFGPGSRIITARDVHLLLEIHGVHGTC 145
 LLVLDDVSE+ LENL Q+WFGPGSR+IIT RD HLL+ HG VH TC
 Sbjct 99 LLVLDDVSELSQLENLAGKQEWFGPGSRVIITTRDKHLLKTHGVHLTC 146

AS-P216MCTC-27A (403 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 74231152 gb ABA00702.1 disease resistance protein [Phaseolus	75.5	6e-13
gi 33341541 gb AAQ15191.1 resistance protein [Vitis vinifera]	72.0	6e-12

Alignments

>gi|74231152|gb|ABA00702.1| **disease resistance protein [Phaseolus vulgaris] Length=753**
 Score = 75.5 bits (184), Expect = 6e-13
 Identities = 44/94 (46%), Positives = 59/94 (62%), Gaps = 4/94 (4%)
 Frame = +2
 Query 128 DDIYKVALESSDAQLLNLHAFDGTGDGHCLERWHNLVSKMVEYAKGIPLGLKVLGSFL 307
 DDIY+ VK L D+Q+L NLHAF E ++ L +++ YA+GIPL L++ LGS L
 Sbjct 324 DDIYEVKELNLLDSDQKLFNLHAFKQKSS--AEKSYMDLSEEVLYGAEIPLALQILGSL 381
 Query 308 YGKCFEDWGCQLRKLQKMPQLQEIHNV--LSYSGL 403
 YG+ E W +L+KL+K I NV LSY GL
 Sbjct 382 YGRTREAWESELQKLKKGQHLGIFNVLKLSYDGL 415

AD-P2378MCTA-23A (312 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 37534632 ref NP_921618.1 putative pol polyprotein [Oryza ...	85.5	5e-16
gi 78708691 gb ABB47666.1 retrotransposon protein, putative,...	85.5	5e-16

Alignments

>gi|37534632|ref|NP_921618.1| **putative pol polyprotein [Oryza sativa (japonica cultivar-group)] Length=1688**
 Score = 85.5 bits (210), Expect = 5e-16
 Identities = 43/78 (55%), Positives = 52/78 (66%), Gaps = 0/78 (0%)
 Frame = -1
 Query 270 KGSLQHTFEMKDLGFLSYFLGLEVISTDDDIYLSQAKYASDLLTRAGIIDSHTPTLEP 91
 KG L F M DLG LSYFLG+EV ST D YLSQ+ Y DLL ++ G+DS T TP+E
 Sbjct 968 KGKLSQFMMSDLGPLSYFLGIEVTSTVDGYLSQHRYIEDLLAQSGLTDSRTTTTPMEL 1027
 Query 90 NVRFTPMDDGTVDNPTLY 37
 +VR DGT LD+P+Y
 Sbjct 1028 HVRLRSTDDGTPLDDPSRY 1045

AD-P2391MGAG-16B (262 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 1853970 dbj BAA13542.1 CPRD46 protein [Vigna unguiculata]	67.0	2e-10
gi 77556128 gb ABA98924.1 lipoygenase, putative [Oryza sativa]	64.7	1e-09

Alignments

>gi|1853970|dbj|BAA13542.1| CPRD46 protein [Vigna unguiculata] Length=899

Score = 67.0 bits (162), Expect = 2e-10
 Identities = 43/52 (82%), Positives = 47/52 (90%), Gaps = 0/52 (0%)
 Frame = -1
 Query 157 QSYLPDTPXXXXXXXXXXXXXXXXXGNGHGERKTFDRIYDYDVYNDLGDPDND 2
 +SYLPS+TPEG+KR+R EELQ LRGNG GERK FDRYDYDVYNDLGDPD D
 Sbjct 205 KSYLPSETPEGVKRIRREEELQHLRGNGQGERKKFDRIYDYDVYNDLGDPDKD 256

AD-P2414MCTC-10C (364 letters)

Sequences producing significant alignments:	Score	E
gi 16118447 gb AAL12248.1 heat shock transcription factor [Phas	(Bits)	Value
gi 20162459 gb AAM14595.1 heat shock transcription factor [M...	102	3e-21
	69.7	3e-11

Alignments

>gi|16118447|gb|AAL12248.1| heat shock transcription factor [Phaseolus acutifolius] Length=402

Score = 102 bits (255), Expect = 3e-21
 Identities = 50/91 (54%), Positives = 68/91 (74%), Gaps = 1/91 (1%)
 Frame = -1
 Query 295 VTLTLTENMERKRRLPKGSSHTDEASIEDPMEISQVLPSENAESSDFVTSCIERMNQL 116
 V+L LTE M+RKRRRLP+ S H+ DE+SID +E SQ+LP ENAE++ + T + ER++QL
 Sbjct 192 VNILPLTETMDRKRRLPR-SGHYYDESSIEDAIETSQMLPRENAENTTVLTLNVERLDQL 250

Query 115 ESTLVFWEYIAQDVNETLVQSHSNLDVDEST 23
 ES++ FWE IA D+ + Q SN+D DEST
 Sbjct 251 ESSVAFWEAIAHDIGDNFAIQSNMDFDEST 281

AS-NBSMC-21A (395 letters)

Sequences producing significant alignments:	Score	E
gi 23508536 ref NP_701205.1 hypothetical protein PF11_0345 [...	(Bits)	Value
gi 68058920 ref XP_671435.1 hypothetical protein PB108262.00...	34.3	1.4
	33.1	3.2

Alignments

>gi|23508536|ref|NP_701205.1| hypothetical protein PF11_0345 [Plasmodium falciparum 3D7]

gi|23496271|gb|AAN35929.1| hypothetical protein PF11_0345 [Plasmodium falciparum 3D7] Length=392
 Score = 34.3 bits (77), Expect = 1.4
 Identities = 18/42 (42%), Positives = 26/42 (61%), Gaps = 4/42 (9%)
 Frame = -2
 Query 271 KGNENLQNVYKS---YKRMGEFIEIRRYCLFGLNPAHILFNH 155
 K ENLQN+YKS Y + I++ CL+ LN HIL+++
 Sbjct 112 KSEENLQNIYKSSNLYNINP-IYLSKICLYNLNEYHILYSN 152

AS-NBSMCTC-14D (356 letters)

Sequences producing significant alignments:	Score	E
gi 48057663 gb AAT39962.1 putative late blight resistance prote	(Bits)	Value
gi 37531744 ref NP_920174.1 putative disease resistance prot...	43.1	0.003
	43.1	0.003

Alignments

>gi|48057663|gb|AAT39962.1| putative late blight resistance protein [Solanum demissum] Length=1159

Score = 43.1 bits (100), Expect = 0.003
 Identities = 20/32 (62%), Positives = 24/32 (75%), Gaps = 1/32 (3%)
 Frame = -2
 Query 355 DDVWNTQVWNKLYLLPKHNNGSRILVTTRN 260
 DDVW+ + VW+ LK P NNGSRIL+TTRN
 Sbjct 617 DDVWDCVWDDLKC-CFPDRNNGSRILLTTRN 647

AS-NBSMCTT-24D(357 letters)

Sequences producing significant alignments:	Score	E
gi 9964416 ref NP_064884.1 Hypothetical protein [Amsacta moo...	(Bits)	Value
gi 28563969 gb AAO32363.1 ARE2 [Saccharomyces bayanus] >gi 3...	37.4	0.17
	34.7	1.1

Alignments

>gi|9964416|ref|NP_064884.1| Hypothetical protein [Amsacta moorei entomopoxvirus]

gi|9944625|gb|AAG02808.1| AMV102 [Amsacta moorei entomopoxvirus] Length=148
 Score = 37.4 bits (85), Expect = 0.17
 Identities = 18/60 (30%), Positives = 33/60 (55%), Gaps = 3/60 (5%)
 Frame = -2
 Query 260 NIFCKKKEKTTKNHVNMRK*SNILLHYIKEYKRVSYNNNDILYSISVSMKVVKRIKS 81

+ C ++KN+ H M K S+ LH +++ YK Y+N D+LY ++ + +K K+
 Sbjct 54 DFLCLGDNRLSKNILHLMNKYSTVNLHLKLEQGYK---YFNMDMLYDLFIPFRYIKIYKN 110

AS-NBSMCTT-24A (512 letters)

Sequences producing significant alignments:	Score (Bits)	E Value
gi 37221985 gb AAN85384.1 resistance protein [Arachis cardenasii	215	5e-55
gi 39636771 gb AAR29073.1 blight resistance protein B149 [Solan	145	8e-34

Alignments

>**gi|37221985|gb|AAN85384.1| resistance protein [Arachis cardenasii] Length=339**

Score = 215 bits (548), Expect = 5e-55
 Identities = 109/173 (63%), Positives = 134/173 (77%), Gaps = 7/173 (4%)
 Frame = -3

Query 510 DVWENDKWEDLMSVLNSRGSETKGVLSILVTRDQIVASAMETFPQTQSHHLQPLPKDENWS 331
 DVW DKWEDL SVL G TKG ++LVTRR + VAS M T P +HHL PL + D+NW
 Sbjct 57 DVWSTDKWEDLKSIVLLCGGG-TKGAALVTVRVSASVMGTCP--AHHLSPLSEDDNWL 113

Query 330 LFTHYAFGQNKQPAKLVEIGKEIVRRCVGNPLASKVVGSLLRNKKEEKEWLNVLKESKFW 151
 LF ++ AFG + K + +L IGK+IV++C G+PLASK + GSLLRNKKEE +W+NVLESKFW
 Sbjct 114 LFKYHAFGSDKVERTELEAIGKIVKKCGGSPLASKALGSLLRNKKEEIQWVNVLESKFW 173

Query 150 DI---DA-VMGAL*LSYFHLKPSSRQCFCFCALYPEDFRISKEQLIHLWMANG 4
 DI DA ++ AL +SYFHLK S RQCF FCA++ PED+R+ KEQLIHLWMANG
 Sbjct 174 DILEDDAIIVRALKISYFHLKLSLRQCFAFCAIFPEDYRMEKEQLIHLWMANG 226

AS-NBSMCTT-24H (397 letters)

Sequences producing significant alignments:	Score (Bits)	E Value
gi 51535606 dbj BAD37549.1 receptor protein kinase PERK1-lik...	164	1e-39
gi 50905775 ref XP_464376.1 receptor protein kinase PERK1-li...	152	4e-36

Alignments

>**gi|51535606|dbj|BAD37549.1| receptor protein kinase PERK1-like [Oryza sativa (japonica cultivar-group)] Length=392**

Score = 164 bits (414), Expect = 1e-39
 Identities = 81/105 (77%), Positives = 93/105 (88%), Gaps = 4/105 (3%)
 Frame = +3

Query 93 QIAVKKLLKAMN-SKAEMEFAVEVEVLGRVRHKNNLLGLRGYCAGN---EERLIVYDYMANL 260
 QIAVKKLLKA N SKAEMEFAVEVEVL RVRHKNNLLGLRGYCAG ++R+IVYDYM NL
 Sbjct 68 QIAVKKLLKATNTSKAEMEFAVEVEVLARVRHKNNLLGLRGYCAGGAAGDQRMIVYDYMPNL 127

Query 261 SLLSHLHGQFANEVQLDWKRRMKIAIGSAEGVLYLHHDVTPHIIIH 395
 SLLSHLHGQFA +V+LDW RRM +A+G+AEG+++LHH+ TPHIIIH
 Sbjct 128 SLLSHLHGQFAADVRLDWARRMAVAVGAAEGLVHLHHEATPHIIIH 172

AS-NBSMGGC-8C (409 letters)

Sequences producing significant alignments:	Score (Bits)	E Value
gi 62632823 gb AAX89382.1 NBS-LRR type disease resistance prote	169	3e-41
gi 62632825 gb AAX89383.1 NBS-LRR type disease resistance prote	167	8e-41

Alignments

>**gi|62632823|gb|AAX89382.1| NBS-LRR type disease resistance protein Rps1-k-1 [Glycine max] Length=1229**

Score = 169 bits (428), Expect = 3e-41
 Identities = 82/135 (60%), Positives = 105/135 (77%), Gaps = 1/135 (0%)
 Frame = -1

Query 409 VVDDVWSESYEDWNKLLKPFKGV-KGSKILITTRSKRVASVVQTVSPYELLSLSEEDCW 233
 V+DDVW+E Y DW+ L KPF+ G+ + SKIL+TTRS++ ASVVQTV Y L+ LS EDCW
 Sbjct 267 VLDDVWTEDEVVDWLLKPFQCGIIRRSKILLTTRSEKTASVVQTVQTYHLNQLSNEEDCW 326

Query 232 LLFSKHARLSTVSMENPTLKKVGRDLVKKCDGLPLAAQSLGGLLRGNSDIKYWNHLLKSE 53
 +F+ HA LS S EN TL+K+G+++ VKKCDGLPLAAQSLGG+LR DI W ++L S+
 Sbjct 327 SVFANHACL SLESENTTLEKIGKEIVKKCDGLPLAAQSLGGMLRRKHDIKDWDYNYLNLSND 386

Query 52 IWELSDDTIKVPAL 8
 IWELS+ KV+ PAL

Sbjct 387 IWELSESECKVIPAL 401

AS-NBSMGGC-8J (173 letters)

Sequences producing significant alignments:	Score (Bits)	E Value
---	-----------------	------------

gi 37221985 gb AAN85384.1 resistance protein [Arachis cardenasi	71.2	1e-11
gi 5817341 gb AAD52714.1 putative NBS-LRR type disease resis...	50.8	2e-05

Alignments

>**gi|37221985|gb|AAN85384.1| resistance protein [Arachis cardenasi]** Length=339

Score = 71.2 bits (173), Expect = 1e-11

Identities = 36/54 (66%), Positives = 39/54 (72%), Gaps = 1/54 (1%)

Frame = +3

```
Query 3  VDDVVDNVKQWEDLKSVLNCRGSGIKGVSVLVTTRDPSVASVMTPCNSHHLQPYS 164
          +DDVW  KWEDLKSVL C G G KG +  VLVTR SVASVM C +HHL P S
Sbjct 55  LDDVWSTDKWEDLKSVMLLC-GGGTKGA AVLVTTRVESVASVMGTCPAHHLSPLS 107
```

AS-P216MCAA-26A (249 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 37221893 gb AAO20347.1 resistance protein [Arachis stenosp	63.2	3e-09
gi 1663547 gb AAC49508.1 disease resistance protein homolog	62.0	6e-09

Alignments

>**gi|37221893|gb|AAO20347.1| resistance protein [Arachis stenosp** Length=156

Score = 63.2 bits (152), Expect = 3e-09

Identities = 27/38 (71%), Positives = 33/38 (86%), Gaps = 0/38 (0%)

Frame = +2

```
Query 2  LVLDDVSEVKCLENLGENQDWFGPGSRRIITARDMHLI 115
          +VLDDV++V  LENL  NQDWFGPGSR++IT RDMHL+
Sbjct 70  IVLDDVNDVSQLENLAGNQDWFGPGSRVMITTRDMHLL 107
```

AS-P216MCTC-27D (395 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 8778513 gb AAF79521.1 F21D18.16 [Arabidopsis thaliana]	56.2	4e-07
gi 15221080 ref NP_175246.1 hydrolase/ protein serine/threon...	56.2	4e-07

Alignments

>**gi|8778513|gb|AAF79521.1| F21D18.16 [Arabidopsis thaliana]** Length=1340

Score = 56.2 bits (134), Expect = 4e-07

Identities = 32/83 (38%), Positives = 42/83 (50%), Gaps = 11/83 (13%)

Frame = +2

```
Query 92  QFIWTPYMGVGVDPDLATHLVIC-----ST*SPLVSFECIEWHAADRVRQFGI*QLP 250
          Q IW PY  PD+LA  +IC  T +PL+ F+ +  EWH  DRV RQFG+ Q
Sbjct 318  QVIWQPY----TPDLLAKIPLICVSGENIWRVAPLICFDVVEWHRPDRVLRQFGLHQT 373
```

```
Query 251 PSPAFLGHDHCKRLTGAQNHDW 319
          P+P +  H  G  +DW
Sbjct 374  PAPCDNEKALHAIDKRGKSEYDW 396
```

AS-P2378MCTC-11B (431 letters)

Sequences producing significant alignments:	Score	E
	(Bits)	Value
gi 47824972 gb AAT38745.1 putative polyprotein, 3'-partial [Sol	222	4e-57
gi 49533772 gb AAT66771.1 putative polyprotein [Solanum demissu	221	8e-57

Alignments

>**gi|47824972|gb|AAT38745.1| putative polyprotein, 3'-partial [Solanum demissum]** Length=1622

Score = 222 bits (565), Expect = 4e-57

Identities = 101/138 (73%), Positives = 117/138 (84%), Gaps = 0/138 (0%)

Frame = -1

```
Query 416 DIDEFPPNREVDFAIELVPGAGPISSAPYRMSPLEMVELKSQLEDLLGKNFIRPSISPWG 237
          D+  PP R++DF IEL PG  PIS  PYRM+P E+  EL  QL+DLLGK FIRPS+SPWG
Sbjct 789  DLPGLPPERDIDFPIELEPGTRPISIPPYRMAPAELKELSVLQDQLLGGKGFIRPSVSPWG 848
```

```
Query 236 APMLLVKKKDGSMRLCVDYRQLNKVTIKNKYPLPRIDDLMDQLQGAEIFSKIDLRSDYHQ 57
          AP+L VKKKDG+MR+C+DYRQLNKVT+KN+YPLPRIDDL  DQLQGA +FSKIDLR DYHQ
Sbjct 849  APVLFVKKKDGTMRCIDYRQLNKVTIKNRYPLPRIDDLFDQLQGASVFSKIDLRFDYHQ 908
```

```
Query 56  IRVRGEDIPKTAFRTRYG 3
          +R+R  DIPKTAFRTRYG
Sbjct 909  LRIRAADIPKTAFRTRYG 926
```

AS-P2414MCTT-19A (387 letters)

	Score	E
Sequences producing significant alignments:	(Bits)	Value
gi 50930529 ref XP_474792.1 OSJNBa0014F04.3 [Oryza sativa (j...	43.5	0.002
gi 37531452 ref NP_920028.1 putative retroelement [Oryza sat...	43.5	0.002

Alignments

>gi|50930529|ref|XP_474792.1| OSJNBa0014F04.3 [Oryza sativa (japonica cultivar-group)]
 gi|39545593|emb|CAE02837.3| OSJNBa0014F04.3 [Oryza sativa (japonica cultivar-group)] Length=230
 Score = 43.5 bits (101), Expect = 0.002
 Identities = 18/32 (56%), Positives = 22/32 (68%), Gaps = 0/32 (0%)
 Frame = +2
 Query 200 LATFLKVNPPNFRGATNPMEADNWFQAMERAL 295
 L FL+ V PP F TNPMEA + WF A +E+ L
 Sbjct 21 LLEFLRVRPPTFSSTTNPMEAKDWFHAIKLL 52

AD-P212MCT-9C (102 letters)

	Score	E
Sequences producing significant alignments:	(Bits)	Value
gi 37221969 gb AAN85376.1 resistance protein [Arachis carden...	71.6	8e-12
gi 37221989 gb AAN85386.1 resistance protein [Arachis cardenasi	70.1	2e-11

Alignments

>gi|37221969|gb|AAN85376.1| resistance protein [Arachis cardenasii]
 gi|37221911|gb|AAO20356.1| resistance protein [Arachis stenoperma] Length=159
 Score = 71.6 bits (174), Expect = 8e-12
 Identities = 32/33 (96%), Positives = 32/33 (96%), Gaps = 0/33 (0%)
 Frame = -2
 Query 101 EDLCGGHTWFGASSRIIVTTRDKHVLVTADADH 3
 EDLCGGHTWFGASSRIIVTTRDKHVLV ADADH
 Sbjct 84 EDLCGGHTWFGASSRIIVTTRDKHVLVKADADH 116

AS-P2378MCTA-23E (246 letters)

	Score	E
Sequences producing significant alignments:	(Bits)	Value
gi 15724009 gb AAL06413.1 reverse transcriptase [Pisum sativum]	70.1	2e-11
gi 15724020 gb AAL06416.1 reverse transcriptase [Sorghum bicolo	69.3	4e-11

Alignments

>gi|15724009|gb|AAL06413.1| reverse transcriptase [Pisum sativum] Length=254
 Score = 70.1 bits (170), Expect = 2e-11
 Identities = 30/38 (78%), Positives = 35/38 (92%), Gaps = 0/38 (0%)
 Frame = +2
 Query 17 TLIQNEKNELIPTRTITGWRMCIDYMLNLTATRKDHSP 130
 ++I NEKNELIPTRT+TGWR+CID+ RLNTATRKDH P
 Sbjct 36 SVILNEKNELIPTRVTGWRVCIDHRLNLTATRKDHFP 73

AD-P216MCTC-27B (397 letters)

	Score	E
Sequences producing significant alignments:	(Bits)	Value
gi 33341541 gb AAQ15191.1 resistance protein [Vitis vinifera]	112	6e-24
gi 74231152 gb ABA00702.1 disease resistance protein [Phaseolus	104	9e-22

Alignments

>gi|33341541|gb|AAQ15191.1| resistance protein [Vitis vinifera] Length=292
 Score = 112 bits (279), Expect = 6e-24
 Identities = 61/136 (44%), Positives = 86/136 (63%), Gaps = 7/136 (5%)
 Frame = +1
 Query 1 VLDDVSTSDQITNLVGGHNWLGPGRSRIIVTTRDKH--AICKADDIYKVKALESSDARQLL 174
 VLDDV Q+ L+G WLG GSR+I+TTR+KH A+ + DD+ Y+ V+ L +A + L
 Sbjct 53 VLDDVDDLSQLLEYLLGHREWLGEGRSRIIVTTRDKHVLAVQEVDLLYEVEGLNFEEACELF 112

 Query 175 SLHAFDGTGDGHCLLHWHNLVSKMVEYAKGIPLGLKVLGSLYKCFEWEQCQFQKLQKM 354
 SL+AF + + + NL ++ V Y+ G+PL LKVLGS L+ K + WE + KL K
 Sbjct 113 SLYAFKQ---NLPKSDYRNLAYRVVGYCQGLPLALKVLGSLFNKRIPQWESELNKLDKE 169

 Query 355 PFPEIHNVN--SYSGL 396
 P + IHNV+ SY GL
 Sbjct 170 PEMKIHNVLKRSYDGL 185

A.C.V.F, José · S.C, Bertioli · P.G, Messenberg · D.J, Bertioli .

Desenvolvimento e Mapeamento de Marcadores Moleculares RGA em Espécies Silvestres de *Arachis*

A.C.V.F, José
Universidade Católica de Brasília
Campus II 916 Norte
e-mail: acvfj@hotmail.com
Tel: + 55 – 61 – 34487213

S.C, Bertioli
EMBRAPA Recursos Genéticos
SAINT Parque Rural W3 norte
e-mail: soraya@cenargen.embrapa.br
Tel: +55-61-34484703

P. G, Messenberg
EMBRAPA Recursos Genéticos
SAINT Parque Rural W3 norte
e-mail: messenbe@cenargen.embrapa.br
Tel: +55-61-34484787

D. J, Bertioli (✉)
Universidade Católica de Brasília
Campus II 916 Norte
e-mail: david@cenargen.embrapa.br
Tel: + 55 – 61 – 34487213

Resumo

O amendoim cultivado (*Arachis hypogaea*), pertencente a família das leguminosae é extremamente susceptível a doenças causadas por fungos e nematóides, não mostrando defesas naturais a esse tipo de patógeno dado a sua baixa variabilidade genética. Uma alternativa para descobrir novos genes de resistência é a construção de mapas genéticos visando a utilização dos marcadores em programas de seleção assistida. Para isso, o uso de espécies silvestres diplóides de *Arachis* é uma ferramenta de grande auxílio, pois as mesmas são fontes naturais de genes de resistência a patógenos. Através da técnica RGA-Display, *primers* degenerados e específicos baseados em regiões conservadas do domínio NBS (Nucleotide Binding Site – que é o maior grupo de genes de resistência em plantas) foram utilizados para isolar um total de 139 novos marcadores moleculares. Destes, 19 marcadores são análogos a genes de resistência (RGAs) uma população de mapeamento obtida através do cruzamento de duas espécies diplóides silvestres de *Arachis* (*A. duranensis* e *A. stenosperma*). Esses marcadores foram separados em dois grupos baseados na estrutura da porção N-terminal do domínio: TIR e NãoTIR. Esses novos RGAs foram seqüenciados e alinhados juntamente com 67 RGAs previamente isolados para o gênero e a partir deste alinhamento uma árvore filogenética foi construída. Ao mesmo tempo, esses RGAs previamente isolados foram usados na técnica RFLP para a identificação de nove novos marcadores para a população de mapeamento. Um total de 148 novos marcadores foram isolados através das duas técnicas e 28 deles foram identificados como sendo RGAs. 39 marcadores, incluindo 15 RGAs, puderam ser mapeados em uma mapa genético recém construído para o gênero com um LOD mínimo de 5,0 e fração de recombinação de 0,35 em oito dos dez grupos de ligação do mapa.

Palavras-chave: Amendoim; RFLP; RGA-Display; espécies silvestres; mapa genético.

Abstract

Peanut seeds (*Arachis hypogaea*), which belong to leguminosae family are extremely susceptible to diseases caused by fungi and nematodes, not showing typical defenses. An alternative to discovery novel resistant genes is the construction of genetic map of the wild type that is nature resources of pathogenic resistance. Degenerated and specific *primers* base on conserved regions of the nucleotide binding site (NBS) domain (encoded by the largest group of plant dieses resistance genes) were used in RGA-Display to isolate a set of 139 new markers, including 19 resistance gene analogs (RGAs) from a mapping population derived fro a cross between two wild diploid species of *Arachis* (*A. duranensis* and *A. stenosperma*). These markers grouped into two classes of N-terminus structure TIR or NãoTIR group. These RGAs markers were sequenced and aligned with 67 previously isolated RGAs from *Arachis* and a phylogenetic tree was constructed. At the same time, these previously isolated RGAs from *Arachis* were used in RFLP technique to identify nine new markers for this population. A total of 148 new markers were identified by the two techniques and 28 were identified as RGAs. 39 markers were mapped in a recently constructed linkage map for the genus and 15 of these were RGAs. These markers were mapped with a minimum LOD of 5,0 and 0,35 of recombination fraction in eight of ten LG of the map.

Keywords: Peanut; RFLP, RGA-Display, Genetic Map; Wild Species.

Introdução

O gênero *Arachis*, pertencente à família Fabaceae, é nativo da América do sul e abrange nove sessões (Valls e Simpson, 1994). Na sessão *Arachis* existem 25 espécies diplóides e 2 tetraploides representando três genomas, A, B e D (Krapovickas e Gregory, 1994). Dentro desta sessão encontra-se o *Arachis hypogaea*, ou amendoim cultivado que é uma leguminosa de grande importância na alimentação humana em países da África e Ásia (Krapovickas e Gregory, 1994). O amendoim cultivado apresenta um genoma tetraplóide muito extenso ($n=1.74 \times 10^9$ bp, Bennet & Smith, 1976) e apesar de ser morfológicamente variável, apresenta baixa diversidade

genética, além de não possuir fontes de resistência a diversas pragas como *Cercospora arachidicola*, *Cercosporidium personatum*, e *Meloidogyne* spp. (Kochert et al., 1996).

O amendoim cultivado originou-se a partir de um único evento de alotetraploidização de um híbrido estéril de duas espécies silvestres. A planta resultante, contendo dois genomas distintos, A e B, é reprodutivamente isolada dos seus parentais silvestres *Arachis duranensis* e *Arachis ipaensis* (Seijo et al., 2004). Provavelmente por causa desta origem, *A. hypogaea* tem uma base genética extremamente estreita (Kochert et al., 1991). Essa baixa variabilidade genética no *Arachis hypogaea* pode causar um grande problema com a geração de marcadores moleculares polimórficos. Marcadores como RAPDs (Random Amplified Polymorphic DNAs), AFLPs (Amplified Fragment Length Polymorphism), RFLPs (Restriction Fragment Length Polymorphisms) e SSRs (simple sequence repeats) têm apresentado baixo polimorfismo dentro do germoplasma de *A. hypogaea* e isto torna muito difícil a construção de mapas usando populações derivadas de cruzamentos de *A. hypogaea* x *A. hypogaea* (Kochert et al., 1996, Halward et al., 1991 e Herselman et al., 2003). Em contraste, o amendoim silvestre apresenta alta diversidade genética e é uma grande fonte de resistência a fatores bióticos (pragas e patógenos) e abióticos (stress hídrico, toxidez do solo) (Galgaro et al., 1997). Essa alta diversidade está sendo usada para facilitar a construção de mapas. A Embrapa - Recursos Genéticos e Biotecnologia possui o maior banco de germoplasma de amendoim silvestre do mundo, o qual pode ser explorado visando o melhoramento do amendoim (Simpson et al., 2001).

A transferência de resistências de espécies selvagens para o amendoim cultivado tem sido extremamente dificultada pela diferença de ploidia, e conseqüente barreira de fertilidade. Tradicionalmente, a transferência destes genes tem sido feita através de cruzamentos complexos, visando a introgressão destas resistências seguida de vários retrocruzamentos com *A. hypogaea* (Simpson et al., 2001). Recentemente, uma maior compreensão sobre a origem de *A. hypogaea* tem permitido a re-síntese da espécie *A. hypogaea* a partir de espécies selvagens (Fávero, 2004). Isso, juntamente com um melhor entendimento da estrutura de genes de resistência tem o potencial de acelerar a transferência dos mesmos para cultivares de interesse através da seleção assistida por marcadores moleculares (Hash et al., 2003). Para o uso dessa seleção assistida, é necessário identificar marcadores proximamente ligados a genes de resistência a doenças.

Muitos dos genes de resistência (R-genes) são caracterizados pela presença do domínio NBS-LRR (Nucleotide Binding Site - Leucine Rich Repeat) (Meyers et al. 1999), que em plantas, a única função associada é em resistência a doenças, que é frequentemente manifestado como uma resposta de hipersensibilidade envolvendo morte celular programada (Tameling et al. 2002). A presença de regiões conservadas dentro deste domínio é uma ferramenta muito importante para a geração de marcadores moleculares, pois propiciam o desenvolvimento de *primers* que amplificam regiões análogas aos genes de resistência (RGAs) (Leister et al., 1996, Yu et al., 1996). Essas regiões são denominadas motivos e são caracterizadas especialmente pelo P-loop, Kinase-2, Kinase-3a e GLPLA.

Em algumas espécies, observou-se que RGAs estão fortemente ligados a genes de resistência, ou parte dos mesmos (Collins et al., 1998). Em *Arachis*, um total de 79 RGAs já foram isolados do amendoim cultivado e de espécies silvestres diplóides, o que representa o maior banco de RGAs de um genoma de leguminosa não seqüenciado (Bertioli et al., 2003). A inclusão destas seqüências em um mapa genético de *Arachis* sp. recém construído, possibilitaria a identificação de genes de resistência ligados a estes RGAs, auxiliando os programas de melhoramento da cultura.

Para o desenvolvimento de marcadores moleculares baseados em RGAs algumas estratégias podem ser utilizadas. O RFLP (Yu et al. 1996) utiliza clones de RGAs como sonda para hibridizações em Southern-blot. Uma segunda técnica, RGA-Display, combina as características do AFLP (Vos et al., 1995) e S-SAP (Sequence Specific amplified polymorphism - Waugh et al., 1997). Nessa abordagem, DNA genômico é cortado com enzimas de restrição e ligado a um adaptador nas extremidades. A amplificação por PCR é feita com um *primer* relativo ao adaptador e outro à região codificante de motivos no domínio NBS em RGAs. O tamanho do produto amplificado varia com a distância do sítio de restrição até o motivo, causando polimorfismo entre indivíduos. A similaridade com RGAs já descritos pode ser detectada por seqüenciamento das bandas polimórficas. A segunda técnica apresenta como vantagens à praticidade e rapidez inerentes ao uso de marcadores do tipo PCR.

Em *Arachis* spp., poucos mapas genéticos foram publicados (Halward et al., 1993; Garcia et al., 1995; Burrow et al., 2001; Herselman et al., 2004). Em 2005, Moretzhon e colaboradores publicaram um mapa baseado em Microsatélite (SSR) para uma população F₂ de 93 indivíduos obtidos através do cruzamento de duas espécies silvestres diplóides de genoma AA (*A. stenosperma* e *A. duranensis*). Um total de 271 novos marcadores Microsatélite foram mapeados em 11 grupos de ligação cobrindo uma distância de 1.230,89 cM no mapa e com a distancia média de 7.24 cM entre os marcadores.

Com o auxílio da Bioinformática com ferramentas de processamento e análise de seqüências, a identificação de marcadores RGAs se torna possível, facilitando a inclusão dos mesmos em mapas genéticos. Juntamente com programas de seleção assistida por marcadores (MAS), esses RGAs vão permitir uma estratégia mais rápida para a transferência de genes de resistência a diversas pragas e doenças do gênero e outras leguminosas (Leal-Bertioli et al., 2004; Guimarães et al., 2005)

Material e Métodos

Material vegetal e população de mapeamento

A população de mapeamento foi obtida através do cruzamento de duas espécies silvestres: *Arachis duranensis* (acesso: K7988) e *A. stenosperma* (acesso: V10309). As duas espécies são diplóides e possuem genoma A, sendo a segunda resistente aos nematóides das galhas *Meloidogyne arenaria* raça 2 e *M. javanica* (Leal-Bertioli et al., 2000) e aos fungos de mancha foliar *Cercospora arachidicola* e *Cercosporidium personatum* (Proite et al., 2002). A planta (F₁) híbrida gerada foi auto-fecundada produzindo a população de mapeamento usada nesse trabalho (população F₂). Essa população consiste de 93 plantas mantidas em vasos de concreto, dentro de casas de vegetação na EMBRAPA Recursos Genéticos e Biotecnologia (Brasília-DF).

Para os experimentos realizados neste trabalho, DNA total de *A. duranensis*, *A. stenosperma*, da planta híbrida e de 93 plantas da população F₂ foi extraído de folhas jovens seguindo um protocolo de extração de DNA baseado em CTAB publicado por Doyle et al. (1991).

RFLP

O Southern-blot foi feito de acordo com Southern, (1975), onde DNA genômico (10 µg) de *A. duranensis*, *A. stenosperma*, da planta híbrida e da população F₂ foi clivado com *EcoRI* ou *HindIII*, separados em gel de agarose 0,8%, transferido para uma membrana de Nylon (N+ Hybond), fixado em luz ultravioleta em um forno “Cross-link” e hibridizado radioativamente com RGAs isolados por Bertoli et al., 2003 (Tabela 1). As sondas foram amplificadas por PCR e marcadas com [α -³²P]dATP com o Kit “Random Primed DNA labeling” (Roche, Mannheim, Germany). As membranas hibridizadas foram auto-radiografadas em filme para Raio-X (Kodak-Bio). Foram escolhidos 9 RGAs representativos para a hibridização com os parentais.

RGA-Display

Cerca de 100 ng de DNA genômico de *A. duranensis*, *A. stenosperma*, da planta híbrida e da população F₂ foram clivados simultaneamente com *PstI* e *MseI* por 4 h a 37°C e adaptadores foram ligados às extremidades dos fragmentos como descrito por Vos et al., 1995 e Maughan et al., 1996. Em uma primeira rodada de PCR, *primers* complementares aos adaptadores (Tabela 3) ligados às extremidades (sem bases seletivas) foram utilizados para a amplificação dos fragmentos digerido e ligados. Em uma reação de 50 µl foram utilizados: 0,9 mM de MgCl₂, 0,6 µM de cada *primer*, 0,25 mM dNTPs, 0,5 U de *Taq* polymerase (Platinum Invitrogen) e 150 ng de cada DNA em um programa de trinta ciclos de: 30s a 95° C, 1 min. a 52° C, 1 min. a 72° C. Antes dos ciclos foi adicionado um passo de 95°C por 15 min para a ativação da enzima. Uma alíquota equivalente a 25 ng do produto dessa primeira amplificação foi usada em uma segunda rodada de PCR. O *Primer MseI* (complementar ao adaptador para *MseI*) com uma, duas ou três bases seletivas foi usado em um sentido enquanto um *primers* desenhado para o motivo Kinase-2 no domínio NBS foi usado no outro sentido. Essa segunda PCR foi conduzida usando o mesmo programa descrito anteriormente em uma reação de 15 µl contendo 10% de tampão para PCR sem magnésio; 0,6 µM de cada *primer*; 0,9 mM de MgCl₂; 0,25 mM de dNTPs e 1 unidade de Platinum *Taq* DNA polimerase (Invitrogen).

Os produtos obtidos na segunda rodada de PCR foram separados em gel desnaturante de poliácridamida 4% por 1,5 h a 60W. O gel foi corado com Nitrato de Prata (Creste et al., 2001) e as bandas polimórficas foram retiradas do gel e eluídas segundo Upender et al. (1995). O produto foi re-amplificado usando-se as mesmas condições já descritas.

Seqüenciamento e processamento das amostras

As reações de seqüenciamento foram otimizadas utilizando um dos dois *primers* usado para a amplificação do fragmento polimórfico. Foi usado o Kit de seqüenciamento BigDie Terminator (Applied Biosystems, CA, USA) em um seqüenciador automático ABI377 (Applied Biosystems). As amostras foram seqüenciadas nos dois sentidos.

As seqüências obtidas foram processadas e analisadas utilizando-se a ferramenta Staden Package (Staden et al., 2003), onde os dois sentidos de cada seqüência puderam ser montados para se conseguir uma melhor qualidade de leitura. A ferramenta “Pregap4” foi utilizada para o pré-processamento das amostras como controle da qualidade das seqüências, conversão de formatos de e estimativa da qualidade de cada base (Phred \geq 20). Após esse pré-processamento, a ferramenta “Gap4” foi utilizada para o processamento propriamente dito, onde os pares de seqüência foram juntados quando conveniente, onde também se juntou as seqüências provenientes das repetições do seqüenciamento.

As seqüências de DNA processadas foram submetidas aos programas BLASTX e BLASTN (Altschul et al., 1997), contra dados de todas as prováveis proteínas de *Arabidopsis* e ao banco de dados total do *GenBank*. Os “Hits” foram considerados significativos quando tiveram um “E-value” menor que 1×10^{-1} . Para uma melhor análise, as amostras processadas também foram traduzidas para sua seqüência correspondente protéica, com o auxílio da ferramenta de tradução para aminoácidos “Translate – TrEMBL” no site www.expasy.ch/tools/dna.html. As seqüências traduzidas foram então submetidas à ferramenta “Motif-Scan” (Falquet et al., 2002 - www.expasy.ch/prosite) onde regiões conservadas do genoma são rastreadas e comparadas com seqüências já conhecidas. Essa ferramenta utiliza o banco de dados do “Pfam” para a busca dos motivos e domínios.

Alinhamento das Seqüências e Árvore de Similaridade

Para o alinhamento das seqüências o programa de domínio público JalView 2.07 (www.jalview.org) foi utilizado, onde as seqüências protéicas foram alinhadas de acordo com sua similaridade. 68 RGAs isolados por Bertoli et al. em 2003 para *Arachis* também foram utilizados para os alinhamentos realizados nesse trabalho. O alinhamento foi do tipo “ClustalW Multiple Sequence Alignment Realign” e as regiões conservadas entre as amostras identificadas. A árvore de similaridade foi construída por distância média por porcentagem de similaridade.

Genotipagem da população F₂ e Mapeamento

Cada combinação de *primer* que produziu pelo menos uma banda polimórfica na técnica RGA-Display foi utilizada para genotipar toda a população F₂ (93 indivíduos) juntamente com seus parentais e a planta híbrida. Através do programa Excel (Microsoft) foram geradas tabelas contendo informações relativas à presença ou ausência da banda polimórfica obtida para cada indivíduo da população F₂.

Os marcadores desenvolvidos por RGA-Display e os desenvolvidos por RFLP foram separados de acordo com o parental de origem. Esses marcadores polimórficos foram analisados na população de mapeamento (população F₂) consistindo de 93 plantas. Um teste χ^2 foi feito para testar a hipótese nula de segregação de 1:2:1 (no caso dos marcadores co-dominantes) e de 3:1 (no caso dos marcadores dominantes) em todos os marcadores genotipados. Apenas os marcadores que não mostraram distorção ($P > 0,05$) da segregação esperada foram usados na construção inicial do mapa. As análises de ligação foram feitas usando o programa Mapmaker Macintosh versão 2.0 (Lander et al., 1987). Um LOD mínimo utilizado foi de 5,0 e o máximo de

fração de recombinação (h) de 0,35 foi escolhido para os grupos de ligação (LG). O máximo de ligação entre os marcadores dentro de cada grupo de ligação foi estimado pelo método de correlação de matrizes pela análise de três pontos para grupos contendo mais de sete marcadores. A ordem dos marcadores foi confirmada por comparação de probabilidade da possível ordem usando a análise de múltiplos pontos e permutando todas as ordens das trincas adjacentes. *Loci* que puderam ser ordenados com a probabilidade de proporção $\geq 1,00:1$ foram usados para estabelecer a estrutura do mapa. *Loci* de marcadores co-dominantes que não mostraram distorção na segregação em $P > 0,01$ foram usados nesta análise. Em um segundo passo, o LOD diminuído para que novos marcadores fossem incluídos em cada grupo. O comando “try” foi, então, usado para determinar a exata posição dos novos marcadores em cada grupo. Em um passo seguinte, marcadores distorcidos e marcadores segregando 3:1 (marcadores RGA-Display) foram incluídos. A ordem dos novos marcadores foi novamente confirmada e as frações de recombinação foram convertidas em distância de mapa em centimorgans (cM) usando a função “Kosambi’s mapping”. Os tamanhos de genoma no mapa foram estimados de acordo com Hulbert et al. (1988) levando em consideração apenas a comparação dos pares de marcadores colocados na estrutura do mapa.

Resultados

RFLP

Foram feitas, no total, dezessete membranas contendo somente os parentais e dezoito membranas contendo a população F_2 para mapeamento. Primeiramente, cada membrana contendo os parentais (digeridos com uma das duas enzimas) foi hibridizada com uma das nove sondas diferentes para a seleção de quais seriam polimórficas. Das sondas testadas, todas se mostraram polimórficas para os parentais, sendo elas: S5-A-375, S5-A-384, S1-A-36, S1-A-37, C8-V-434, S4-A-164, S1S2-A-152, T-A-44 e S5C8-AXY-370. A Tabela 2 mostra um resumo dos resultados dos testes aplicados nos parentais. A enzima *EcoRI* produziu 12 fragmentos polimórficos, quatro na planta *A. duranensis* e oito na planta *A. stenosperma*, enquanto a enzima *HindIII* produziu 24 fragmentos polimórficos (sete com a planta *A. duranensis* e 17 com a planta *A. stenosperma*) representando o dobro de marcadores. O RGA que produziu mais polimorfismo foi o S1-A-37, onde 12 marcadores poderão ser genotipados na população F_2 entre as duas enzimas. O parental *A. stenosperma* produziu 25 bandas polimórficas no total, enquanto o *A. duranensis* produziu apenas 11 bandas polimórficas.

Nos testes preliminares realizados nos parentais também foi possível a contagem do número de *loci* de cada RGA no genoma de cada planta. A Tabela 2 também representa o número de cópias dos RGAs utilizados como sonda nos parentais (*A. duranensis* e *A. stenosperma*). O RGA S5-A-375 possui 13 cópias no genoma da *A. duranensis* digerido com a enzima *EcoRI*, enquanto que no genoma de *A. stenosperma*, clivado com *EcoRI*, só existe 1 cópia do RGA S5-A-384.

Após todos os testes realizados e de posse das informações obtidas sobre as sondas e enzimas com maior potencial de polimorfismo, três sondas foram genotipadas na população completa (S1-A-36, S1-A-37 e S1S2-A-152) e duas sondas foram genotipadas em 20 plantas da população F_2 (S4-A-164 e T-A-44). A Figura 1

mostra um Southern-blot com vinte plantas da população F_2 digeridas com *EcoRI* e hibridizadas com a sonda S1-A-36. As bandas observadas segregam independentemente e neste caso, foram genotipadas como sendo dominantes, ou seja, cada banda representa um marcador. Após a genotipagem da população F_2 , nove bandas polimórficas puderam ser contabilizadas para as cinco sondas testadas nessa população (S1-A-36: 3 bandas, S1-A-37: 1 banda, S1S2-A-152: 2 bandas, S4-A-164: 2 bandas e T-A-44: nenhuma banda genotipada).

RGA-Display

Nas reações de amplificação dos fragmentos, de 140 combinações possíveis de *primers* (sete *primers* relativos ao motivo Kinase-2 e vinte relativos ao adaptador ligado ao sítio de *MseI*) 130 foram testadas contra os parentais *A. duranensis* e *A. stenosperma*. Dessas combinações, 100 apresentaram algum grau de polimorfismo entre os parentais e as 24 melhores combinações foram selecionadas para serem genotipadas na população F_2 de 93 indivíduos (Tabela 3).

As amplificações foram testadas com três tipos de *Taq* DNA polimerase: Hot Star *Taq* (Quiagen), *Taq* comum (PhT) e a Platinum *Taq* (Invitrogen). As enzimas Hot Star *Taq* e Platinum mostraram alta fidelidade nos resultados, apresentando o mesmo perfil de bandas. Já a enzima comum, amplificou uma grande quantidade de fragmentos inespecíficos (dados não mostrados). Esses fragmentos foram seqüenciados e não apresentaram homologia com nenhuma seqüência já depositada no GenBank, por estes motivos, todas as reações posteriores foram feitas com *Taq* Hot Star ou *Taq* Platinum.

Os *primers* desenhados para reconhecer o sítio de restrição da enzima *MseI* possuem bases seletivas ao final de sua seqüência afim de restringir o número de fragmentos produzidos pela PCR. Uma, duas ou três bases foram adicionadas ao final destes *primers*. Testes com esses *primers* contendo essas bases seletivas foram feitos nos parentais e quase todas as combinações produziram polimorfismo entre as espécies estudadas. A Figura 2 contém um gel de poliácridamida 4% corado com Nitrato de Prata com amostras de *A. duranensis* e *A. stenosperma* amplificadas com *primers* P2-378 e *Mse* com uma, duas ou três bases seletivas. Pode-se observar que as reações contendo o *primer Mse* com apenas uma base seletiva possuem mais bandas (43) do que as amplificadas com duas bases seletivas (26 bandas amplificadas), que por sua vez, possuem mais fragmentos que as amplificadas com três bases seletivas (12 bandas amplificadas). A melhor reação para ser analisada, baseando-se nesse teste, é a amplificada com duas bases seletivas, pois gerou um número satisfatório de fragmentos sem formar um rastro como acontece em alguns casos (dados não apresentados).

Após os testes iniciais de temperatura de ligação, *Taq* polimerase e número de bases seletivas, foram feitos testes (*Screenings*) para a detecção de melhores combinações de *primers* para serem utilizados na população F_2 de mapeamento. Foi observado que os *primers* RGA específicos produzem em média, duas vezes mais bandas do que os *primers* degenerados, significando que os *primers* degenerados necessitam de um decréscimo na temperatura de anelamento para que se liguem adequadamente aos fragmentos. Nas 100 combinações de *primer* polimórficas testadas nos parentais, foram amplificados 3.095 fragmentos, sendo 293 (9,5%) polimórficos entre os parentais. Para a genotipagem da população F_2 , foram escolhidas as 24 melhores combinações de *primers* polimórficas. Um exemplo da genotipagem pode ser observado na Figura 3 que representa um gel de poliácridamida corado com Nitrato de Prata contendo RGA-Display na população F_2 . Pode-

se observar que oito bandas polimórficas foram visualizadas neste gel. As amostras amplificadas com o DNA dos parentais mostradas nesta figura, não forneceram exatamente o mesmo resultado do encontrado na população F₂. As bandas principais são as mesmas, mas nos parentais houve uma qualidade menor de amplificação dos fragmentos mais fracos neste caso.

No total, 765 fragmentos (polimórficos e monomórficos) foram amplificados com as 24 combinações de *primers* usadas na população F₂. Destes, 143 foram considerados polimórficos entre os parentais, mas somente 139 marcadores puderam ser analisados devido à falta de consistência de algumas bandas. Dentre as combinações genotipadas, houve uma taxa de polimorfismo de 18,7 %, representando uma média de seis bandas polimórficas por combinação de *primer*.

Dois marcadores RGAs identificado na técnica RGA-Display (AD-NBSMGGC-8B e AD-NBSMGGC-8B) foram escolhidos para a hibridização com DNA genômico dos parentais em Southern-blot para a identificação do número de *loci* dos RGAs. Esses marcadores foram re-amplificados, resolvidos em gel de agarose, purificados com o Kit “QIAquick Gel Extraction” (Quiagen) e usados como sonda contra DNA dos parentais *A. duranensis* e *A. stenosperma*. O RGA AD-NBSMGGC-8B apresentou dez cópias no DNA digerido com *Hind*III e 11 cópias no DNA digerido com *Eco*RI enquanto que o RGA AD-NBSMGGC-8B apresentou 9 cópias tanto no DNA digerido com *Hind*III quanto no digerido com *Eco*RI (dados não apresentados). Nenhum polimorfismo foi observado entre os parentais com esses RGAs hibridizados.

Seqüenciamento e análise das bandas polimórficas

As seqüências obtidas através do seqüenciamento das bandas polimórficas na técnica RGA-Display obtidas foram processadas com o auxílio do programa Staden Package (Staden et al., 2003). No total foram seqüenciadas 109 bandas polimórficas. Após a junção de todas as leituras de seqüência de cada amostra, 100 seqüências completas foram utilizadas para as análises. Primeiro, as seqüência de nucleotídeos foram comparadas com o GenBank utilizando o BLASTX (Altschul et al., 1997). Essas seqüências também foram comparadas com um banco local de genes de *Arabidopsis thaliana*. Considerando que as bandas polimórficas produzidas e usadas para seqüenciamento eram de tamanho pequeno (entre 80 e 400pb), as seqüências produzidas também eram pequenas. Quando foi feita uma comparação destas seqüências com outras já depositadas no GenBank de tamanhos maiores, os valores de *score* e *e-value* se apresentam pequenos em vista dessa diferença de tamanho entre as seqüências. Devido a esse problema, outro tipo de análise também foi feito para que a homologia com os RGAs fosse confirmada. As seqüências foram traduzidas para sua seqüência protéica correspondente utilizando a ferramenta “Translate – TrEMBL” no site www.expasy.ch/tools/dna.html. A escolha da melhor tradução foi baseada na presença da seqüência do *primer*. As traduções foram submetidas ao programa de domínio público “Motif-Scan” (Falquet et al., 2002) no site www.expasy.ch/prosite onde a ferramenta HAMAP procura no banco de dados do Pfam por domínios e motivos conhecidos dentro da seqüência estudada. A Tabela 4 representa uma comparação entre os resultados obtidos no BLASTX contra o banco não redundante do NCBI e o banco do Pfam com a ferramenta “Motif-Scan” para os prováveis marcadores RGAs.

Outro aspecto analisado nas seqüências RGAs foi o tipo de estrutura (TIR ou NãoTIR) encontrado na porção N-terminal das amostras. Por meio de comparação com as seqüências completas já depositadas no GenBank, pôde-se prever qual estrutura encontra-se em cada amostra. Os RGAs 9A, 9B, 9C, 12F, 26A, 26B, 27A, 27B e 28G possuem a estrutura do tipo TIR. Os RGAs 8D, 8J, 14A 14D e 24E não possuem essa estrutura, sendo então denominadas NãoTIR. As outras amostras RGAs não puderam ter o seu grupo definido por comparação com seqüências conhecidas.

Para que os RGAs encontrados nesse trabalho pudessem ser comparados entre si, foi feito um alinhamento das seqüências protéicas através dos programas Jalview do tipo ClustalW (www.ebi.ac.uk). Em uma primeira análise, apenas os RGAs encontrados foram alinhados, mostrando claramente as regiões conservadas na seqüência e a região de anelamento do *primer* (Figura 4). As amostras 9B, 9C, 14A e 28G não possuem a região de anelamento do *primer* nesse alinhamento, o que significa que a seqüência não está completa, foi usada apenas a leitura relativa ao *primer* reverso (complementar ao sítio de *MseI*), onde o mesmo não conseguiu alcançar a região Kinase-2 no começo da seqüência. Apesar dos baixos valores de significância para algumas amostras, no alinhamento é possível observar a presença das regiões conservadas que caracterizam o domínio NBS.

Em um segundo momento, as seqüências RGAs traduzidas foram comparadas com seqüências já descritas no Genbank. Para a comparação foram utilizados 67 RGAs descritos por Bertoli et al., 2003 (Figura 5). O motivo Kinase-2 é extremamente conservado em quase todas as seqüências, só não aparecendo nas seqüências incompletas, como no caso das amostras 9B, 9C, 14A e 28G. O motivo Kinase-3a também está extremamente conservado em todas as amostras, menos na 14A. Este motivo está representado pelos aminoácidos.

Após o alinhamento múltiplo das seqüências foi feita uma árvore de similaridade com as dezenove seqüências RGAs encontradas neste trabalho. A ferramenta utilizada para a confecção da árvore foi a “Distância Média por Porcentagem de Identidade” do programa Jalview (Figura 6). Uma segunda árvore de similaridade foi construída para o alinhamento contendo também as seqüências RGAs obtidas por Bertoli et al., 2003. Nessa árvore, 67 seqüências RGAs depositadas no GenBank para *Arachis* foram utilizadas para uma comparação com as 19 seqüências RGAs encontradas aqui, onde através de distância média por porcentagem de identidade, as amostras foram separadas em grupos de similaridade de seqüência (Figura 7)

Algumas amostras seqüenciadas mostraram homologia com proteínas do tipo Kinase-2 quando submetidas ao programa BLASTX e comparadas com o banco de dados não redundante. Essas amostras (AS-NBSMCTT-24H, AD-P2378MC-20B, AD-P2391MCTA-13D, AS-P2391MCTA-13C, AD-P2378MCTA-23A) foram, então, alinhadas juntamente com outras quatro seqüências com homologia para essa mesma proteína já depositadas no GenBank (gi42567050 kinase - *Arabidopsis thaliana*; gi4826399 wall-associated kinase 2 - *Arabidopsis thaliana*; gi15218170 WAK5 - Wall Associated Kinase e gi15218959 WAK2; ATP binding / calcium ion binding) (Figura 8). Outras amostras mostraram homologia com transposons, retrotransposons e proteínas como, por exemplo, B-amilase. A Tabela 5 é um resumo dos principais *Hits* encontrados para os marcadores que não tiveram homologia com RGAs.

Genotipagem da população F_2 e Mapeamento

Os marcadores desenvolvidos no presente trabalho, foram inseridos em um mapa de ligação já existente, baseado em marcadores Microsatélite (Moretsohn et al., 2005). Neste mesmo mapa, foram incluídos

marcadores AFLP desenvolvidos em outro laboratório (Sainsbury Institute, na Inglaterra) e marcadores âncora, desenvolvidos na Universidade de Aarhus, na Dinamarca. No total, foram mapeados 335 marcadores, sendo 180 Microsatélite, 65 marcadores âncoras, 46 AFLPs, 3 RAPDs, 39 RGA-SAP (sendo 15 RGAs), usando-se um LOD mínimo de 5,0 e uma fração de recombinação máxima (θ) de 0,35 (Figura 9).

Como esperado, dez grupos de ligação foram obtidos. O número de marcadores por grupo de ligação variou de 4 (Grupo 10) a 51 (Grupo 1). A inserção de novos marcadores, em sua maioria dominantes, aumentou sobremaneira o tamanho dos grupos e do mapa total, em cM. Isso pode ter sido causado por alguns erros de genotipagem ou, mais provavelmente, por dificuldades do programa de mapeamento (Mapmaker) em trabalhar com marcadores co-dominantes e dominantes em conjunto. Uma evidência desse fato é que, dos 206 marcadores AFLP polimórficos incluídos na análise, apenas 46 mapearam e em somente cinco dos 10 grupos de ligação (Grupos 1, 4, 6, 8 e 9). Com isso, os novos marcadores devem ter sido inseridos na ordem correta, mas os valores das distâncias entre marcadores foram, possivelmente, superestimados. Devido a esse problema a numeração dos grupos de ligação não obedeceu à ordem de tamanho dos mesmos.

Dos nove marcadores obtidos através da técnica RFLP, quatro puderam ser mapeados. Embora esses marcadores possam ser bem mais informativos do que os RGA-Display se foram mapeados como co-dominantes, não apresentaram esse tipo de característica nesse caso. Devido a uma falta de resolução apropriada dos fragmentos hibridizados, os marcadores foram analisados como dominantes, embora tenham mapeado muito próximos de seus pares. Os marcadores S1-A-36B e S1-A-36D mapearam a uma distância de 5,3 cM no Grupo de Ligação 4, enquanto que os marcadores S1-A-37A e S1-A-36C mapearam a uma distância de 14,3 cM no Grupo de Ligação 3.

Dos 658 marcadores polimórficos incluídos na análise, 317 (48,2%) desviaram das proporções 1:2:1 ou 3:1 esperadas, ao nível de 5%. Essa proporção de marcadores distorcidos é maior do que a detectada em dois mapas de *Arachis*, baseados em marcadores RFLP, que foi de 25% em média (Halward et al., 1993; Burrow et al., 2001). Com exceção do Grupo 8, os demais nove grupos de ligação apresentaram marcadores distorcidos. Distorções de segregação têm sido descritas para a maioria dos mapas genéticos construídos em plantas, tanto intra como interespecíficos e isso tem sido atribuído a diferentes fatores, tais como: alelos recessivos deletérios (Berry et al., 1995), alelos auto-incompatíveis (Barzen et al., 1995), rearranjos estruturais (Quillet et al., 1995) ou diferenças no conteúdo de DNA (Jenczewski et al., 1997).

Discussão

Geração de marcadores através de RFLP

O mapeamento genético baseado em marcadores RFLP pode ser muito informativo (Halward et al., 1993, Garcia et al., 1995, Burrow et al., 2001) embora muito dispendioso e demorado, além de não conter um número adequado de marcadores. Para a obtenção das membranas a serem hibridizadas, muito DNA precisa ser concentrado em um volume muito pequeno, que, no caso de espécies de *Arachis* se torna um problema, pois o DNA dessas plantas é altamente contaminado por compostos que o deixa viscoso em altas concentrações e de

fácil degradação. Consequentemente, esta não é uma ferramenta ideal para um trabalho de análise e implementação de seleção assistida por marcadores (MAS) em programas de melhoramento do amendoim (Moretzhon et al., 2005).

Os marcadores obtidos com a técnica são gênicos, embora o número (quatro marcadores) é muito pequeno quando colocado em um mapa genético. Técnicas baseadas em PCR fornecem rapidamente um número muito mais significativo de marcadores.

A técnica RGA-Display

Hayes e Saghai Maroof (2000) já descreveram a utilização de um AFLP direcionado para RGAs, encontrando marcadores ligados a um gene de resistência em soja. A modificação proposta por Linden et al. (2004) e utilizada no presente trabalho, foi a utilização de *primers* para diversos motivos do domínio NBS, com diversos graus de degeneração e o seqüenciamento dos fragmentos para a confirmação da sua homologia com genes de resistência.

Essa técnica, que tem as vantagens do AFLP, mostrou ser uma ferramenta muito útil na construção de marcadores moleculares em programas de melhoramento, pois utiliza partes de genes de resistência conservadas para a amplificação de produtos, que, provavelmente são regiões ligadas a esses genes (Soriano et al., 2005; Calenge et al., 2005). Esse é o primeiro estudo deste tipo para o gênero *Arachis*, onde RGAs e outros genes foram mapeados com o auxílio dessa técnica. Anteriormente, apenas marcadores como RFLP, AFLP, RAPD, SSR haviam sido incluídos em mapas genéticos de alta resolução.

Através da técnica, 139 marcadores foram desenvolvidos e 19 mostraram ser RGAs após o seqüenciamento das bandas polimórficas. Quinze seqüências apresentaram forte homologia com outros tipos de proteínas, como β -amilase, Hidrolase, Kinase e outras. Semelhanças com Retrotransposons também foram encontradas como no caso de treze seqüências. Cinqüenta e três marcadores não mostraram homologia nenhuma com outras seqüências depositadas no GenBank, contudo, não podemos descartar a hipótese que essas seqüências sejam genes ou outras proteínas importantes ainda não estudadas.

Em uma análise do mesmo tipo feita para Damasco (*Prunus armeniaca* L.) 27 marcadores foram identificados, e destes, sete (26%) mostraram homologia para genes de resistência do tipo TIR (Soriano et al., 2005). Com a técnica, 16 marcadores puderam ser mapeados em um mapa previamente construído baseado em marcadores Microsatélite (SSR). Entre os marcadores identificados, seis mostraram homologia para proteínas depositadas no GenBank e quatro com Retrotransposons. Dez seqüências não tiveram nenhuma homologia com seqüências já depositadas. Em outro trabalho (Calenge et al., 2005) foram mapeados 43 marcadores pela técnica RGA-Display em um mapa construído para maçã, onde 23 (53%) eram RGAs. Os marcadores foram mapeados em 10 dos 17 Grupos de Ligação do mapa e muitos estavam organizados em pequenos grupos “*clusters*”. Vinte e cinco marcadores mapearam perto de genes ou QTLs para resistência à duas das maiores pragas da espécie.

As seqüências mostrando homologia com genes de resistência quando comparadas com o GenBank através de BLASTX (Altschul et al., 1997) causaram alguma dúvida se realmente eram ou estavam ligadas a genes de resistência pelo discreto valor de Nível de Significância encontrado (e-value). Após o alinhamento das seqüências proteicas, foi confirmado que as mesmas são RGAs, pois os principais motivos que caracterizam o

domínio NBS em genes de resistência mostraram-se claramente conservados em todas as seqüências previamente ditas como sendo RGAs. Esse argumento também foi utilizado como base de confirmação em outros trabalhos (Meyers et al., 2003; Linden et al., 2004; Calenge et al., 2005). Não se pode deixar de pensar que as seqüências com homologia a RGAs possam ser pseudogenes que apareceram durante a geração de novos genes de resistência (Soriano et al., 2005). De fato, já foi proposto que uma proporção significativa de genes do tipo NBS podem ser pseudogenes (Meyers et al., 1999).

No total, 35 marcadores RGA-Display foram mapeados em oito de dez Grupos de Ligação do mapa, onde os mesmos estavam bem distribuídos dentro dos grupos. Alguns marcadores mostraram uma leve tendência à formação de agrupamentos em alguns grupos de ligação. Segundo Meyers et al., 1999, é comum a formação desses agrupamentos dentro dos Grupos de Ligação. Dezenove marcadores mostraram ser RGAs, contudo, essa proporção não atendeu o esperado de 50 a 90% proposto por Van der Linden em 2004 para batata, tomate e maçã. A proporção de seqüências não identificadas (53%) também é muito diferente da encontrada por esse autor (21%), onde o mesmo diz que essas seqüências não podem ser ignoradas, e que uma grande proporção das mesmas devem sim ser RGAs como também proposto por Meyers et al., 2003.

Após todas as análises, os resultados sugerem que o desenvolvimento e o mapeamento desses marcadores RGA-Display poderão contribuir para a identificação de regiões gênicas contendo prováveis genes de resistência e facilitar o longo processo de melhoramento do gênero *Arachis* e possivelmente de outras leguminosas que são a base da alimentação humana e animal em todo o mundo.

Mapeamento Genético

Dos 148 marcadores desenvolvidos no presente trabalho (139 marcadores obtidos pela técnica RGA-Display e 9 por RFLP), 39 mapearam em oito Grupos de Ligação (Grupos 1, 2, 3, 4, 5, 6, 8 e 9), sendo quinze deles considerados RGAs por homologia no BLASTX. Os demais marcadores incluídos não mapearam. Além disso, dos 39 marcadores RGA-Display mapeados, 20 apresentaram valores de BLASTX significativos. Dentre essas homologias destacam-se RGAs, Proteínas Kinase-2, β -amilases e Retrotransposons. Com isso, abre-se a perspectiva de usar essas informações na seleção assistida, visando a introgressão de genes de interesse para o amendoim cultivado, com a vantagem de se ter marcadores pertencentes a esses genes e não apenas ligados a eles.

Considerações Gerais

Os programas de melhoramento de plantas cultivadas têm sido grandemente auxiliados por Seleção Assistida por Marcadores Moleculares (MAS). A obtenção de mapas genéticos, possuindo marcadores gênicos ou ligados a Genes de Resistência, é um passo muito importante para esse processo. Sendo assim, o desenvolvimento de marcadores do tipo RGA é uma ferramenta promissora para a identificação de *loci* associados a esses genes, uma vez que já foi estabelecido que uma grande proporção de RGAs de várias espécies está ligado ou faz parte de genes de resistência. (Collins et al., 2001, Donald et al., 2002, Leal-Bertioli et al., 2004). Esses genes podem estar agrupados no cromossomo, possibilitando também a piramidização de genes de resistência (Aarts et al., 1998). Além disso, vários trabalhos demonstraram que genes de resistência podem

manter sua atividade quando transferidos para outras espécies (Hammond-Kosack et al., 1998), o que também abre a perspectiva de que, uma vez identificados os genes de resistência em *Arachis* spp., estes possam ser transferidos para outras espécies, especialmente leguminosas como soja, feijão e ervilha.

Com as duas técnicas utilizadas, muitos marcadores gênicos foram incluídos no mapa genético recém construído de alta resolução (Moretzsohn et al., 2005). Para isso, foi utilizada uma população F₂ originada do cruzamento entre acessos (*A. duranensis* – K7988 e *A. stenosperma* – V10309) que contrastam para diversas pragas e patógenos como os nematóides das galhas *Meloidogyne arenaria* raça 1 e raça 2, *M. javanica* raça 4 e *M. hapla* (Fávero, 2004). No momento, bioensaios estão sendo conduzidos para a obtenção de QTLs (*Quantitative Trait Loci*) e resistências ao nematóide *M. arenaria* raça 1 que é uma das maiores pragas do amendoim cultivado nos Estados Unidos com grande potencial de infectar as culturas brasileiras.

Referências Bibliográficas

Aarts MGM, Hekkert BL, Holub EB, Beynon JL, Stiekema WJ, Pereira A. Identification of R-gene homologous DNA fragments genetically linked to disease resistance *loci* in *Arabidopsis thaliana*. *Mol Plant-Microbe Interact*, 11:251-258 (1998).

Altschul SF, Thomas L, Madden A, Schäffer A, Zhang J, Zhang Z, Miller W, Lipman, DJ. Gapped BLAST and PSIBLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402 (1997).

Barzen E, Mechelke W, Ritter E, Schulte-Kappert E, Salamini F. An extended map of the sugar beet genome containing RFLP and RFLP *loci*. *Theor Appl Genet* 90:189–193 (1995).

Berry ST, Leon AJ, Hanfrey CC, Challis P, Burkholz A, Barnes SJ, Rufener GK, Lee M, Caligari PDS. Molecular marker analysis of *Helianthus annuus* L. 2. Construction of an RFLP linkage map for cultivated sunflower. *Theor Appl Genet* 91:195–199 (1995).

Bertioli DJ, Leal-Bertioli SC, Lion MB, Santos VL, Pappas G. Jr, Cannon SB, Guimarães PM. A Large Scale Analysis of Resistance Gene Homologues in *Arachis*. *Molecular Genetics and Genomics* 270:34-45 (2003).

Burow MD, Simpson CE, Starr JL, Paterson AH. Transmission genetics of chromatin from a synthetic amphidiploid to cultivated peanut (*Arachis hypogaea* L.): Broadening the gene pool of a monophyletic polyploid species. *Genetics* 159:823-837 (2001).

Calenge F, Van der Linden CG, Van de Weg E, Schouten HJ, Van Arkel EG, Denace EC, Durel CE. Resistance gene analogues identified through the NBS-profiling method map close to major genes and QTL for disease resistance in apple. *Theor Appl Genet* 110: 660–668 (2005).

Collins NC, Webb CA, Seah S, Ellis JG, Hulbert SH, Pryor A. The isolation and mapping of disease resistance gene analogs in maize. *Mol Plant Microbe Interact* 11:968–978 (1998)

Collins NC, Park R, Spielmeier W, Ellis J, Pryor T. Resistance gene analogs in barley and their relationships to rust resistance genes. *Genome* 44:375–381 (2001).

Creste S, Tulmann A, Figueira A. Detection of single sequence repeats polymorphisms in denaturing polyacrylamide gels by silver staining. *Plant Mol. Biol. Reporter* 19:299-306 (2001).

Donald TM, Pellerone F, Adam-Blondon AF, Bouquet A, Thomas MR, Dry IB Identification of resistance gene analogs linked to a powdery mildew resistance locus in grapevine. *Theor Appl Genet* 104:610–618 (2002).

- Doyle JJ, Doyle JL. Isolation of plant DNA from fresh tissue. *Focus* 12:13-15 (1991).
- Falquet L, Pagni M, Bucher P, Hulo N, Sigrist CJ, Hofmann K, Bairoch A. The PROSITE database, its status in 2002. *Nucleic Acids Res.* 1,30(1):235-8. (2002).
- Fávero AP. Cruzabilidade entre espécies silvestres de *Arachis* visando a introgressão de genes de resistência a doenças no amendoim cultivado. Tese de Doutorado. ESALQ, Piracicaba (2004).
- Ferreira ME, Grattapaglia D. Introdução ao uso de marcadores RAPD e RFLP em análise genética. Brasília: Embrapa Cenargen, 220p (1995).
- Galgaro L, Lopes CR, Gimenes M, Valls JFM, Kochert G. Genetic variation between species of sections *Extranervosae*, *Caulorrhizae*, *Heteranthae*, and *Triseminatae* (genus *Arachis*) estimated by DNA polymorphism. *Genome* 41:445-454 (1997).
- Garcia GM. Evaluating efficiency of germplasm introgression from *Arachis* species to *A. hypogaea* L. Ph.D. Dissertation, N.C. State Univ., Raleigh, 197 pp. (1995).
- Guimarães PM, José ACVFI, Proite K, Bertoli DJ, Leal-Bertoli SCM. Desenvolvimento de marcadores moleculares para análogos a genes de resistência em *Arachis* spp. silvestres. *Fitopatologia Brasileira* 30:663-667 (2005).
- Halward TM, Stalker HT, Kochert G. Development of an RFLP linkage map in diploid peanut species. *Theor Appl Genet* 87:379-384 (1993).
- Halward TM, Stalker HT, LaRue EA, Kochert G. Genetic variation detectable with markers among unadapted germplasm resources of cultivated peanut and related wild species. *Genome*, 34: 1013-1020 (1991).
- Hammond-Kosack KE, Tang S, Harrison K, Jones JD. The tomato Cf-9 disease resistance product avr 9 gene functions in tobacco and potato to confer responsiveness to the fungal avirulence gene. *Plant Cell* 10:1251-1266 (1998).
- Hash CT, Bhasker Raj AG, Lindup S, Sharma A, Beniwal CR, Folkertsma RT, Mahalakshmi V, Zerbini E, Blümmel M. Opportunities for marker-assisted selection (MAS) to improve the feed quality of crop residues in pearl millet and sorghum. *Field Crops Research* 84: 79 – 88 (2003).
- Hayes AJ, Saghai Maroof MA. Targeted resistance gene mapping in soybean using modified AFLPs. *Theor Appl Genet* 100:1279–1283 (2000).
- Herselman L, Thwaites R, Kimmins FM, Courtois B, van der Merwe PJA, Seal SE. Identification and mapping of AFLP markers linked to peanut (*Arachis hypogaea* L.) resistance to the aphid vector of groundnut rosette disease. *Theor Appl Genet* 109:1426–1433 (2004).
- Hulbert SH, Ilott TW, Legg EJ, Lincoln SE, Lander ES, Michelmore RW. Genetic analysis of the fungus, *Bremia lactucae*, using Restriction Fragment Length Polymorphisms. *Genetics* 120:947–958 (1988).
- Jenczewski E, Gherardi M, Bonnin I, Prosperi JM, Olivieri I, Huguet T. Insight on segregation distortions in two intraspecific crosses between annual species of *Medicago* (Leguminosae). *Theor Appl Genet* 94:682–691 (1997).
- Krapovickas A, and Gregory WC. Taxonomia del género *Arachis* (Leguminosae). *Bonplandia*, 8: 1-186 (1994).
- Kochert G, Halward TM, Branch WD and Simpson CE. RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. *Theor. Appl. Genet.*, 81: 565-570 (1991).
- Kochert G, Stalker HT, Gimenes M, Galaro L, Lopes CR and Moore K. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). *Am. J. Bot.*, 83:1282-1291(1996).

- Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Etoh J. Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174–181 (1987).
- Leal-Bertioli SCM, Guimarães PM, Fávero AP, Moretzsohn MC, Proite K & Bertioli DJ. Amendoim selvagem: uma fonte de resistência a pragas. *Biotecnologia Ciência e Desenvolvimento*. Site: www.biotecnologia.com.br (2004).
- Leister D, Ballvora A, Salamini S, Gebhardt C. A PCRbased approach for isolating pathogen resistance genes from potato with potential for wide application in plants. *Nat Genet* 14:421–429 (1996).
- Maughan PJ, Saghai Maroof MA, Buss GR, Huestis GM. Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near-isogenic line analysis. *Theor Appl Genet* 93:392–401 (1996).
- Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral BW, Young ND. Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant J* 20:317–332 (1999).
- Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW. Genome-wide analysis of NBS–LRR-encoding genes in *Arabidopsis*. *Plant Cell* 15:809–834 (2003).
- Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral B, Young ND. Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant J* 20:317–332 (1999).
- Moretzsohn MC, Leoi L, Proite K, Guimarães PM, Leal-Bertioli SCM, Gimenes MA, Martins WS, Valls JFM, Grattapaglia D, Bertioli DL. A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (Fabaceae) *Theor Appl Genet* (2005)
- Proite K, Dias JGO, Guimarães PM, Bertioli DJ, Carneiro RG, Valls JFM, Leal-Bertioli SCM. Avaliação de Resistência a Fitonematoides e Fungos em diversas espécies de *Arachis* Silvestre. VI encontro Talento Estudantil, Embrapa (2000)
- Quillet MC, Madjidian N, Griveau Y, Serieys H, Tersac M, Lorieux M, Bervillé A. Mapping genetic factors controlling pollen viability in an interspecific cross in *Helianthus* sect. *Helianthus*. *Theor Appl Genet* 91:1195–1202 (1995).
- Seijo GJ, Lavia GI, Fernandez A, Krapovickas A, Ducasse E, Moscone DEA. Physical mapping of the 5s and 18s-25s rRNA genes by fish as evidence that *Arachis duranensis* and *A. ipaensis* are the wild diploid progenitors of *A. hypogaea* (leguminosa) *American Journal of Botany* 91(9): 1294-1303 (2004).
- Simpson CE, Krapovickas A and Valls JFM. History of *Arachis* including evidence of *A. hypogaea* L. progenitors. *Peant Science* 28: 78-80 (2001).
- Soriano J M, Vilanova E S, Romeiro C, Llácer G, Badenes M L. Characterization and mapping of NBS-LRR resistance gene analogs in apricot (*Prunus armeniaca* L.) *Theor Appl Genet* 110: 980–989 (2005).
- Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Mol. Biol.* 98:503-517.
- Staden R, Judge DP, Bonfield JK. Managing sequencing projects in the GAP4 environment. In: Krawetz SA, Womble DD. *Introduction to Bioinformatics. A Theoretical and practical approach*. Human Press Inc, Totawa (2003).
- Tameling WIL, Elzinga SDJ, Darmin PS, Vossen JH, Takken FLW, Harling MA, Cornelissen B. The tomato R gene products I-2 and Mi-1 are functional ATP binding proteins with ATPase activity. *Plant Cell* 14:2929–2939 (2002).
- Uppender M, Raj L, Weir M. Rapid method for the elution and analysis of PCR products separated on high-resolution acrylamide gels. *Biotechniques* 18:33–34 (1995).

Valls JFM, Simpson CE. Taxonomy, natural distribution and attributes of *Arachis*. In: Kerridge, PC, Hardy, B, Biology and agronomy of forage *Arachis*. Centro Internacional de Agricultura Tropical. p. 1-18 (1994).

Linden CG Van der, Wouters DC, Mihalka V, Kochieva EZ, Smulders MJM, Vosman B. Efficient targeting of plant disease resistance *loci* using NBS profiling Theor Appl Genet 109:384-393 (2004).

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407-4414 (1995).

Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BBT, Powell W. Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (SSAP). Mol Gen Genet 253:687-694 (1997).

Yu YG, Buss GR, Maroof MA. Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. Proc Natl Acad Sci USA, 93:11751-11756 (1996).

Documentos e Fontes em Meio Eletrônico

Home Page EXPASY (Ferramenta TRASLATE e MotiScan) Disponível em:
<<http://www.expasy.ch/prosite>>

Programa JalView 2.07 Disponível em:
<<http://www.jalview.org>>

Home Page FAO (Food and Agriculture Organization) Disponível em:
<<http://www.Fao.org>>

Home Page EBI (European Bioinformatics Institute) – Ferramenta ClustalW. Disponível em:
<http://www.ebi.ac.uk>

Tabela 3 Combinações de *primers* utilizadas para genotipagem da população F₂, Número de bandas amplificadas, polimórficas, seqüenciadas e RGAs produzidas por cada combinação testada

Combinação ^a	Número de bandas				% de	% de
	Amplificadas	Polimórficas	Seqüenciadas	RGAs ^e	Polimorfismo	RGAs
P2-12 + Mct (9)	13	3	3	3	23,1	100
P2-12 + Mcta (1)	8	1	1	0	12,5	0
P2-16 + Mc (28)	42	9	9	1	21,4	11
P2-16 + Mcaa (26)	17	2	2	2	11,8	100
P2-16 + Mctc (27)	35	7	7	2	20	29
P2-19 + Mcac (6)	45	8	0	0	17,8	0
P2-19 + Mcta (12)	22	7	7	1	31,8	14
P2-19 + Mctc (5)	56	10	0	0	17,9	0
P2-19 + Mgecg (7)	49	5	0	0	10,2	0
P2-378 + Mct (20)	43	9	9	0	20,9	0
P2-378 + Mcta (23)	39	6	4	0	15,4	0
P2-378 + Mctc (11)	25	9	8	0	36	0
P2-378 + Mctt (24)	45	15	15	3	33,3	20
P2-391 + Mcta (13)	39	4	4	0	10,3	0
P2-391 + Mgag (16)	31	2	2	0	6,5	0
P2-414 + Mc (17)	40	8	6	0	20	0
P2-414 + Mcta (15)	28	3	3	0	10,7	0
P2-414 + Mctc (10)	14	4	4	1	28,6	25
P2-414 + Mctt (19)	28	3	3	0	10,7	0
NBS5 + Mc (21)	55	9	9	0	16,4	0
NBS5 + Mctc (14)	19	4	4	2	21,1	50
NBS5 + Mgag (22)	36	3	2	0	8,3	0
NBS5 + Mgct (25)	19	2	2	0	10,5	0
NBS5 + Mgecg (8)	17	10	5	4	58,8	40
Total	765	143	109	19	18,7	13
Média por combinação	32	6	5	-	-	-

^a refere-se aos dois *primers* utilizados para a amplificação dos fragmentos referentes ao teste realizados nos parentais. O número entre parênteses ao lado das combinações de *primers* diz respeito ao número da combinação utilizada. ^b Número de bandas amplificadas por cada combinação de *primers*. Inclui bandas polimórficas e monomórficas. ^c Número de bandas que diferem entre os parentais e indivíduos da população F₂. ^d refere-se ao número de bandas polimórficas genotipadas que foram seqüenciadas e caracterizadas. ^e Número de bandas seqüenciadas que tiveram homologia com RGAs no GenBank. A comparação com o banco de dados foi feita através de BLASTX e BLASTN.

Tabela 4 Comparação entre homologias encontradas por BLASTX e Pfam para seqüências RGA-Display

Primer RGA	Marcador	BLASTX ^a		Motif-Scan ^b	
		Hits	E-Value	Hits	E-Value
P2-12	AD-P212MCT-9A	Disease resistance protein (TIR-NES-LRR...	0.003	no hits	
P2-12	AD-P212MCT-9B	Disease resistance protein (TIR-NES-LRR...	0.001	no hits	
P2-12	AD-P212MCT-9C	Disease resistance protein (TIR-NES-LRR...	0.001	no hits	
P2-16	AS-P216MCAA-26A	Disease resistance protein (TIR-NES-LRR...	4,00E-07	no hits	
P2-16	AD-P216MCAA-26B	Disease resistance protein (TIR-NES-LRR...	3,00E-10	no hits	
P2-16	AS-P216MCTC-27A	Disease resistance protein (TIR-NES-LRR...	6,00E-14	no hits	
P2-16	AD-P216MCTC-27B	Disease resistance protein (TIR-NES-LRR...	5,00E-22	pfam_fs:NB-ARC Domain	3,20E-10
P2-16	AD-P216MC-28G	Disease resistance protein (TIR-NES-LRR...	2,00E-04	no hits	
P2-19	AS-P219MCTA-12F	Disease resistance protein (TIR-NES-LRR...	1,00E-06	pfam_fs:NB-ARC Domain	8,70E-06
P2-414	AS-P2414MCTC-10A	Disease resistance protein, RPP13-like ...	3,00E-24	pfam_fs:NB-ARC Domain	1,70E-45
NBS 5	AD-NBSMGCC-8B	Disease resistance protein (NES-LRR cla...	4,00E-22	pfam_fs:NB-ARC Domain	1,80E-27
NBS 5	AS-NBSMGCC-8C	Disease resistance protein (NES-LRR cla...	7,00E-26	pfam_fs:NB-ARC Domain	3,30E-34
NBS 5	AD-NBSMGCC-8D	Disease resistance protein (CC-NES clas...	3,00E-11	pfam_fs:NB-ARC Domain	1,30E-11
NBS 5	AS-NBSMGCC-8J	Disease resistance protein (CC-NES clas...	0.003	pfam_fs:NB-ARC Domain	1,40E-08
NBS 5	AD-NBSMCTC-14A	Disease resistance protein (CC-NES clas...	0.005	pfam_fs:NB-ARC Domain	1,30E-08
NBS 5	AS-NBSMCTC-14D	Disease resistance protein (CC-NES-LRR ...	0.002	pfam_fs:NB-ARC Domain	5,40E-07
NBS 5	AS-NBSMCTT-24A	Disease resistance protein (NES-LRR cla...	2,00E-23	pfam_fs:NB-ARC Domain	4,50E-25
NBS 5	AS-NBSMCTT-24E	Disease resistance protein (CC-NES clas...	7,00E-14	pfam_fs:NB-ARC Domain	5,80E-30
NBS 5	AD-NBSMCTT-24F	Putative late blight resistance protein (Solanum...	0.29	pfam_fs:NB-ARC Domain	3,70E-04

^a Homologias encontradas com seqüências de nucleotídeos traduzidas para aminoácidos pelo próprio BLASTX e comparadas com o GenBank. ^b seqüências traduzidas pelo programa “Translate Tool - TrEMBL” da Home Page: www.expasy.ch/tools/dna.html e processadas pela ferramenta “Motif-Scan” da Home Page: www.expasy.ch/prosite.

Tabela 5 Bandas polimórficas seqüenciadas e suas principais homologias com seqüências conhecidas no GenBank.

Primer RGA	Marcador ^a	Hit ^b	E-value ^c
P2-16	AS-P216MCTC-27D	gi 15221080 ref NP_175246.1 hydrolase/ protein serine/threon...	4,00E-07
P2-19	AD-P219MCTA-12B	At1g64260.1 68414.m07281 zinc finger protein-related contains Pf...	3,00E-10
P2-378	AS-P2378MCTC-11B	GI:49533772 putative polyprotein [Solanum demissum].	4,00E-57
P2-378	AS-P2378MCTC-11C	GI:7267154 putative transposon protein [Arabidopsis thaliana].	1,00E-09
P2-378	AD-P2378NCTC-11D	gi 2995405 emb CAA73042.1 polyprotein [Ananas comosus]	6,00E-31
P2-378	AD-P2378MC-20B *	At4g23160.1 68417.m03342 protein kinase family protein contains ...	5,00E-04
P2-378	AA-P2378MC-20D	GI:49533772 putative polyprotein [Solanum demissum].	2,00E-27
P2-378	AS-P2378MC-20H	gi 15724009 gb AAL06413.1 reverse transcriptase [Pisum sativum]	9,00E-11
P2-378	AD-P2378MCTA-23A*	At4g23160.1 68417.m03342 protein kinase family protein contains ...	3,00E-09
P2-378	AS-P2378MCTA-23E	gi 15724009 reverse transcriptase [Pisum sativum]	2,00E-11
P2-391	AS-P2391MCTA-13C *	At1g21230.1 68414.m02653 wall-associated kinase, putative simila...	6,00E-07
P2-391	AD-P2391MCTA-13D *	At1g21230.1 68414.m02653 wall-associated kinase, putative simila...	2,00E-11
P2-391	AD-P2391MGAG-16B	At1g72520.1 68414.m08386 lipoxygenase, putative similar to lipox...	2,00E-10
P2-414	AD-P2414MCTC-10C	GI:16118447 heat shock transcription factor [Phaseolus acutifolius].	3,00E-21
P2-414	AD-P2414MC-17B	gi 6691193 gb AAF24531.1 F7F22.17 [Arabidopsis thaliana]	4,00E-36
P2-414	AS-P2414MC-17D	GI:16118447 heat shock transcription factor [Phaseolus acutifolius].	4,00E-20
P2-414	AS-P2414MCTT-19A	GI:50930529 OSJNBa0014F04.3 [Oryza sativa (japonica cultivar-group)].	0.002
NBS 5	AS-NBSMCTC-14B	At2g32290.1 68415.m03947 beta-amylase, putative / 1,4-alpha-D-gl...	5,00E-36
NBS 5	AD-NBSMCTC-14C	At2g32290.1 68415.m03947 beta-amylase, putative / 1,4-alpha-D-gl...	3,00E-41
NBS 5	AD-NBSMGAG-22C	GI:4567295 putative retroelement pol polyprotein [Arabidopsis thaliana].	3,00E-18
NBS 5	AD-NBSMCTT-24B	gi 54644620 gb EAL33361.1 GA11129-PA [Drosophila pseudoobscura]	3,90E-01
NBS 5	AS-NBSMCTT-24H *	At1g52540.1 68414.m05931 protein kinase, putative contains prote...	4,00E-36
NBS 5	AD-NBSMCTT-24N	At3g16190.1 68416.m02044 isochorismatase hydrolase family protei...	0.020
NBS 5	AS-NBSMGGC-8E	gi 55770171 ref XP_550075.1 hypothetical protein [Oryza sati...	0.001

*Provável proteína do tipo Kinase. ^a Nome dado a cada marcador identificado no trabalho. O nome possui as iniciais da planta de origem (AD – *A. duranensis*; AS – *A. stenosperma*), os dois *primers* utilizados na amplificação por PCR (por exemplo: P212MCT – P2-12 e Mse-ct) e o número da combinação de *primers* associado com a letra da banda extraída do gel. ^b Principal homologia encontrada no GenBank. O programa utilizado para a comparação com o banco de dados foi o BLASTX. ^c E-Value é o número correspondente ao nível de significância. Quanto menor o valor, melhor a qualidade da homologia. O nome do marcador pode ser resumido ao número e à letra final do nome, por exemplo: 10A, 24C.

Figuras

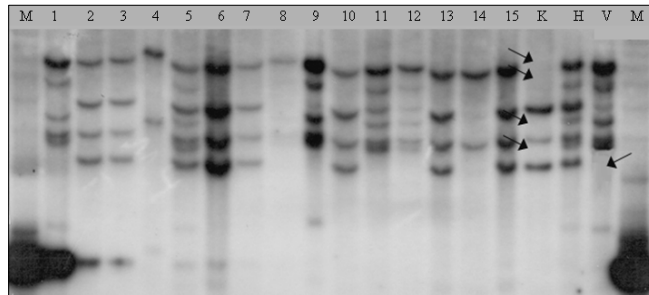


Figura 1 Auto-radiografia de Southern-blot contendo 15 plantas da população F₂ de mapeamento de *Arachis* representadas pelos números. As amostras foram digeridas com a enzima de restrição *EcoRI* e hibridizadas com a sonda RGA S1-A-36. “M” representa o marcador 1 Kb Ladder (Promega), “K” representa a planta *A. duranensis*, “H” representa a planta híbrida e “V” representa a planta *A. stenosperma*. As setas mostram as bandas polimórficas entre os parentais e que segregam nas plantas da população F₂.

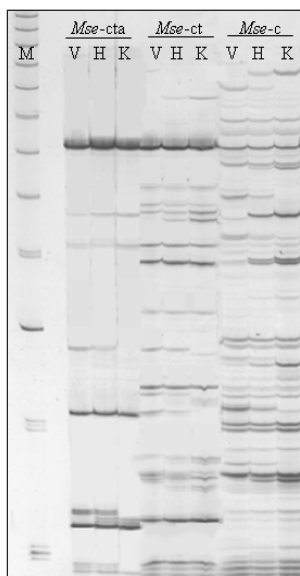
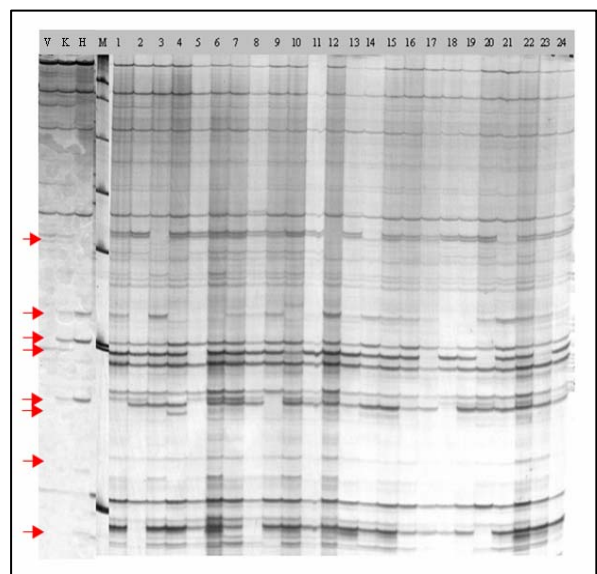


Figura 2 Gel de poliacrilamida 4% corado com Nitrato de Prata contendo amostras de RGA-Display amplificadas com o *primer* P2-378 relativo ao motivo Kinase-2 no Domínio NBS no sentido direto e outro relativo ao sítio de restrição da enzima *MseI* com três, duas ou uma base seletivas. As amostras da esquerda foram amplificadas com o *primer* *Mse-cta* (3 bases seletivas), as do centro do gel foram amplificadas com o *primer* *Mse-ct* (2 bases seletivas) e as da direita com o *primer* *Mse-c* (1 base seletiva). V representa *A. stenosperma*, H representa a planta híbrida e K representa a planta *A. duranensis*. M é o marcador 1 kb plus (Promega).

Figura 3 Gel de poliacrilamida corado com Nitrato de prata contendo amostras de RGA-Display com parte da população F₂ de *Arachis*. As amostras foram amplificadas com os *primers* P2-414 (degenerado) e *M-c*. “M” representa o marcador 1Kb Plus. A planta *A. duranensis* está representada por “K”, *A. stenosperma* por “V” e a planta híbrida por “H”. Os números na parte superior do gel representam as 24 primeiras plantas da população F₂ genotipadas com essa combinação de *primers*. As setas indicam as bandas polimórficas entre os parentais e que segregaram na população F₂.



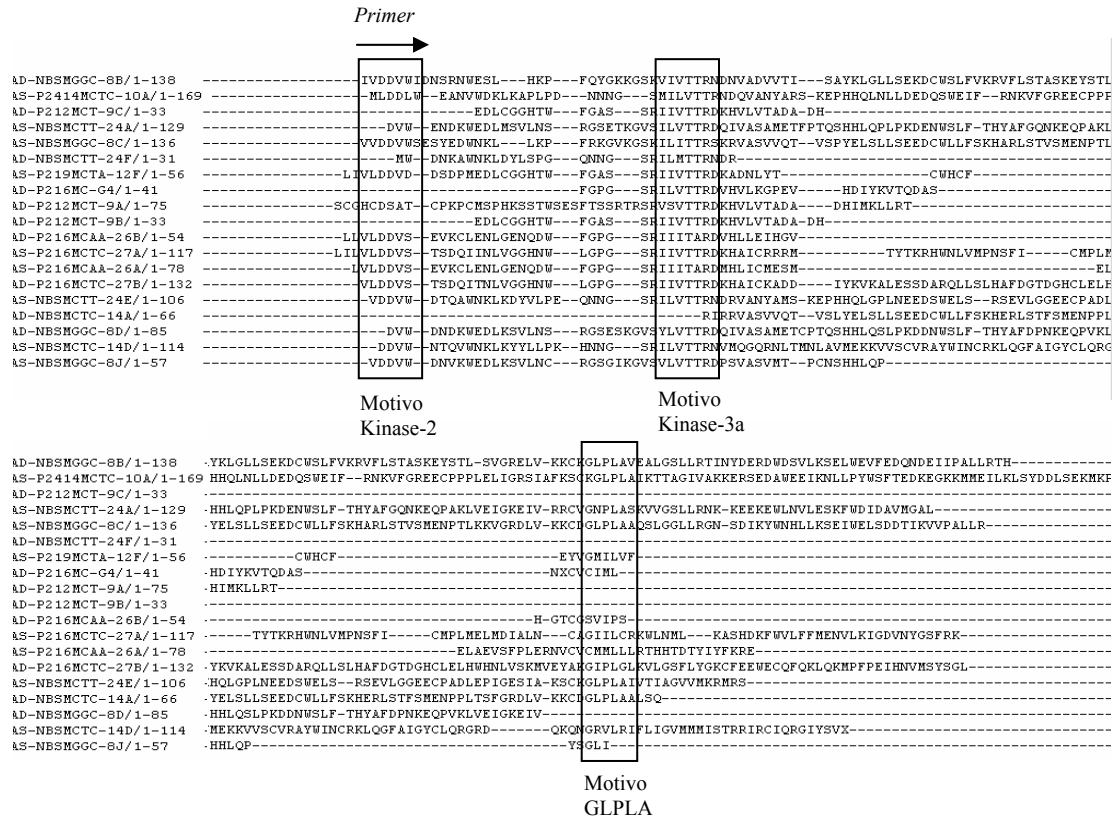


Figura 4 Imagem de alinhamento feita pelo programa Jalview com tipo de alinhamento ClustalW. As 19 seqüências traduzidas de RGAs encontrados neste trabalho foram alinhadas e tiveram as regiões conservadas destacadas. As caixas mostram os motivos Kinase-2, Kinase-3a e GLPLA do domínio NBS. A seta representa o sentido do primers NBS e a região de anelamento do mesmo. Os números após o nome dos RGAs representam o números de aminoácidos em cada seqüência.

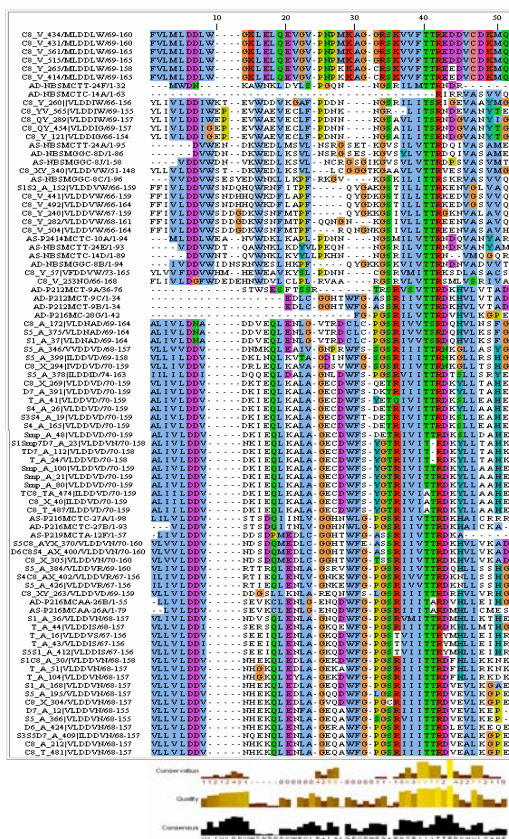


Figura 5 Imagem Jalview com alinhamento das seqüências RGAs, proteínas de resistência de *Arachis* spp. e domínio NBS. A parte inferior da figura mostra o grau de conservação da seqüência consensus e a qualidade do alinhamento.

Figura 6 Árvore de similaridade baseada em um alinhamento ClustalW de seqüências de proteínas dos RGAs encontrados neste trabalho. Seqüências do motivo Kinase-2 até GLPLA foram usadas para o alinhamento. Os números indicam a distância média calculada através da similaridade das seqüências. As amostras foram divididas de acordo com sua estrutura na porção N-terminal TIR ou NãoTIR.

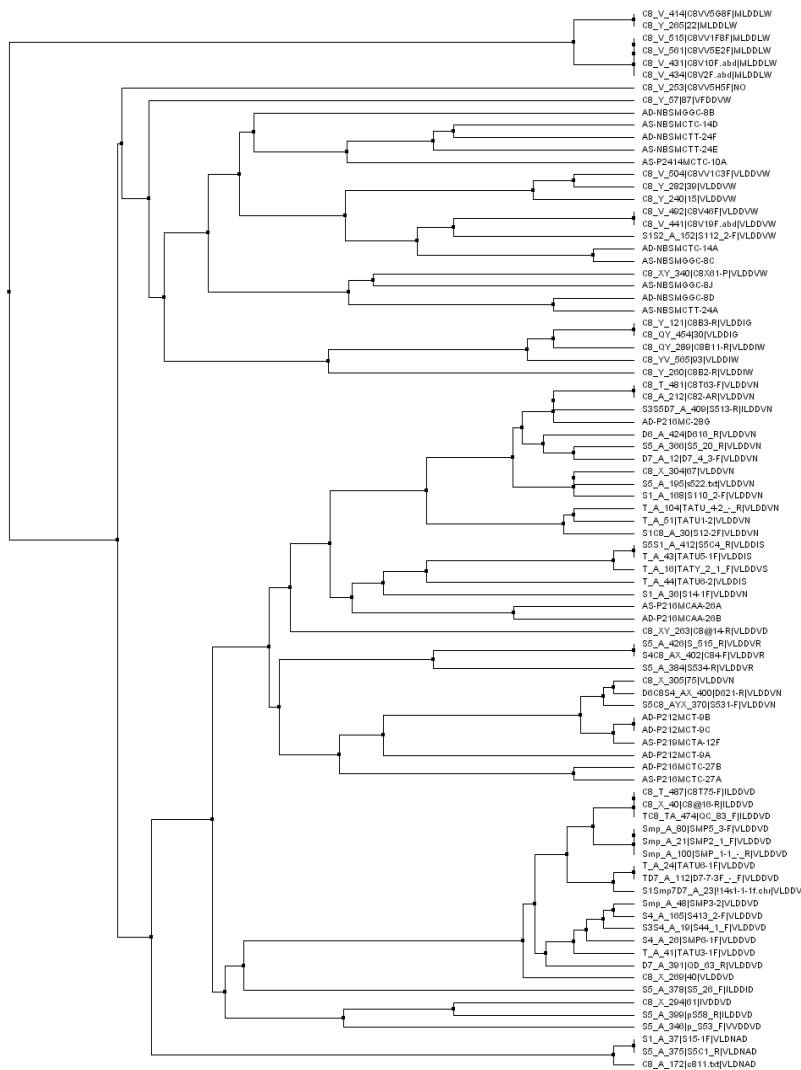
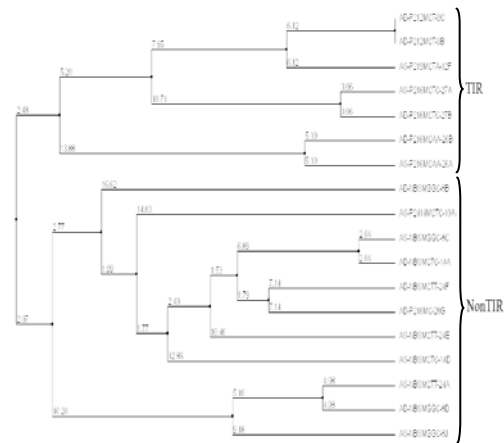


Figura 7 Árvore de similaridade das 19 amostras RGAs obtidas nesse trabalho e 67 obtidas em Bertoli et al., 2003. A árvore foi construída baseada em um alinhamento do tipo ClustalW de seqüências protéicas cortadas entre o motivo Kinase-2 e Kinase-3a.

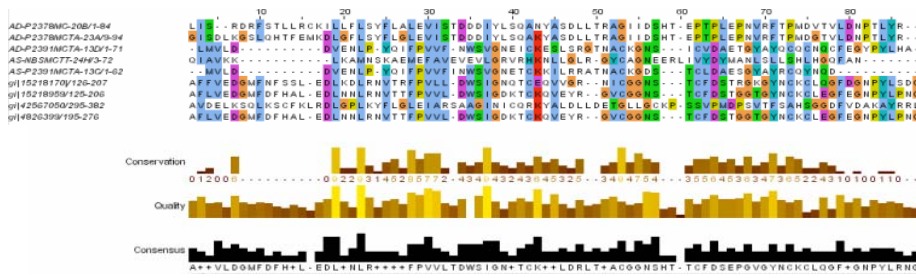


Figura 8 Alinhamento ClustalW utilizando o programa JalView de seqüências de proteínas do tipo kinase. As cinco primeiras amostras são seqüências obtidas do processamento de bandas polimórficas da técnica RGA-Display e as quatro últimas são proteínas do tipo kinase de domínio público depositadas no GenBank (www.ncbi.nlm.nih.gov). A parte inferior da figura representa os gráficos de conservação das seqüências, a qualidade do alinhamento e a seqüência *consensus* obtida do alinhamento.

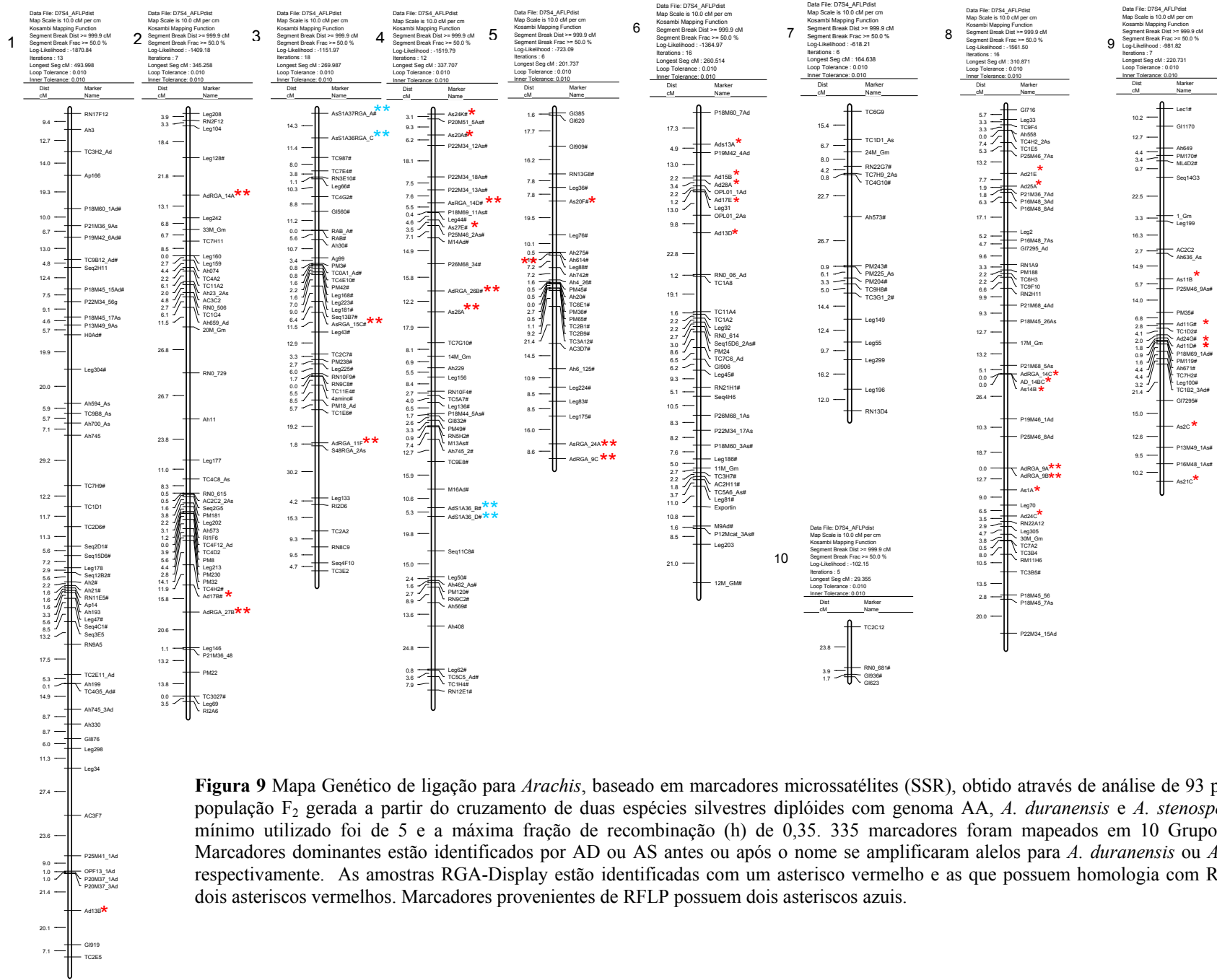


Figura 9 Mapa Genético de ligação para *Arachis*, baseado em marcadores microssatélites (SSR), obtido através de análise de 93 plantas de uma população F₂ gerada a partir do cruzamento de duas espécies silvestres diplóides com genoma AA, *A. duranensis* e *A. stenosperma*. O LOD mínimo utilizado foi de 5 e a máxima fração de recombinação (h) de 0,35. 335 marcadores foram mapeados em 10 Grupos de Ligação. Marcadores dominantes estão identificados por AD ou AS antes ou após o nome se amplificaram alelos para *A. duranensis* ou *A. stenosperma* respectivamente. As amostras RGA-Display estão identificadas com um asterisco vermelho e as que possuem homologia com RGAs possuem dois asteriscos vermelhos. Marcadores provenientes de RFLP possuem dois asteriscos azuis.

