

Universidade Católica de Brasília

PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
STRICTO SENSU EM CIÊNCIAS GENÔMICAS E BIOTECNOLOGIA

Mestrado

**CARACTERIZAÇÃO E MAPEAMENTO DE RGAS EM
*Arachis***

Autor: Bruna Vidigal dos Santos
Orientador: David John Bertioli
Co-orientadora: Patrícia Messenberg Guimarães

BRASÍLIA

2009

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

BRUNA VIDIGAL DOS SANTOS

CARACTERIZAÇÃO E MAPEAMENTO DE RGAS EM *Arachis*

Dissertação apresentada ao Programa de Pós-graduação Stricto Sensu em Ciências Genômicas e Biotecnologia da Universidade Católica de Brasília, como requisito parcial para obtenção do Título de Mestre em Ciências Genômicas e Biotecnologia

Orientador: David John Bertioli
Co-orientadora: Patrícia Messenberg
Guimarães

**Brasília
2009**



Dissertação de autoria de Bruna Vidigal dos Santos, intitulada “CARACTERIZAÇÃO E MAPEAMENTO DE RGAS EM *Arachis*”, apresentada como requisito parcial para obtenção do grau de Mestre em Ciências Genômicas e Biotecnologia de Universidade Católica de Brasília, em 05 de junho de 2009, defendida e aprovada pela banca examinadora abaixo assinada:

Prof. Dr. David John Bertioli
Orientador
Universidade Católica de Brasília - UCB

Dra. Patrícia Messenberg Guimarães
Co-orientadora
Embrapa – Recursos Genéticos e Biotecnologia

Prof. Dr. Robert Neil Gerard Miller
Universidade Católica de Brasília – UCB

Dra. Gláucia Salles Cortopassi Buso
Embrapa – Recursos Genéticos e Biotecnologia

Brasília
2009

A minha família pelo carinho, incentivo e apoio
constantes,

DEDICO.

iv

AGRADECIMENTOS

A Deus, pela minha existência, pela saúde, força e amor que me permitiu todos os dias respirar e finalizar este trabalho;

À Universidade Católica de Brasília e ao Curso de Pós-Graduação em Ciências Genômicas e Biotecnologia pela estrutura excepcional, além da oportunidade de realização do Mestrado;

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudos;

Ao Professor e Orientador David John Bertioli pela oportunidade, apoio e confiança durante o desenvolvimento deste trabalho;

À minha co-orientadora Patrícia Messemberg pelos construtivos ensinamentos e opiniões acerca do trabalho desenvolvido;

Aos membros da banca examinadora Robert Miller e Gláucia Buso, pelas excelentes sugestões para finalização deste manuscrito;

Aos doutores e consultores da Embrapa Recursos Genéticos e Biotecnologia: Soraya Bertioli, Simone Ribeiro, Ana Cristina Brasileiro, Ana Cláudia Guerra e Stephan Nielen;

Aos meus AMIGOS de vida e de laboratório: Iugo, Pedro, Andressa, Talles, Nina, Rafaela, Zélia, Leandro, Daiane, Carolina, Valéria, Fernando e Dione.

A toda a equipe envolvida nos projetos de amendoim, pela colaboração;

Aos Professores da UCB pelos valiosos ensinamentos durante o curso. Sempre serei grata a vocês;

À minha família, pelo amor e apoio em toda a minha vida;

A todos que contribuíram, mas não foram mencionados... Obrigada

" A imaginação é mais importante que o conhecimento."

" Toda nossa ciência, comparada com a realidade, é primitiva e infantil – e, no entanto, é a coisa mais preciosa que temos."

(Albert Einstein)

RESUMO

O amendoim (*Arachis hypogaea*) está entre as principais oleaginosas cultivadas em todo o mundo. É bastante susceptível a diversos estresses bióticos e abióticos, para os quais muitas das espécies silvestres são resistentes. O processo de introgressão de genes úteis de espécies silvestres em amendoim pode ser acelerado pelo uso de seleção assistida por marcadores moleculares, através da identificação de marcadores proximamente ligados a genes de interesse como os de resistência a doenças. QTLs (*Quantitative Trait locus*) para resistência à mancha foliar causada pelo fungo *C. personatum* foram mapeados em mapa genético de referência construído para o cruzamento entre *A. duranensis* e *A. stenosperma*. O RGA (*Resistance Genes Analog*) S1_A_36 que co-segrega com um destes QTLs, localizado no Grupo de Ligação 4, foi utilizado para identificar clones positivos nas bibliotecas BAC (*Bacterial Artificial Chromosome*) de espécies silvestres do gênero *Arachis*. A subclonagem e sequenciamento do BAC que contém este RGA possibilitou a caracterização da sequência, identificando, com auxílio do Programa de predição RiceGAAS, 20 ORFs distribuídas em oito *contigs*. Destas, cinco ORFs estavam relacionadas a genes de resistência pertencentes à família TIR-NBS-LRR e quatro ORFs a elementos de transposição, sendo um transposon e três retrotransposons LTR do tipo Ty1-Copia. Este trabalho possibilitou ainda, o mapeamento de cinco marcadores microsatélites derivados do BAC, que mapearam no mesmo grupo de ligação dos (RGAs). A utilização de marcadores microsatélites exhibe vantagens adicionais à utilização de RGAs no processo de seleção assistida, como a praticidade operacional, alta reprodutibilidade, por ser baseado em PCR, e principalmente por ser um marcador transferível entre espécies relacionadas. A identificação de marcadores mais robustos e transferíveis associados a locos que conferem resistência possibilitará a seleção assistida nos programas de melhoramento genético, o que acelerará o processo de introgressão dos genes de resistência no amendoim.

ABSTRACT

Peanut (*Arachis hypogaea*) is one of the major oilseed crops worldwide. It is very susceptible to various biotic and abiotic stresses, to which many wild species are resistant. Introgression of useful genes from wild species in peanut can be accelerated by the use of molecular marker assisted selection (MAS), through the identification of markers closely linked to genes of interest such as resistances to diseases. QTLs for resistance to leaf-spot caused by the fungus *C. personatum* were mapped in the reference genetic map from the cross between *A. duranensis* and *A. stenosperma*. The RGA S1_A_36 that co-segregated with one QTL, located in Linkage Group 4, was used to identify positive clones in BAC libraries of two wild species of *Arachis*. The subcloning and sequencing of a full BAC containing the RGA enabled the characterization of this sequence, and its annotation using the gene prediction program RiceGAAS. In total 20 ORFs were found distributed in eight *contigs*. Of these, five ORFs contained resistance genes analogs of the TIR-NBS-LRR family and four contained transposable elements, one DNA transposon and three retrotransposons of the LTR Ty1-Copia type. This work also enabled the mapping of five microsatellite markers, which were BAC derived, in the same linkage group of the RGA studied. Microsatellite markers are attractive as they are PCR based and show good reproducibility and transferability among related species. The identification of such markers associated with resistance loci is of great importance as it has the potential of speeding up the process of introgression of resistance genes in peanut.

LISTA DE FIGURAS

INTRODUÇÃO

Figura 1: (A) *Arachis hypogaea* L. (B) Fruto desta espécie. Retirado do endereço eletrônico: <http://www.agraria.org/coltivazionierbacee/arachide.htm>.6

Figura 2: Domínio NBS-LRR representado esquematicamente. As regiões TIR e Não-TIR também estão ilustradas. Os motivos P-loop, Kinase-2, Kinase-3a e GLPL estão representados em destaque na parte inferior da figura.12

Figura 3: Vetor de clonagem pIndigoBAC-5.16

CAPÍTULO I

Figura 1: Grupo de Ligação 4 do mapa genético de *Arachis* (Moretzsohn *et al.*, 2005). A seta indica a posição do RGA S1_A_36. Em verde estão indicados dois QTLs para resistência ao fungo *C. personatum*.23

Figura 2: Metodo de subclonagem de BAC através da técnica de *shotgun*.25

Figura 3: Esquema de sequenciamento do BAC AD25F09.....28

Figura 4: Desenho esquemático do nebulizador. Retirado de www.invitrogen.com..33

Figura 5: Marcador 1 Kb plus (A). Digestão do BAC AD25F09 com a enzima de restrição *HindIII* (B). *Southern Blotting* realizado com o BAC AD25F09 digerido e hibridizado com uma sonda construída a partir de um produto de PCR referente ao RGA S1_A_36 (C). Imagens sobrepostas (D).....399

Figura 6: Gel de agarose 1% corado com brometo de etídeo, contendo o marcador molecular 1Kb plus (Invitrogen) e o BAC AD25F09 após a nebulização.....41

Figura 7: Gel de agarose 1% corado com brometo de etídeo referente à minipreparação de plasmídeos do BAC AD25F09 no vetor de clonagem pCR 4Blunt-TOPO (3.957 pb) (Invitrogen). O marcador molecular utilizado foi o 1Kb plus (Invitrogen).42

Figura 8: Janela da Interface GAP (Staden Package) mostrando um editor de contig. A Figura exemplifica a formação de uma sequência consenso a partir das sequências obtidas através das duas metodologias de subclonagem. O círculo grande mostra sequências advindas de placas confeccionadas pela fragmentação do BAC através de nebulização e o círculo pequeno mostra uma sequência advinda da fragmentação do BAC pela enzima de restrição *SauBAI*.....43

Figura 9: Janela da Interface GAP (Staden Package) mostrando todos os *contigs* produzidos a partir da subclonagem do BAC AD25F09. Em marron, a escala de

formação dos 505 *contigs*, do maior até o menor. Em verde, os *contigs* selecionados para anotação. Em vermelhos, os *contigs* não analisados em virtude do pequeno tamanho.43

Figura 10: Disposição das possíveis regiões codificadoras de proteínas nos *contigs* montados a partir do BAC AD25F09...52

CAPÍTULO II

Figura 1: Esquema de geração de marcadores microsatélites a partir BAC AD25F09.....77

Figura 2: Gel de agarose 1% corado com brometo de etídeo referente à quantificação de três amostras de DNA (Amostras 1, 8 e 10) junto com o marcador molecular High Mass Ladder (Invitrogen).....83

Figura 3: Géis de poliacrilamida a 4%. (A) Primer AD25F09-PC-H08; (B) Primer AD25F09-PA-D03.85

Figura 4: Géis de poliacrilamida a 4%. (A) Primer AD25F09-HIII-82; (B) Primer AD25F09-HIII-34.86

Figura 5: Géis de poliacrilamida a 4% mostrando a amplificação com cinco pares de *primers* nos parentais da população A e F1 resultante deste cruzamento. (A) AD25F09-HIII-82; (B) AD25F09-HIII-34; (C) AD25F09-HIII-P6H06; (D) AD25F09-PB-H01; (E) AD25F09-PD-A10.86

Figura 6: Géis de poliacrilamida a 4% mostrando a amplificação com cinco pares de *primers* nas duas cultivares de *A. hypogaea* (IAC-Runner 886 e IAC-Caiapó). (A) AD25F09-HIII-82; (B) AD25F09-HIII-34; (C) AD25F09-PB-H01; (D) AD25F09-PD-A10; (E) AD25F09-HIII-P1E08.87

Figura 7: Grupo de Ligação 4 pertencente ao Mapa Genético de *Arachis* (Moretzsohn *et al.*, 2005). Circulados estão destacados os marcadores SSRs desenvolvidos neste estudo.88

LISTA DE TABELAS

CAPÍTULO I

Tabela 1: Número de colônias recombinantes e sequências obtidas através das diferentes estratégias de subclonagem.....41

Tabela 2: Tabela Composição dos oito *contigs* selecionados da subclonagem do BAC AD25F09.....43

Tabela 3: Lista de ORFs preditas para cada *contig*, através de análises feitas pelo Programa RiceGAAS.....52

CAPÍTULO II

Tabela 1: Lista de espécies do gênero *Arachis* selecionadas para extração de DNA genômico para utilização em PCR. A Tabela mostra o tipo de genoma referente a cada espécie..78

Tabela 2: Quantificação de DNA genômico de espécies de *Arachis* em Nanodrop.. 82

Tabela 3: *Primers* utilizados para amplificação de microssatélites.....84

Tabela 4: Localização de microssatélites nos *contigs* do BAC AD25F09 e resultado da amplificação em alguns genótipos do gênero *Arachis*.....89

LISTA DE GRÁFICOS

CAPÍTULO I

Gráfico 1: Porcentagem de ORFs identificadas no BAC AD25F09.....	51
Gráfico 2: <i>Contig1</i> (AD25F09-PO-F10.FM)	56
Gráfico 3: <i>Contig2</i> (AD25F09-HIII-P2F03.FM)	57
Gráfico 4: <i>Contig3</i> (AD25F09-PC-F09.RM)	58
Gráfico 5: <i>Contig4</i> (AD25F09-PH-B11.RM).....	59
Gráfico 6: <i>Contig5</i> (AD25F09-HIII-P2G03.RM)	60
Gráfico 7: <i>Contig6</i> (AD25F09-PL-B03.FM).....	61
Gráfico 8: <i>Contig7</i> (AD25F09-PN-C10.RM).....	62
Gráfico 9: <i>Contig8</i> (AD25F09-PG-G04.RM)	63
Gráfico 10: Comparação das sequências dos <i>contigs</i> gerados a partir da subclonagem do BAC AD25F09, com bancos de dados de ESTs depositados no GenBank (NCBI).	64
Gráfico 11: Comparação das sequências dos <i>Contigs</i> gerados a partir da subclonagem do BAC AD25F09, com bancos de dados de sequências GSS depositadas no GenBank (NCBI).	64

LISTA DE SÍMBOLOS E ABREVIATURAS

° C – Graus Celsius

– Ohm

ADP – Adenosina difosfato

AFLP – *Amplified Fragment Length Polymorphism* (Polimorfismo de comprimentos de fragmentos amplificados)

cm – Centímetros

cM – Centimorgans

BAC – *Bacterial Artificial Chromosome* (Cromossomo artificial de bactéria)

BAG – Banco Ativo de Germoplasma

*Bam*HI - *Bacillus amyloliquefaciens*H

BLAST – *Basic Local Alignment and Search Tool*

BSA – *Bovine Serum Albumin* (Albumina de soro bovino)

CC – *Coiled coil* (Cauda espiralada)

cDNA – *Complementary DNA* (DNA complementar)

CIP – Calf Intestinal Alkaline Phosphatase (Fosfatase alcalina intestinal de bezerro)

CTAB – *Cetyl-trimethylammoniumbromide* (Brometo de cetiltrimetilamônio)

DNA – Ácido Desoxirribonucleico (*Desoxiribonucleic Acid*)

dNTPs – *Deoxyribonucleoside Triphosphates* (Desoxirribonucleotídeos trifosfatados)

EDTA – *Ethylenediamine tetraacetic acid* (Ácido etilenodiaminotetracético)

EST – *Expressed Sequence Tag* (Etiquetas de sequências expressas)

g – Grama

GL – Grupo de Ligação

GSS – *Genome Survey Sequence* (Sequência de amostragem do genoma)

GET – Glucose/EDTA/TRIS-HCl

*Hind*III – *Haemophilus influenzae* Rd

IPTG – *Isopropyl -D-1-thiogalactopyranoside*

Kb – Kilobase

KOAc – Acetato de Potássio

KW – *kilowatts*

LB – Meio de cultura Luria-Bertani

LRR – *Leucine Rich Repeat*(Repetição rica em leucina)
M – Molar
mg – Miligrama
MgCl₂ – Cloreto de Magnésio
mL – Mililitro
mM – Milimolar
mRNA – RNA mensageiro (*Messenger RNA*)
μFD – *Microfaraday*
μg – Micrograma
μL – Microlitro
μM – Micromolar
NaCl – Cloreto de Sódio
NaOH – Hidróxido de Sódio
NBS – *Nucleotide Binding Site*(Sítio de ligação a nucleotídeos)
NCBI – *National Center for Biotechnology Information* (Centro Nacional para Informação de Biotecnologia)
ng – Nanograma
OD – *Optical Density*(Densidade óptica)
ORF – *Open Reading Frame*(Quadro de leitura aberta)
pb – Pares de bases
PCR – *Polymerase Chain Reaction*(Reação em cadeia da polimerase)
pH – Potencial Hidrogeniônico
PVP – Polivinilpirrolidona
QTL – *Quantitative Trait Locus*(Locus de caracter quantitativo)
RAPD – *Randomly Amplified Polymorphism* (Polimorfismo amplificado randomicamente)
RFLP – *Restriction Fragment Length Polymorphism* (Polimorfismo de tamanho de fragmentos de restrição)
RGA – *Resistance Gene Analogue*(Análogos a genes de resistência)
RiceGAAS – *Rice Genome Automated Annotation System* (Sistema de anotação automatizada do arroz)
RNA – *Ribonucleic Acid*(Ácido ribonucleico)
RNase – Ribonuclease

rpm – Rotações por minuto
RT – *Reverse Transcriptase* (Transcriptase reversa)
SAM – Seleção Assistida por marcador
SAP – *Shrimp Alkaline Phosphatase* (Fosfatase alcalina de camarão)
SAR – *Systemic acquired resistance* (Resistência sistêmica adquirida)
Sau3AI – *Staphylococcus aureus* 3A
SCAR – *Sequence-Characterized Amplified Regions* (Regiões amplificadas de sequências caracterizadas)
SDS – *Sodium Dodecyl Sulfate* (Dodecil sulfato de sódio)
SSR – *Simple Sequence Repeats* (Sequências de repetições simples)
Taq – *Thermus aquaticus*
TAE – Tris-acetato-EDTA
TBE – Tris-Base-EDTA
TIR – *Toll interleukine 1-receptor* (Receptor Toll/interleucina 1)
tRNA – *Transporter RNA* (RNA transportador)
TROLL – *Tandem Repeat Occurrence Locato*
U – Unidade
UV – Radiação Ultravioleta
V – Volts
X-GAL – *5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside*
W – *Watts*

SUMÁRIO

INTRODUÇÃO	4
1. O Gênero <i>Arachis</i>	4
2. O amendoim	5
3. Espécies Silvestres.....	7
4. Doenças em amendoim	8
5. Genes de Resistência.....	10
6. Marcadores moleculares, mapas genéticos e QTLs	13
7. Bibliotecas BAC	15
8. Subclonagem.....	17
OBJETIVO GERAL.....	19
OBJETIVOS ESPECÍFICOS.....	19
APRESENTAÇÃO DOS CAPÍTULOS	20
CAPÍTULO I – SUBCLONAGEM, SEQUENCIAMENTO E ANOTAÇÃO DO BAC AD25F09	21
1. INTRODUÇÃO	22
1.1 Sequenciamento através de <i>shotgun</i>	24
1.2 Anotação de sequências.....	26
2. MATERIAL E MÉTODOS.....	27
2.1 Plantas de <i>Arachis</i> e bibliotecas BAC.....	27
2.2 Isolamento de RGAs da biblioteca BAC de espécies de <i>Arachis ipaënsis</i> e <i>A. duranensis</i>	27
2.3 Extração de BAC	29
2.4 Confirmação da presença do inserto no vetor	30
2.5 Estratégias para subclonagem.....	30
2.5.1 Estratégia de subclonagem 1 – Fragmentação através de Enzimas de Restrição	31
a) Digestão do BAC	31
b) Digestão do vetor de subclonagem e desfosforilação.....	31
c) Ligação dos fragmentos com extremidades coesivas ao vetor pBlueScript SK-	32
2.5.2 Estratégia de subclonagem 2 – Fragmentação através de nebulização.....	32

a)	Nebulização do BAC.....	32
b)	Tratamento dos fragmentos com extremidades não-coesivas.....	33
c)	Desfosforilação	34
d)	Ligação dos fragmentos tratados ao vetor pCR 4Blunt-TOPO	34
2.6	Preparo de células eletrocompetentes.....	34
2.7	Eletroporação.....	35
2.8	Seleção de transformantes e cultura permanente	35
2.9	Minipreparação de plasmídeos em tubo	36
2.10	Minipreparação de plasmídeos em microplaca.....	36
2.11	Sequenciamento	37
2.12	Análise dos dados de sequenciamento	38
3.	RESULTADOS	39
3.1	Confirmação da presença do RGA	39
3.2	Estratégias de subclonagem.....	40
3.3	Confirmação da presença do inserto no vetor	42
3.4	Análise dos dados do sequenciamento	42
3.4.1	<i>Contig1</i> (AD25F09-PO-F10.FM).....	45
3.4.2	<i>Contig2</i> (AD25F09-HIII-P2F03.FM).....	46
3.4.3	<i>Contig3</i> (AD25F09-PC-F09.RM)	47
3.4.4	<i>Contig4</i> (AD25F09-PH-B11.RM)	48
3.4.5	<i>Contig5</i> (AD25F09-HIII-P2G03.RM)	48
3.4.6	<i>Contig6</i> (AD25F09-PL-B03.FM)	49
3.4.7	<i>Contig7</i> (AD25F09-PN-C10.RM)	50
3.4.8	<i>Contig8</i> (AD25F09-PG-G04.RM).....	50
4.	DISCUSSÃO.....	65
CAPÍTULO II - DESENVOLVIMENTO E MAPEAMENTO DE MARCADORES		
MICROSSATÉLITES A PARTIR DO BAC AD25F09		
1.	INTRODUÇÃO	74
2.	MATERIAL E MÉTODOS.....	77
2.1	Material vegetal e população de mapeamento	77
2.2	Extração do DNA genômico de <i>Arachis</i> e desenho de primers para microssatélites	78
2.3	Amplificação por PCR.....	80

2.4	Mapeamento de microssatélites.....	80
3.	RESULTADOS.....	82
3.1	Extração de DNA e qualidade das amostras.....	82
3.2	Desenho de <i>primers</i> para regiões contendo microssatélites e amplificação destas seqüência.....	83
3.3	Análise dos marcadores microssatélites e mapeamento.....	84
4.	DISCUSSÃO.....	90
	CONCLUSÃO GERAL.....	94
	PERSPECTIVAS FUTURAS.....	96
	REFERÊNCIAS BIBLIOGRÁFICAS.....	97
	ANEXO 1.....	111
	ANEXO 2.....	116

INTRODUÇÃO

1. O Gênero *Arachis*

O gênero *Arachis* L., família Fabaceae, é nativo da América do Sul. O Brasil central é seu provável centro de origem, em uma região que se estende do sudoeste do estado de Mato Grosso do Sul até o sul de Goiás (Valls, 2000). Sua distribuição natural é restrita ao Brasil, Bolívia, Paraguai, Argentina e Uruguai (Valls & Simpson, 1994). Este gênero contém 80 espécies descritas e 64 destas ocorrem no Brasil, sendo 47 exclusivas da flora brasileira (Valls & Simpson, 2005).

A maioria das espécies do gênero é diplóide ($2n=2x=20$ cromossomos), algumas aneuplóides ($2n=2x=18$ cromossomos) e cinco espécies tetraplóides ($2n=4x=40$ cromossomos), incluindo a espécie cultivada *Arachis hypogaea* (Krapovickas & Gregory, 1994; Lavia, 1998; Valls & Simpson, 2005).

As 80 espécies já conhecidas estão reunidas em nove seções taxonômicas (*Arachis*, *Erectoides*, *Heteranthae*, *Caulorrhizae*, *Rhizomatosae*, *Extranervosae*, *Triseminatae*, *Procumbentes* e *Trirectoides*) de acordo com a sua morfologia, distribuição geográfica e compatibilidade sexual (Krapovickas & Gregory, 1994; Valls & Simpson, 2005). A seção *Arachis* é o maior grupo do gênero, totalizando 31 espécies descritas até o momento (Krapovickas & Gregory, 1994; Valls & Simpson, 2005). Possui uma ampla distribuição, desde os Andes até áreas costeiras do Atlântico (Creste *et al.*, 2005).

De acordo com dados citogenéticos, existem três tipos de genomas que diferenciam as espécies em relação aos cromossomos. As espécies caracterizadas com genoma A possuem um pequeno par de cromossomos, denominado par A. As espécies diplóides que não apresentam o par A, possuem o genoma B, com exceção de *A. glandulifera* que possui o genoma denominado D, caracterizado pela presença de seis pares de cromossomos subteloentéricos (Stalker, 1991; Robledo & Seijo, 2008). Existem três espécies que possuem $2n=18$ cromossomos (*A. decora*, *A. palustris* e *A. praecox*) (Lavia, 1998). *A. hypogaea* e *A. monticola* são espécies alotetraplóides com genoma AB.

2. O amendoim

A espécie cultivada, *Arachis hypogaea* (Figura 1A e 1B), está entre as oleaginosas mais cultivadas no mundo, ocupando a 5ª posição dentre as mais produzidas (USDA, 2008).

No Brasil, grande parte da cadeia de produção de amendoim está localizada na região Sudeste, sendo o Estado de São Paulo o maior produtor, respondendo por quase 90% do total produzido no Brasil (Godoy *et al.*, 1999). Porém, o amendoim representa um mercado relativamente pequeno quando comparado a outros produtos agrícolas (Valls, 2005). O cultivo no Brasil representa menos de 1% da produção mundial, mas apesar disso, a produção nacional tem aumentado nos últimos anos, estimando-se na safra 2008/2009, um total de 291,4 mil toneladas em 113,1 mil ha (CONAB, 2009).

Devido ao seu grande valor energético e nutricional, o amendoim é matéria-prima para a indústria alimentícia. Seus grãos são ricos em óleo (40 – 50%) composto em sua maioria por ácidos oléicos (média de 55%), linoléico (25%), palmítico (10%) dentre outros ácidos graxos encontrados em menor quantidade (Knauft *et al.*, 1993; Lopez *et al.*, 2000), proteínas (20 – 30%), carboidratos (10 – 20%), sais minerais e vitaminas (Fagundes, 2002).

O amendoim possui um genoma alotetraplóide. Um alopoliplóide é uma planta híbrida de duas ou mais espécies, contendo duas ou mais cópias de cada um dos genomas componentes. Os alopoliplóides formam-se apenas entre espécies muito correlatas e na natureza, isso pode ser considerado como uma força importante na evolução de novas espécies de plantas (Griffiths *et al.*, 2008).

Vários estudos já foram realizados acerca da origem genética do amendoim. Por ser um alotetraplóide, acredita-se que o amendoim cultivado tenha surgido do cruzamento entre duas espécies silvestres diplóides diferentes, *A. duranensis* (Genoma AA) e *A. ipaënsis* (genoma BB) (Halward *et al.*, 1991; Kochert *et al.*, 1996; Seijo *et al.*, 2004; 2007). O resultado gerou um híbrido, no qual o número de cromossomos foi espontaneamente duplicado levando à restauração da fertilidade e provável seleção do controle genético da meiose. Assim, a planta alotetraplóide resultante tinha vigor híbrido, mas era reprodutivamente isolada de seus parentes selvagens (Halward *et al.*, 1991; Kochert *et al.*, 1996; Seijo *et al.*, 2004; 2007).

Os estudos citológicos de maneira geral têm apoiado os estudos clássicos de cruzamentos, que são realizados com a intenção de refazer o caminho evolutivo que originou o amendoim. Cruzamentos entre diferentes espécies diplóides, no geral resultam em híbridos estéreis. Os híbridos são tratados com colchicina, duplicando o número de cromossomos, podendo então haver um cruzamento com o amendoim cultivado tetraplóide. Estudos acerca da evolução do amendoim cultivado através da cruzabilidade entre *A. duranensis* e *A. ipaënsis* já foram realizados experimentalmente (Fávero, 2004; 2006). Além disso, recentemente, estudos realizados por Seijo *et al.* (2007), utilizando a técnica de GISH (*Genomic in Situ Hybridization*) corroboraram a hipótese de que os mais prováveis parentais da espécie cultivada são *A. duranensis* e *A. ipaënsis*.

A origem de *A. hypogaea* através de um único e recente evento de poliploidização seguido por seleções sucessivas durante os cruzamentos, resultou em um genoma altamente conservado (Young *et al.*, 1996), o que levou a uma baixa diversidade em relação às características agronômicas importantes. Já a variação morfológica observada entre acessos de *A. hypogaea* é provavelmente devido à variação em poucos genes (Kochert *et al.*, 1991). Em virtude dessa baixa variabilidade genética, o amendoim cultivado encontra grandes problemas relacionados à busca de resistência a pragas e patógenos, além da dificuldade na geração de marcadores moleculares polimórficos para a construção de mapas genéticos (Kochert *et al.*, 1996; Halward *et al.*, 1991).

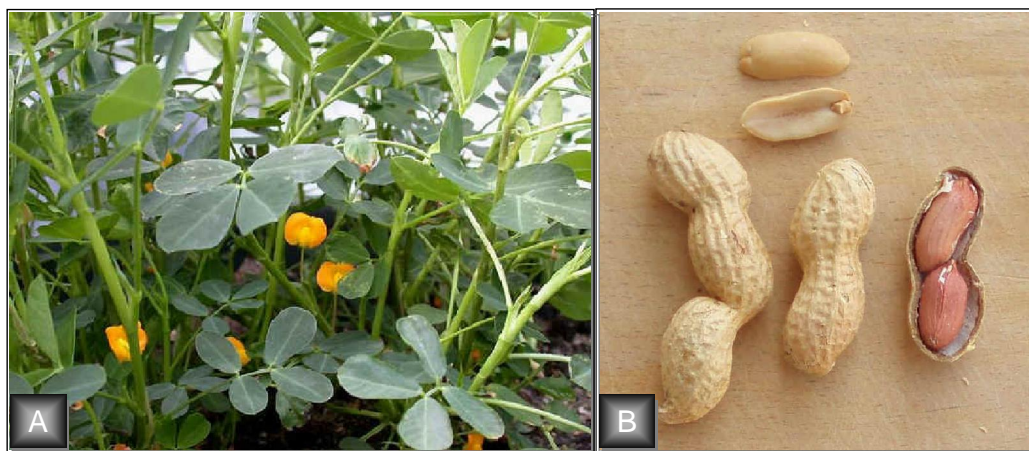


Figura 1: (A) *Arachis hypogaea* L. (B) Fruto desta espécie. Retirado do endereço eletrônico: <http://www.agraria.org/coltivazionierbacee/arachide.htm>.

3. Espécies Silvestres

As espécies silvestres da seção *Arachis*, em sua maioria, são diplóides ($2n=20$) e caracterizam-se por apresentarem alto polimorfismo genético (Kochert *et al.*, 1991; Halward *et al.*, 1991; Hilu & Stalker, 1995; Stalker *et al.*, 1997).

As espécies silvestres diplóide apresentam uma rica fonte de diversidade alélica selecionada durante a evolução em diferentes ambientes e estresses bióticos e abióticos, proporcionando uma alta variação nas características agrônômicas de relevância, como resistência a pragas, patógenos e estresses hídricos (Nelson *et al.*, 1990; Stalker & Simpson, 1995; Burow *et al.*, 2001; Kameswara *et al.*, 2003; Moretzsohn *et al.*, 2004).

Por ser o Brasil central o provável centro de origem do gênero *Arachis*, o Brasil é um grande detentor de germoplasma destas espécies silvestres. 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), e por constituírem importante fonte de genes úteis, espécies silvestres de *Arachis* vêm sendo introduzidas em programas de melhoramento do amendoim. Em estudos realizados de forma pioneira por Isleib & Wynne em 1992, a espécie tetraplóide silvestre *A. monticola* foi utilizada em cruzamentos interespecíficos com *A. hypogaea* resultando nas primeiras cultivares de amendoim com genes de espécies silvestres, Spancross e Tamnut 74, porém sem nenhuma característica fenotípica que pudesse ser identificada como derivada de espécies silvestres.

A primeira cultivar de amendoim com genes identificáveis que foram transferidos de espécies silvestres de *Arachis* foi denominada COAN (resistente a *Meloidogyne javanica* e *M. arenaria*), sendo esta, obtida a partir de retrocruzamentos (RC_5) de *A. hypogaea* cv. Florunner, utilizada como parental recorrente, com um complexo híbrido interespecífico, chamado TxAG-6, resultante do cruzamento entre [*A. batizocoi* x (*A. cardenasii* X *A. diogo*)] e posterior duplicação de cromossomos por colchicina (Simpson & Starr, 2001).

Uma das dificuldades no uso de espécies silvestres no melhoramento genético de plantas é que, ao se cruzar um parente silvestre com uma espécie cultivada, além da transferência de uma característica específica, inevitavelmente,

introduz-se na progênie uma série de características indesejáveis. Para a remoção de tais características, utiliza-se o retrocruzamento, processo muitas vezes demorado, e mesmo quando espécies compatíveis são utilizadas, barreiras de esterilidade estão presentes, devido a diferentes níveis de ploidia, incompatibilidades genômicas e diferenças genéticas (Holbrook & Stalker, 2003). O processo de introgressão apenas dos genes úteis pode ser acelerado através da utilização de marcadores moleculares.

4. Doenças em amendoim

O amendoim é bastante susceptível a uma grande variedade de pragas, incluindo-se os nematóides das galhas (*Meloidogyne* sp.) e doenças fúngicas de parte aérea (Subrahmanyam *et al.*, 1985; Pezzopane *et al.*, 1998).

As doenças fúngicas de parte aérea são consideradas um grande problema da cultura de amendoim no Brasil. Em escala global, as doenças foliares causadas pelos fungos são consideradas limitantes para a cultura, devido aos seus variados graus de severidade. Elas reduzem a área das folhas disponíveis para a fotossíntese, resultando em um decréscimo na habilidade para produzir e converter os produtos da fotossíntese, o que pode reduzir a produção das sementes (Nutter & Shokes, 1995).

No Brasil, os maiores problemas de natureza biótica são: a ferrugem, causada pelo fungo *Puccinia arachidis* Speg, e as manchas foliares *Cercospora arachidicola* Hori (Mancha castanha ou *Early leafspot*), *Cercosporidium personatum* (Berk. e Curt.) Deighton (Mancha preta ou *Late leafspot*). Todas estas doenças podem causar desfolhação e em combinação, podem causar perdas consideráveis na produção.

As cercosporioses são as doenças que mais causam prejuízos à cultura do amendoim, cujos danos podem ser superiores a 50% (Godoy *et al.*, 1994; Moraes *et al.*, 1994; Santos, 1995; Soares *et al.*, 1996).

Ao contrário de várias doenças do amendoim que ocorrem em regiões geográficas específicas, as cercosporioses estão presentes em quase todos os campos de cultivo do Brasil e no mundo. No entanto a predominância e a intensidade de cada uma variam com a localidade e as épocas de plantio. No Estado

de São Paulo, maior produtor do Brasil, a mancha preta tem predominado e se mostrado mais severa (Moraes *et al.*, 1988).

Com relação aos nematóides, estes são responsáveis por sérias perdas nas regiões basais da planta. Os danos causados por nematóides são, muitas vezes, agravados pela associação com outros patógenos, causando podridões em caules, ginóforos e vagens (Kokalis-Burelle & Rodrigues-Kabana, 1997).

Os fungos *Aspergillus flavus* Link e *A. parasiticus* Speare, são de grande importância devido à produção de aflatoxinas, quando infectam o amendoim. As aflatoxinas pouco afetam a produtividade da cultura, mas trazem problemas sérios à saúde humana e animal, influenciando significativamente na comercialização do produto (Bolonhezi, 2005). Um dos meios de se evitar a contaminação por aflatoxinas em amendoim é através do uso de cultivares resistentes. No Brasil, é comum a ocorrência de altos teores de aflatoxinas, devido à ineficiência de fiscalização e às condições ambientais desfavoráveis em que amendoim é produzido e armazenado (Godoy *et al.*, 2001).

O Brasil é constituído, em sua grande maioria, por pequenos produtores de amendoim, o que torna o controle químico, tanto de doenças fúngicas, como de outras pragas, caro e economicamente inviável. O uso de germoplasma resistente pode propiciar um melhor rendimento na produção, além de custos mais baixos. O gênero *Arachis* vem sendo muito estudado, pelo alto potencial de suas espécies silvestres como fonte de genes de resistência para o melhoramento da cultura. Como algumas espécies silvestres são caracterizadas por apresentarem resistências múltiplas a determinadas doenças, elas têm como maior potencial agrônomico e de grande impacto econômico a possibilidade de aumentar a resistência da espécie cultivada (Stalker, 1992). A espécie silvestre *A. stenosperma* Krapov e WC Gregory, pode ser considerada como uma importante fonte de genes para o melhoramento genético do amendoim. Entre as características de interesse que podem ser transferidas para o amendoim, estão o alto teor lipídico, e as resistências ao nematóide *M. arenaria* e a fungos foliares (Guimarães *et al.*, 2005; Leal-Bertioli *et al.*, 2005).

5. Genes de Resistência

Após ser infectada por um patógeno, a planta resistente disponibiliza um amplo espectro de defesa contra o microorganismo invasor. Uma defesa comum é a resposta de hipersensibilidade, na qual as células adjacentes ao local de infecção morrem rapidamente, privando o patógeno de nutrientes e impedindo a sua propagação. Esta resposta de hipersensibilidade pode ser a causa de outras respostas secundárias que, conjuntamente, são conhecidas como resistência sistêmica adquirida (SAR – *Systemic Acquired Resistance*). As bases genéticas das respostas de hipersensibilidade envolvem uma interação “gene-a-gene” (Flor, 1971), na qual o produto de um alelo da planta (um gene *R*, para resistência) interage com o produto de um alelo do patógeno (um gene *Avr* para avirulência). As proteínas codificadas por genes de resistência (genes *R*) são capazes de, direta ou indiretamente, reconhecer proteínas elicitoras produzidas pelos patógenos e ativar respostas de defesa (Taken & Joosten, 2000; Hammond-Kosack & Jones, 2000; Martin *et al.*, 2003; Hammond-Kosack & Parker, 2003).

A caracterização molecular dos genes de resistência, em mono e dicotiledôneas, revelou que estes genes codificam proteínas com motivos estruturais comuns, indicando que o modo de reconhecimento de diferentes patógenos e a sinalização de defesa da planta foi conservado durante sua evolução e diversificação. Até hoje, cinco principais famílias de genes R já foram identificadas, baseados nos domínios conservados das proteínas (Martin *et al.*, 2003). A família mais abundante é denominada NBS-LRR (Rommens & Kishore, 2000). A maioria destas proteínas contém domínios C-terminais com repetições ricas em aminoácido leucina (LRR – *Leucine-Rich Repeat*), um domínio central com sítios de ligação a nucleotídeos trifosfatados (NBS – *Nucleotide Binding Site*) e domínios N-terminais variáveis (Figura 2). Esta família é largamente conhecida, e inclui mais de 20 genes R identificados em várias espécies de plantas (Meyers *et al.*, 1999; Van der biezen *et al.*, 2002; Cordero & Skinner, 2002). Genes desta família conferem resistência a diversos patógenos como fungos, bactérias, viroses, nematóides e insetos (Baker *et al.*, 1997).

O domínio LRR está envolvido no reconhecimento do patógeno pelo hospedeiro durante o processo de infecção. Proteínas que contém este domínio

apresentam o aminoácido leucina, repetida em número de 10 a 40 intervalos regulares, podendo conter ainda outros resíduos hidrofóbicos, asparaginas e prolinas (Collins *et al.*, 2001; Cannon *et al.*, 2002). Esses motivos geram uma estrutura terciária semelhante a uma mola, com cada hélice apresentando um motivo de leucina repetido. A função do domínio LRR pode estar relacionada com a mediação da interação entre proteínas, podendo atuar extracelularmente como receptor da molécula produzida pelo patógeno ou intracelularmente, em um passo anterior à via de transdução dos sinais envolvidos no processo de resistência (Banerjee *et al.*, 2001).

O domínio NBS está presente em diversas proteínas e são responsáveis por atividade de ligação ao ATP ou GTP, tais como subunidades da ATP sintetase e fatores de alongação do ribossomo. A presença desses domínios nas proteínas codificadas pelos genes de resistência sugere que estes sítios estão envolvidos na ativação de proteínas quinases ou proteínas transportadoras (Hammon-Kosack & Jones, 1997) e mutações em sua estrutura, impedem que ocorra a resposta de hipersensibilidade, ressaltando a importância da função deste domínio na sinalização (Zhou *et al.*, 2004).

A família NBS-LRR é dividida 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 Não-TIR, possui o domínio CC (*Coiled-coil*, Cauda espiralada) na região N-terminal e é responsável pelo reconhecimento das moléculas elicitórias (Young *et al.*, 2000). O segundo grupo, TIR, é composto por genes que codificam produtos contendo uma estrutura homóloga às proteínas receptoras humanas “Interleucina 1” e “*Toll*” de *Drosófila*. As espécies do gênero *Arachis* possuem as duas formas da estrutura (Bertioli *et al.*, 2003).

O domínio LRR pode sofrer várias mutações em seus aminoácidos representando uma contribuição para a variabilidade genética, necessária para reconhecer as diversas especificidades patogênicas. A região NBS por outro lado, apresenta motivos conservados (Kobe & Deisenhofer, 1994). Estas regiões conservadas podem facilitar a construção de *primers* para amplificação de RGAs (*Resistance Gene Analogs*) ou genes similares. Alguns desses motivos são: P-loop, Kinase-2, Kinase-3a, GLPL e RNBS-D.

Os RGAs são sequências análogas às sequências codificadas por genes R, mas sem função de resistência ainda estabelecida. Alguns *primers* degenerados baseados nos domínios conservados de genes de resistência têm sido utilizados para amplificação e clonagem de RGAs em diferentes espécies de plantas. Alguns estudos de mapeamento de RGAs no genoma de plantas comprovaram ligação com locos que contêm genes R conhecidos, como em soja, cevada, alface e milho (Kanazin *et al.*, 1996; Yu *et al.*, 1996; Collins *et al.*, 2001; Brueggeman *et al.*, 2001; Shen *et al.*, 1998; Collins *et al.*, 1998).

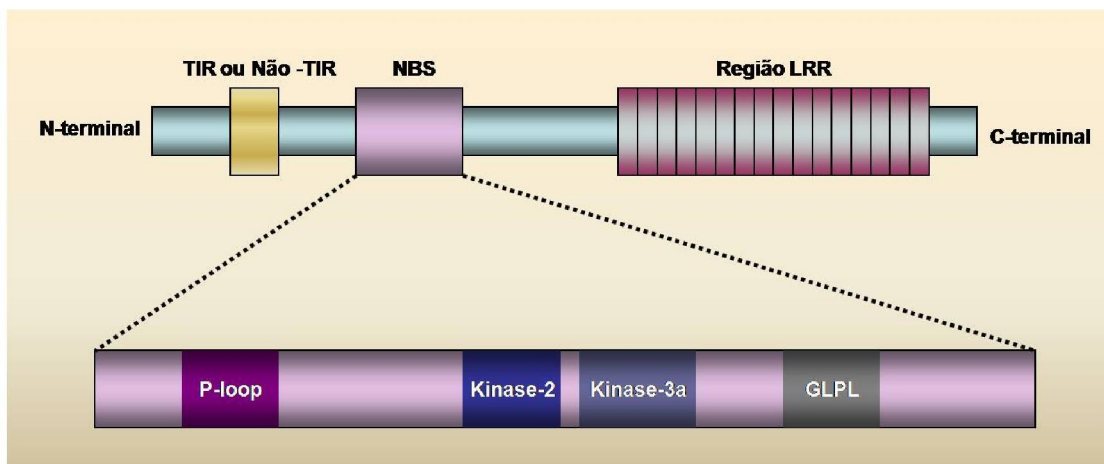


Figura 2: Domínio NBS-LRR representado esquematicamente. As regiões TIR e Não-TIR também estão ilustradas. Os motivos P-loop, Kinase-2, Kinase-3a e GLPL estão representados em destaque na parte inferior da figura.

A mais expressiva caracterização de genes ligados à resistência (RGAs) no gênero *Arachis* foi conduzida por Bertoli *et al.* (2003). O estudo gerou 78 RGAs a partir de *primers* degenerados baseados no domínio NBS da espécie *A. hypogaea* var. Tatu e quatro espécies silvestres (*A. duranensis*, *A. cardenasii*, *A. stenosperma* e *A. simpsonii*). ORFs (*Open Reading Frames*) indicaram a proximidade com RGAs de outras espécies de leguminosas incluindo *Glycine max*, *Lotus japonicus*, *Medicago truncatula* e *Phaseolus vulgaris* (citado por Yuksel *et al.*, 2005). Neste trabalho, realizado por Yuksel e colaboradores (2005) um total de 234 novos RGAs foi isolado de *A. hypogaea* a partir de *primers* desenhados para regiões conservadas no domínio NBS.

6. Marcadores moleculares, mapas genéticos e QTLs

Mapas genéticos têm sido desenvolvidos desde 1980 (Botstein, 1980). O desenvolvimento de mapas genéticos é considerado uma das aplicações de maior impacto da tecnologia de marcadores moleculares na análise genética de espécies e, potencialmente, no melhoramento de plantas. O mapeamento genético possibilita a cobertura e análise completa de genomas, a decomposição de características genéticas nos seus componentes mendelianos, a localização de regiões genômicas que controlam caracteres de importância econômica (QTL - *Quantitative Trait Loc*), a quantificação dos efeitos dessas regiões na característica estudada e a canalização de toda essa informação para uso em programas de melhoramento (Lee, 1995).

Há vários tipos de marcadores moleculares utilizados para estudos genéticos, como Isoenzimas, RFLP, RAPD, AFLP, SCAR, SSR além de outros. A grande disponibilidade de marcadores e metodologias estatísticas eficientes fez com que, atualmente, estejam disponíveis mapas genéticos para um grande número de espécies cultivadas.

O mapeamento genético toma como base a hipótese de que a co-segregação de dois marcadores indica proximidade entre eles, e que a probabilidade de ocorrer permutas genéticas entre esses dois marcadores é menor, quanto menor for a distância entre eles. Dessa forma, é possível ordenar linearmente a informação genética ao longo dos cromossomos. A probabilidade de ocorrência de permuta genética entre dois marcadores (frequência de recombinação) reflete a distância genética entre locos. Esta distância, expressa em centiMorgans (cM), é calculada com base na frequência de recombinação por meio de funções de mapeamento, as quais corrigem as distorções entre a conversão das unidades. A função Haldane é a mais simples e admite que permutas genéticas ocorram ao acaso e de forma independente (Haldane, 1919). A função de Kosambi considera interferência parcial (I) nos cálculos da distância em centiMorgans, sendo que essa interferência significa que uma permuta genética afeta a ocorrência de outras em regiões adjacentes, considerando a ocorrência de permutas duplas (Kosambi, 1944).

A identificação de marcadores moleculares ligados a genes e QTLs de resistência pode ser de grande utilidade para o melhoramento de uma espécie. A

primeira etapa da construção de um mapa genético contendo marcadores associados a QTLs de resistência é a escolha correta dos genitores para cruzamento e obtenção de uma geração segregante. Esses genitores devem ser divergentes para o caráter. A próxima etapa é a detecção de marcadores moleculares polimórficos entre os genitores resistente e suscetível, analisando-se posteriormente as segregações destes e da resistência em uma população segregante.

De acordo com Tanksley *et al.* (1993), os melhoristas podem utilizar marcadores moleculares para selecionar indiretamente indivíduos em uma população segregante que carrega um gene de interesse, se existir uma ligação forte entre o loco marcador e o gene que controla o caráter. Marcadores fortemente ligados a genes de resistência podem auxiliar muito os programas de melhoramento visando resistência a doenças, permitindo seguir o gene sob seleção através das gerações, ao invés da avaliação da expressão fenotípica do gene de resistência. Esses marcadores podem ser utilizados no processo de seleção assistida por marcadores (SAM), permitindo também que vários genes sejam agrupados em um mesmo genótipo (piramidação dos genes de resistência). Adicionalmente, os marcadores podem ser úteis para a clonagem e sequenciamento desses genes de resistência em associação a outras ferramentas moleculares, como as bibliotecas genômicas.

A limitação da diversidade genética no amendoim, como observado em diferentes estudos utilizando marcadores moleculares (Kochert *et al.*, 1996; Subramaniam *et al.*, 2000; Milla *et al.*, 2005), tem dificultado o mapeamento genético e estudos com QTL, com isso, poucos mapas de ligação têm sido publicados.

Em 1993, Halward e colaboradores desenvolveram um dos primeiros mapas genéticos de amendoim, baseado na técnica de RFLP, a partir de uma população F_2 , derivada de um cruzamento entre duas espécies silvestres diplóides com genoma AA, *A. stenosperma* e *A. cardenasii*. Burow *et al.* (2001) desenvolveram outro mapa baseado em RFLP, porém para uma população tetraplóide. Mais tarde foi desenvolvido um mapa derivado de um cruzamento entre duas espécies silvestres diplóides *A. stenosperma* e *A. cardenasii* baseado em RAPD (Garcia *et al.*, 2005).

Ainda em 2005, Moretzsohn e colaboradores publicaram um mapa de ligação baseado em uma população de 93 indivíduos F_2 obtidos do cruzamento entre duas

espécies silvestres de genoma AA, *A. duranensis* e *A. stenosperma*. Um total de 170 marcadores microssatélites foi mapeado em 11 grupos de ligação cobrindo uma distância de 1.230,89 cM no mapa e com uma distância média de 7.24 cM entre marcadores adjacentes. Este mapa vem sendo aprimorado e, atualmente possui aproximadamente 188 SSRs, 80 marcadores âncoras, 35 RGAs e 10 grupos de ligação (Leal-Bertioli *et al.*, 2009). Recentemente um mapa para o genoma BB, baseado em um cruzamento entre *A. ipaënsis* e *A. magna*, foi publicado por Moretzsohn *et al.*, (2009). Esse mapa contém 10 grupos de ligação, com 149 locos cobrindo uma distância de mapa total de 1.294 cM.

QTLs para resistência à mancha foliar causada pelo fungo *Cercosporidium personatum* foram mapeados em um cruzamento de *A. duranensis* e *A. stenosperma* (Leal-Bertioli *et al.* 2009). Alguns desses QTLs foram localizados em regiões do genoma onde mapearam genes candidatos à resistência a doenças, na região superior do grupo de ligação 4, assim como na região inferior do grupo de ligação 2, indicando que estas regiões podem estar relacionadas ao controle de resistência a doenças.

Recentemente foi publicado um mapa genético para *A. hypogaea* (Varshney *et al.*, 2009) incluindo marcadores SSRs e marcadores âncoras (derivados de Leguminosas) já existentes e mapeados em germoplasma silvestre.

7. Bibliotecas BAC

Bibliotecas de grandes insertos de DNA genômico se tornaram uma ferramenta básica para mapeamento físico, clonagem posicional e sequenciamento genômico (Zhang *et al.*, 1994). Atualmente, o uso do vetor denominado cromossomo artificial de bactéria (BAC - *Bacterial Artificial Chromosome*) tem sido preferido na construção dessas bibliotecas por sua eficácia de transformação, baixo índice de quimerismo além da possibilidade de uso em subclonagens por *shotgun* para sequenciamento (Safár *et al.*, 2004).

A estabilidade de clones BAC é em parte devido à presença do fator F (par A e par B) que impede que mais de um BAC habite simultaneamente uma mesma bactéria (Yuksel & Paterson, 2005). Uma vantagem adicional na utilização de clones BAC é que estes são relativamente fáceis de manipular e propagar, comparados

com vetores virais ou leveduras. Conseqüentemente, BACs têm suplantado YACs (*Yeast Artificial Chromosome*) como os vetores mais comuns utilizados em mapeamento físico em larga escala e sequenciamento (Kelley *et al.*, 1999).

Bibliotecas BAC têm sido construídas para várias culturas economicamente importantes como, por exemplo, *Oryza sativa* (Tao *et al.*, 1994), *Phaseolus vulgaris* (Vanhouten & MacKenzie, 1999), *Vigna radiata* (Miyagi *et al.*, 2004), *Glycine max* (Danesh *et al.*, 1998), *Trifolium pretense* (Sato *et al.*, 2005) e *Medicago truncatula* (Nam *et al.*, 1999). Dentro do gênero *Arachis* uma biblioteca BAC foi construída para a espécie cultivada alotetraplóide (Yuksel & Paterson, 2005), e duas para as espécies silvestres *A. duranensis* (84.096 clones) e *A. ipaensis* (75.648 clones) (Guimarães *et al.*, 2008). Estas últimas bibliotecas foram construídas a partir do vetor de clonagem pIndigoBAC-5, que possui um tamanho de 7.5 Kb (Figura 3) e o tamanho médio dos insertos clonados em ambas as bibliotecas variaram entre 100 ou 110 Kb.

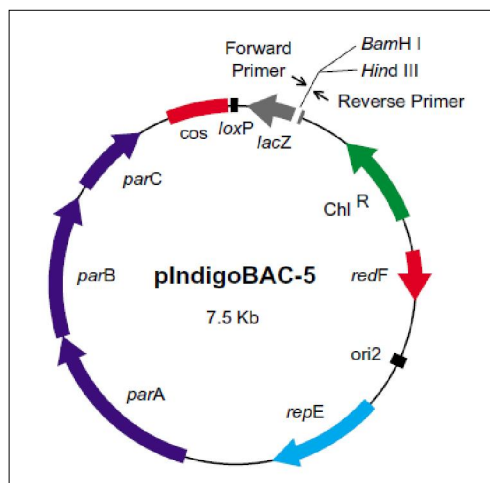


Figura 3: Vetor de clonagem pIndigoBAC-5.

A busca de marcadores ligados a genes de importância agrônômica tem sido objeto de pesquisa nos últimos anos (Hash *et al.*, 2003). O isolamento de genes de resistência ainda não caracterizados pode ser facilitado com a utilização de RGAs. Regiões de resistência têm sido mapeadas utilizando RGAs como sondas de hibridização em Southern Blotting em progênies segregantes de diferentes

parentais. Em um trabalho realizado por Clément *et al.*, (2004), uma biblioteca BAC para a espécie *Theobroma cacao* L. foi construída e um grupo de RGAs previamente mapeados que co-segregam com um QTL para resistência ao patógenos do gênero *Phytophthora* spp, foi utilizado para realizar um screening através de hibridização nesta biblioteca. Mais tarde, Lein *et al.* (2007), identificou RGAs e os agrupou no grupo de ligação 3 em beterraba. Estes co-segregam com o QTL para resistência a doença rizomania, uma das mais importantes em beterraba.

O mapeamento de RGAs demonstrando que eles estão ligados a locos conhecidos por conterem genes de resistência já foi citado também em outras culturas, como batata (Leister *et al.*, 1996), soja (Kanazin *et al.*, 1996), *A. thaliana* (Aarts *et al.*, 1998) e ervilha (Timmerman-Vaughan *et al.*, 2000). Yuksel e colaboradores (2005) exploraram a diversidade em relação aos genes de resistência em amendoim, utilizando grupos de *primers* previamente estudados em outras taxa, identificando e caracterizando clones BAC de *A. hypogaea*.

QTLs para resistência à mancha foliar causada pelo fungo *C. personatum* foram mapeados em um mapa genético construído para um cruzamento de *A. duranensis* e *A. stenosperma* (Leal-Bertioli *et al.*, 2009). O RGA S1_A_36 que co-segrega com um destes QTLs (Figura 1 – Capítulo I) foi utilizado para identificar clones positivos nas bibliotecas BAC destas espécies silvestres do gênero *Arachis* (Proite, 2007). A subclonagem e sequenciamento de um clone BAC (positivo) que contém este RGA possibilitarão a caracterização da sequência, além da identificação de microssatélites na região designada, provendo marcadores mais próximos ao QTL já identificado e ao isolamento do gene de resistência de interesse.

8. Subclonagem

A metodologia mais informativa para estudos de estrutura do genoma é a análise de sequência de DNA, entretanto, embora o sequenciamento tenha se tornado significativamente mais rápido e econômico, a montagem de genomas completos de eucariotos ainda constitui uma tarefa complexa. Uma opção estratégica interessante e informativa para o estudo de estruturas de genoma é o sequenciamento por amostragem (*sample sequencing*) (Elgar *et al.*, 1999). Nesta estratégia, todo o DNA genômico, ou alguns trechos deste, são sonicados,

nebulizados ou cortados com enzimas de restrição (*shotgun*) e os fragmentos resultantes são subclonados em vetores bacterianos, sequenciados e montados para posterior análise. Da mesma forma, clones BAC contendo regiões de interesse podem ser selecionados e sequenciados possibilitando o acesso à sequência completa desses genes e suas regiões vicinais.

OBJETIVO GERAL

Analisar a estrutura genômica da região que contém o RGA S1_A_36, que se encontra co-localizado com o QTL cp4.2 (resistência ao fungo *Cercosporidium personatum*) mapeado no Grupo de Ligação 4 do mapa genético de referência, através do sequenciamento total do clone BAC que contém esse RGA de interesse e desenvolver marcadores moleculares do tipo microssatélites, com intuito de mapear marcadores associados a genes de resistência que sejam transferíveis entre as espécies relacionadas.

OBJETIVOS ESPECÍFICOS

- Subclonar e sequenciar o clone BAC AD25F09 contendo o RGA S1_A_36;
- Realizar a anotação e interpretação da sua sequência;
- Identificar microssatélites nas sequências geradas e desenvolver pares de *primers* para mapeamento destas regiões;
- Analisar os pares de *primers* para microssatélites quanto ao polimorfismo nos parentais das populações AA de mapeamento (*A. stenosperma* e *A. duranensis*);
- Genotipar a população F₂ segregante com os microssatélites polimórficos gerados;
- Posicionar os marcadores no mapa genético de referência;

APRESENTAÇÃO DOS CAPÍTULOS

O presente trabalho foi dividido em dois capítulos. O capítulo I consiste na subclonagem, sequenciamento e anotação do BAC AD25F09. O capítulo II apresenta o desenvolvimento de marcadores microssatélites a partir das informações do BAC sequenciado e sua utilização na genotipagem e mapeamento da população F₂ (genoma AA).

**CAPÍTULO I – SUBCLONAGEM, SEQUENCIAMENTO E ANOTAÇÃO DO BAC
AD25F09**

1. INTRODUÇÃO

O amendoim está entre as principais oleaginosas cultivadas no mundo e as maiores causas de perda na produção são a seca e fungos causadores de manchas foliares como as cercosporioses (Santos, 1995). Embora estejam disponíveis acessos de *A. hypogaea* com diferentes níveis de resistência, o melhoramento das cultivares é dificultado devido à baixa variabilidade genética em relação a características agrônômicas de interesse. Por outro lado, espécies silvestres do gênero *Arachis* possuem resistência a diferentes fatores bióticos e abióticos apresentando uma maior variabilidade genética (Hilu & Stalker, 1995; Moretzsohn *et al.*, 2004; Milla *et al.*, 2005).

Por constituírem importante fonte de genes úteis e apresentarem alto polimorfismo, espécies silvestres vêm sendo introduzidas em programas de melhoramento. O processo de introgressão de genes pode ser acelerado pelo uso da seleção de plantas assistida por marcadores moleculares e para tal, é necessário identificar marcadores proximamente ligados a genes de interesse como resistência a doenças (RGA). Muitos destes genes são caracterizados pela presença do domínio NBS-LRR que, em plantas, tem sua função associada à resistência a doenças.

Bibliotecas genômicas de grandes insertos, como as construídas em BAC, são ferramentas importantes para o isolamento de genes, determinação de agrupamentos gênicos, estudos de sintenia entre espécies, desenvolvimento de marcadores moleculares e na caracterização de genomas de uma maneira geral. Mais informações sobre o genoma diplóide de *Arachis*, através de estudos relacionados às bibliotecas BAC, seriam muito úteis, tanto para a construção de mapas genéticos e físicos, como para caracterização de alelos silvestres.

Primers degenerados para o motivo NBS foram produzidos para diferentes espécies de *Arachis* visando à identificação de sequências do tipo RGA (Bertioli *et al.*, 2003). Alguns desses RGAs foram posicionados em um mapa genético construído para uma população de genoma AA (*A. duranensis* e *A. stenosperma*) (Moretzsohn *et al.*, 2005). QTLs para resistência à mancha foliar causada pelo fungo *C. personatum* foram posicionados no mesmo mapa, sendo que o RGA S1_A_36

apresentou co-segregação com um destes QTLs, situado no Grupo de Ligação 4 do mapa de referência (Figura 1). A sequência deste RGA foi utilizada como sonda para identificar clones positivos nas bibliotecas BAC das espécies silvestres parentais *A. duranensis* (genoma AA) e *A. ipaënsis* (genoma BB) (Proite, 2007). Um clone positivo foi encontrado na biblioteca BAC da espécie silvestre *A. duranensis* e em virtude de sua localização na biblioteca, foi denominado AD25F09 (AD=*Arachis duranensis*, placa 25, posição F09). Este BAC será subclonado e sequenciado, permitindo a reconstituição deste.

Grupo de Ligação 4

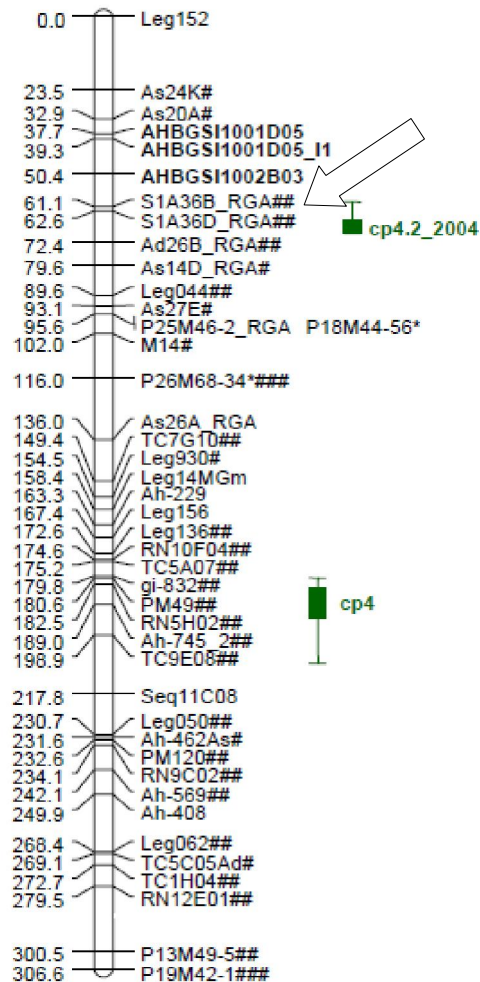


Figura 1: Grupo de Ligação 4 do mapa genético de *Arachis* (Moretzsohn *et al.*, 2005). A seta mostra a posição do RGA S1_A_36. Em verde estão indicados dois QTLs para resistência ao fungo *C. personatum*.

1.1 Sequenciamento através de *shotgun*

Os sequenciadores automáticos de DNA (ABI 377 e ABI 3100, por exemplo), que utilizam o método de sequenciamento pela reação de terminação em cadeia (Sanger *et al.*, 1977), têm limitações técnicas que os impedem de sequenciar regiões muito grandes, estando a cobertura média entre 500 e 800 pb apenas. Desta forma, para viabilizar o sequenciamento completo de um genoma, ou de um BAC, por exemplo, deve-se primeiramente tratar o DNA de forma a criar inúmeros fragmentos. Esses fragmentos devem ser individualmente sequenciados e, posteriormente, montados com o auxílio da bioinformática, como peças de um quebra-cabeça (Luscombe *et al.*, 2001). Este sequenciamento individualizado torna-se viável, pois cada fragmento da sequência é subclonado em um vetor bacteriano.

O sequenciamento aleatório, aliado a uma estratégia de clonagem randômica (*shotgun*) é interessante e possui um relevante poder informativo para gerar um panorama geral do genoma de um organismo de interesse. Nesta estratégia, como demonstrado na Figura 2, todo o DNA genômico, ou alguns trechos deste, como BACs, são sonicados, nebulizados ou cortados com enzimas de restrição, e os fragmentos resultantes são subclonados e sequenciados. As sequências de nucleotídeos dos fragmentos são então montadas e analisadas, com o intuito de descrever a estrutura e composição nucleotídica. Esta estratégia contrasta com o sequenciamento parcial baseado exclusivamente em genes expressos (ESTs – *Expressed Sequence Tags*), pois não busca apenas as sequências expressas dos organismos, e sim uma visão da estrutura do genoma como um todo ou parte dele. Dentre as informações provenientes do sequenciamento aleatório pode-se citar a distribuição de microssatélites, elementos transponíveis e retrotransponíveis, genes, pseudogenes, regiões promotoras, conteúdo nucleotídico, tRNAs, e outros. Esta estratégia foi utilizada inicialmente no estudo de genomas de microorganismo que não possuíam mapas físicos (Wong *et al.*, 1999; Bell *et al.*, 2002) e constituiu a base da abordagem, hoje comum, de sequenciamento por fragmentação randômica de DNA (*shotgun*) para montagem completa de genomas de alguns microrganismos, utilizada de forma pioneira para *Haemophilus* (Fleischmann *et al.*, 1995).

A disponibilidade das sequências é um avanço significativo, mas por si só não é garantia para que se possa entender o funcionamento do organismo, mesmo no

caso dos genomas completos. Isto se deve, primariamente, ao fato dos sistemas biológicos serem extremamente complexos, onde cada gene participa de uma rede intrincada de interações. Para determinar a função de um gene, faz-se necessário uma série de procedimentos experimentais, que, em sua maioria, são trabalhosos e demorados. Por isso, e também pelo fato de que as sequências acumuladas sem interpretação não terem muito valor, a anotação genômica ganhou grande importância nos últimos anos.

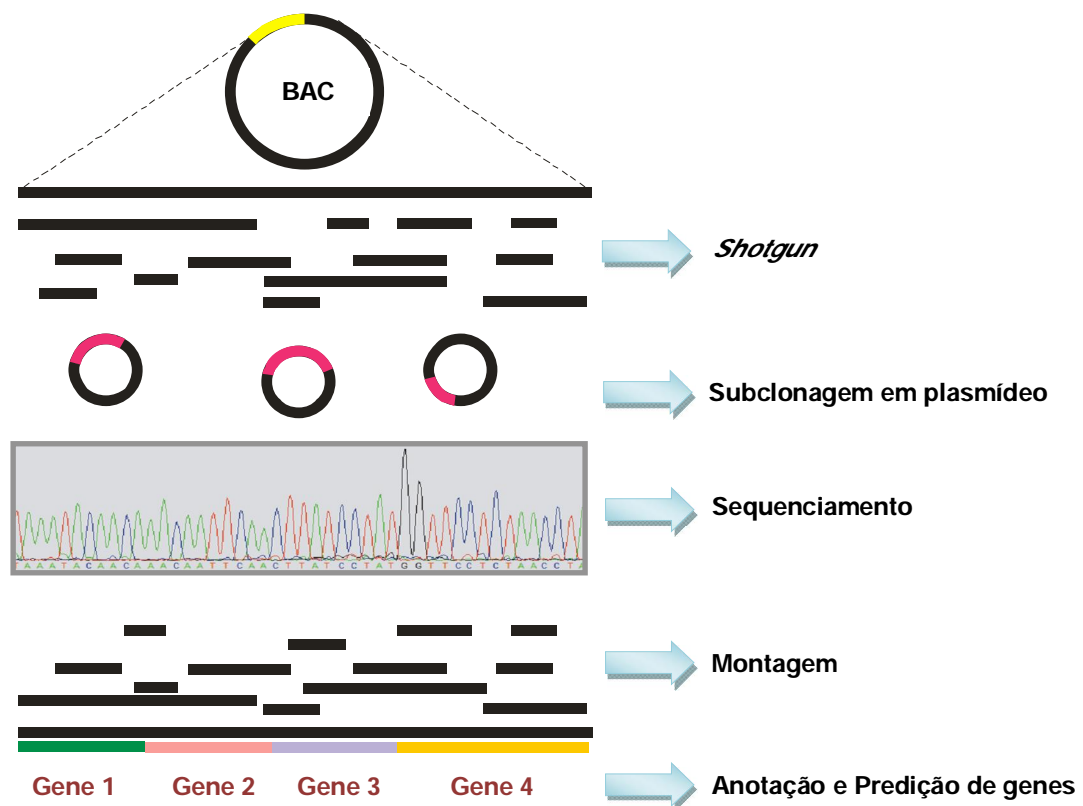


Figura 2: Método de subclonagem de BAC através da técnica de shotgun.

1.2 Anotação de sequências

A anotação genômica, extração de significado biológico de sequências genômicas anônimas, é tarefa bastante complicada, e diversos esforços foram feitos para sua automação (Aubourg & Rouzé, 2001). Este processo consiste em adicionar análises de bioinformática para gerar interpretações biológicas sobre as sequências brutas de DNA. A predição da função de uma proteína é feita com base na comparação com sequências de proteínas previamente caracterizadas e armazenadas em bancos de dados biológicos.

A identificação de genes a partir da comparação de sequências buscando regiões de similaridade já é uma ferramenta amplamente utilizada. Para isto, em geral, pesquisadores utilizam o algoritmo *Basic Local Alignment Sequence Tool* (BLAST) (Altschul *et al.*, 1997). A partir do alinhamento de sequências, identificação das regiões alinhadas e pontuação destas a partir de uma matriz padrão, pode-se descobrir sequências similares e com diversos graus de identidade, permitindo qualificar sequências não-idênticas e descrever possíveis funções (Pertsemlidis & Fondon, 2001).

Outra estratégia bastante útil é a comparação das sequências anônimas com bancos de ESTs e bancos de GSS (*Genome Survey Sequence*), de preferência da mesma espécie, indicando dentre outros fatores genes putativos nas sequências anônimas. A função destes genes seria atribuída diretamente de acordo com a função putativa do EST semelhante. Atualmente, a taxa de crescimento dos bancos de dados biológicos é exponencial, tornando a atividade de anotação funcional dinâmica.

2. MATERIAL E MÉTODOS

2.1 Plantas de *Arachis* e bibliotecas BAC

Os experimentos foram conduzidos no Laboratório de Biotecnologia (Pós-graduação da Universidade Católica de Brasília) e Embrapa Recursos Genéticos e Biotecnologia. As plantas utilizadas neste trabalho foram fornecidas pelo Banco Ativo de Germoplasma (BAG) de Espécies de *Arachis* situado na Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF. Os clones BAC utilizados foram selecionados a partir de biblioteca BAC construídas no CIRAD (Montpellier – França) (Guimarães *et al.*, 2008).

2.2 Isolamento de RGAs das bibliotecas BAC de espécies de *Arachis ipaënsis* e *A. duranensis*

O RGA S1_A_36 (GenBank ref. AY157808; Bertoli *et al.* 2003) localizado no GL 4 do mapa genético AA, construído a partir do cruzamento das espécies silvestres de *A. duranensis* e *A. stenosperma*, foi selecionado por apresentar co-segregação com um QTL para resistência à mancha foliar causada pelo fungo *C. personatum* (Leal Bertoli *et al.*, 2009). Esta sequência foi utilizada como sonda para identificar clones positivos nas bibliotecas BAC de espécies silvestres do gênero *Arachis* através de hibridizações de colônias (<http://www.genome.clemson.edu/protocols>) (Guimarães *et al.*, 2008). O clone selecionado foi denominado AD25F09 situado na placa 25 posição F09 na biblioteca AA de *A. duranensis*. A subclonagem do clone BAC selecionado foi realizada através de dois métodos distintos: com enzimas de restrição e através da técnica de nebulização, com o intuito de gerar um maior número de sequências para análise e posterior geração de marcadores microsatélites.

O esquema do método utilizado para o sequenciamento do BAC AD25F09, está descrito na Figura 3.

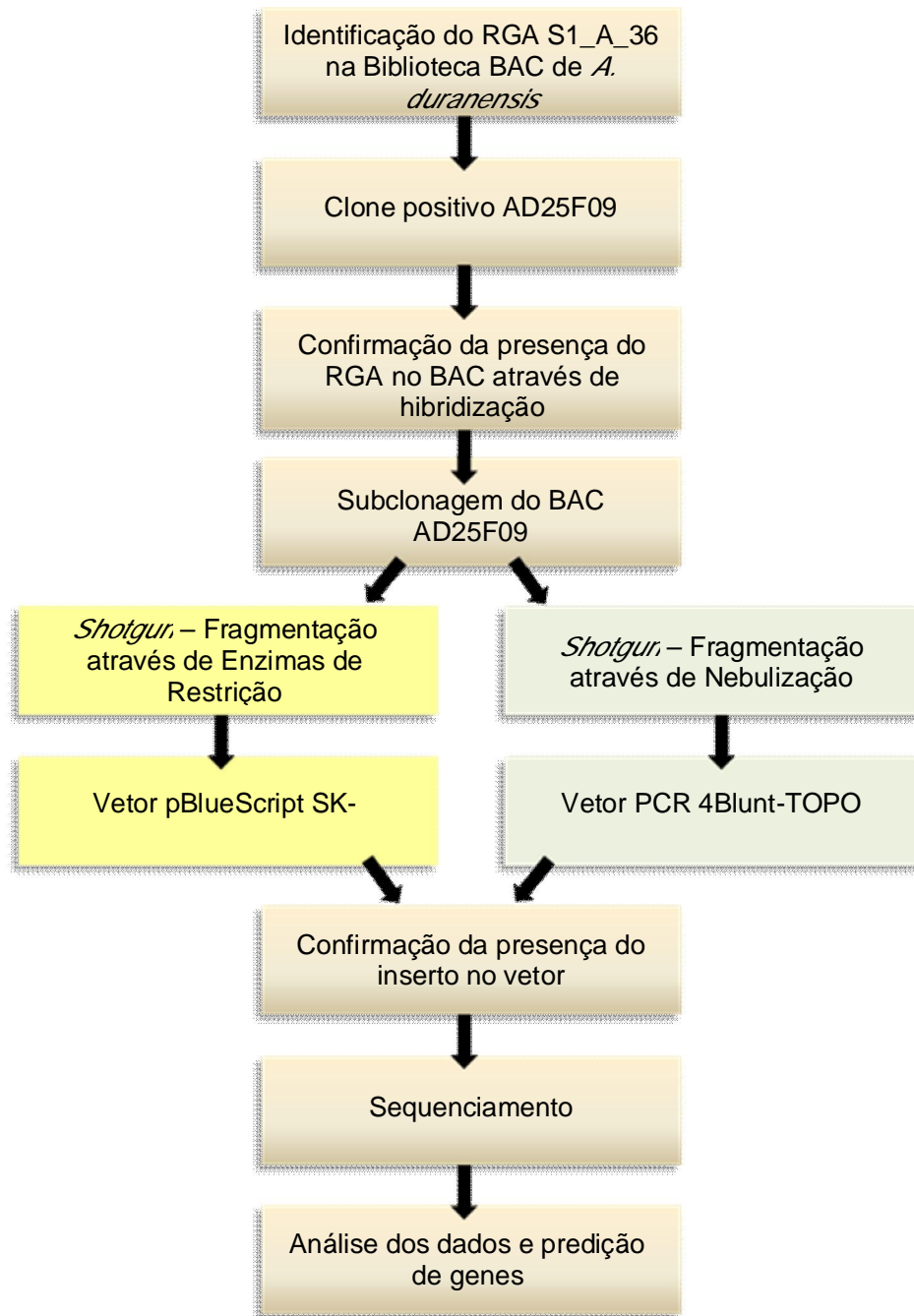


Figura 3: Esquema de sequenciamento do BAC AD25F09.

2.3 Extração do vetor BAC a partir dos clones selecionados das bibliotecas

A partir de uma cultura permanente do clone BAC armazenada em freezer -80° C realizou-se um pré-inóculo: Foram utilizados 3 mL de meio LB (Lúria Bertani – 1% de Triptona; 0,5% de Extrato de Levedura; 1% de NaCl) e 3 µL do antibiótico cloranfenicol (12,5 mg/µL). A incubação foi realizada por 12 horas a 37° C com rotação em torno de 225 rpm. Para a confecção do inóculo, utilizou-se 3 µL da pré-cultura em um tubo contendo 3 mL de meio LB e 3 µL do antibiótico cloranfenicol (12,5 mg/µL). A incubação foi realizada durante 16 horas a 37° C com rotação em torno de 225 rpm.

Para extração do BAC, utilizou-se o Protocolo fornecido pelo Doutor Pierre Lagoda (CIRAD - França) através de comunicação pessoal, com modificações realizadas pelo Doutor Stephan Nielen (Embrapa Recursos Genéticos e Biotecnologia – Brasília). As células (cepa de bactérias DH10B utilizada para a confecção da biblioteca BAC) foram centrifugadas a 13.000 rpm por 1 minuto. O sobrenadante foi descartado e o *pellet* ressuspensionado em 200 µL de tampão a 4° C (glicose 50 mM; Tris-HCl 1M pH 8; EDTA 500 mM; 100 µg de RNase A). A solução foi homogeneizada, acrescida de 200 µL de tampão de lise (NaOH 200 mM; SDS 1%) e misturada gentilmente por 2 minutos à temperatura ambiente. Foram adicionados 200 µL de tampão de neutralização (acetato de potássio 1,32 M, pH 4,8) e após 3 minutos no gelo, os resíduos foram removidos por centrifugação a 13.000 rpm por 15 minutos (4° C) e o sobrenadante transferido para novo tubo. Foram adicionados 400 µL de isopropanol, misturando-se os tubos gentilmente. As amostras foram então centrifugadas a 13.000 rpm por 20 minutos e o sobrenadante descartado. O *pellet* foi ressuspensionado em 100 µL de TE (Tris-HCl 10 mM; EDTA 1 mM), precipitado com 10 µL de acetato de sódio 3 M (4° C) e 250 µL de etanol absoluto (-20° C) e mantidos em freezer -80° C durante 1 hora seguida de centrifugação por 20 minutos a 4° C. O sobrenadante foi descartado e o *pellet* foi lavado com etanol 70% gelado e seco completamente.

O DNA foi ressuspensionado em 50 µL de água *MililiQ* autoclavada e incubado por 10 minutos a 65° C, antes de ser estocado a -20° C.

2.4 Confirmação da presença do RGA no BAC AD25F009

Após a extração do clone BAC, a confirmação da presença do RGA foi realizada através de hibridização pela técnica de *Southern blot* (Southern, 1975). Cerca de 5 µg de BAC foram digeridos com 50 U da enzima de restrição HindIII em um volume final de 150 µL. A amostra foi separada em gel de agarose 1% a 20 V por aproximadamente 16 horas. Após a separação dos fragmentos, o gel foi tratado, adicionando-se a solução de depurinação por 15 minutos, sob agitação leve e constante e, posteriormente, lavado com água destilada. Após a lavagem, adicionou-se a solução de desnaturação, sob agitação durante 30 minutos e lavando novamente com água destilada. Adicionou-se a solução de neutralização sob agitação por 30 minutos. Os fragmentos foram transferidos para membrana Hybond-N⁺ por capilaridade em SSC 20X por aproximadamente 24 horas.

Para a pré-hibridização, a membrana foi colocada em uma garrafa de vidro contendo solução de pré-hibridização a 65° C com agitação lenta por aproximadamente 5 horas. A sonda foi amplificada por PCR do clone *A. stenosperma* (selecionado por amplificar o RGA S1_A_36) e cerca de 50 ng foram utilizados para a marcação utilizando-se o Kit Ready-to-Go d-CTP (Amersham, Pharmacia) juntamente com 5 µL de ³²P. A sonda foi incubada por 30 minutos a 37° C e adicionada à garrafa contendo a membrana com a solução de Pré-hibridização. A hibridização ocorreu a 65° C por 24 horas e logo após foi realizada a lavagem (uma lavagem com Solução I (Soluções em anexo) a 65° C por 15 minutos, duas lavagens com Solução II por 15 minutos). A membrana foi seca e colocada para sensibilizar filme radiográfico (Kodak-Bio).

2.5 Estratégias para subclonagem

Para realizar a subclonagem do BAC AD25F09, foram utilizadas duas estratégias distintas para a fragmentação do DNA. A primeira consistiu na utilização de três conjuntos de enzimas de restrição para uma geração de diferentes perfis, o que ajudaria na sobreposição das sequências. A segunda estratégia de fragmentação foi realizada através do processo de nebulização.

2.5.1 Estratégia de subclonagem 1 – Fragmentação através de Enzimas de Restrição

Foram utilizadas três enzimas de restrição (*Hind*III, *Bam*HI, *Sau*BI) para a geração de fragmentos a partir do BAC. Estas enzimas de restrição do Tipo II clivam o DNA no mesmo sítio do seu reconhecimento, que é uma sequência palindrômica, isto é, ela tem um eixo de simetria e a sequência de bases de uma fita é a mesma da fita complementar, quando lida na direção oposta.

a) Digestão do BAC

O BAC (aproximadamente 10 µg) foi digerido com 2U de enzima de restrição (*Hind*III, *Bam*HI e *Sau*BI (Invitrogen) em três experimentos distintos) de acordo com o protocolo do fabricante e incubado a 37° C por 3 horas. Após a digestão o DNA foi precipitado com 1/10 do volume de acetato de sódio a 3 M, 2,5 x o volume de etanol absoluto e incubado por 1 hora em freezer -80° C. A solução foi centrifugada a 13.000 rpm por 20 minutos e o sobrenadante descartado. Em seguida, o DNA foi ressuspenso em água *MillicQ* autoclavada.

b) Digestão do vetor de subclonagem e desfosforilação

O vetor utilizado foi o pBlueScript SK- (pBS SK-). Para a linearização do vetor utilizou-se aproximadamente 20 µg de vetor; 1 U de enzima de restrição (*Hind*III ou *Bam*HI) (Invitrogen); 10X Buffer (Invitrogen); água estéril para um total de 20 µL na reação final. A reação foi incubada a 37° C por 1 hora e logo após adicionou-se 1 µL da enzima SAP (*Shrimp Alkaline Phosphatase*) (Invitrogen). A reação foi novamente incubada a 37° C por mais 1 hora e submetida a uma eletroforese em gel de agarose 1% corado com brometo de etídeo. A banda linearizada contendo aproximadamente 3 Kb foi retirada do gel e purificada com o auxílio do Kit *GENE CLEAN II*(Q-BIO gene), resultando em vetor linearizado, purificado e pronto para ser utilizado na reação de ligação.

c) Ligação dos fragmentos com extremidades coesivas ao vetor pBlueScript SK-

Para reação de ligação do inserto ao plasmídeo, foram utilizados 6 µL do BAC (aproximadamente 60 ng/ µL); 0,5 µL de plasmídeo (aproximadamente 50 ng/µL); 2µL de 5X Rapid Ligation Buffer-T4 DNA ligase (Invitrogen); 1 U de T4 ligase (Invitrogen) e 0,5 µL de água estéril para 10 µL de volume final. A reação foi incubada 24 horas a 16° C.

2.5.2 Estratégia de Subclonagem 2 - Fragmentação através de nebulização

Um recurso para a obtenção de fragmentos aleatórios de tamanho razoável ao longo de todo o genoma é a fragmentação do DNA por nebulização, um método mecânico, que produz fragmentos com extremidades abruptas (que devem ser reparadas) e posterior clonagem dos fragmentos em um vetor bacteriano aberto. Desta forma pode-se obter mais facilmente uma biblioteca genômica que represente efetivamente todo o genoma.

a) Nebulização do BAC

O BAC foi nebulizado seguindo o protocolo do Kit TOPO[®] *Shotgun Subcloning Kit* (Invitrogen) com algumas modificações: Em um nebulizador (Figura 4), foi adicionada uma solução contendo 20 µg de DNA (BAC) e um tampão (TE pH 8,0; glicerol 20%), para um volume final de 750 µL. Esta solução foi nebulizada, com auxílio de uma bomba compressora (11 psi) durante 15 segundos. Após isso, a solução contendo o DNA fragmentado foi coletada e o resultado foi verificado através de eletroforese em gel de agarose 1% corado com brometo de etídeo. Foram feitas várias preparações até se obter um tamanho de fragmentos entre 400 pb e 4 kb. Um total de 700 µL da solução contendo o DNA fragmentado foi transferido para um tubo. Adicionaram-se 80 µL de acetato de sódio a 3 M, pH 5,2; 4 µL de glicogênio 20 mg/mL e 700 µL de isopropanol. A solução foi armazenada durante 30 minutos em freezer - 80° C e centrifugada a 13.000 rpm por 20 minutos

(4° C). O sobrenadante foi removido e o *pellet* foi lavado com 800 µL de etanol 80%. O DNA foi seco e ressuspenso em água estéril, para uma concentração final de 200 ng/µL.

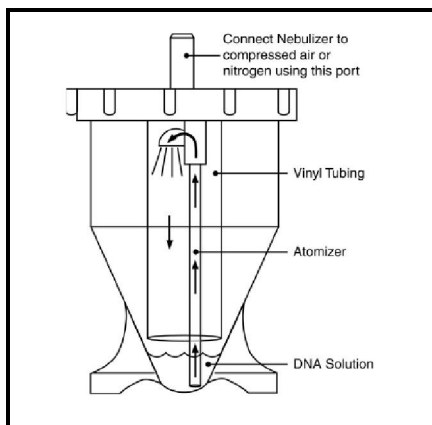


Figura 4: Desenho esquemático do nebulizador. Retirado de www.invitrogen.com.

b) Tratamento dos fragmentos com extremidades não-coesivas

Após a nebulização foram obtidos fragmentos de DNA com pontas não-coesivas, o que inviabiliza sua ligação ao vetor utilizado (pCR 4Blunt-TOPO - Invitrogen). Diante disso, os fragmentos foram tratados com as enzimas T4 DNA polimerase e Klenow (Invitrogen). Nesta reação, as enzimas completam as extremidades dos fragmentos a partir da incorporação de nucleotídeos na direção 5'-3' utilizando a ponta fita-simples como molde, sendo que a atividade exonucleotídica 3'-5' da T4 DNA polimerase retira as fitas simples que resultam da incorporação anterior, tornando as extremidades coesivas para posterior ligação com o vetor.

Os fragmentos de DNA foram tratados de acordo com o seguinte protocolo: Em um tubo, foram adicionados 5 µL de tampão 10X (Tris-HCl 100 mM; MgCl₂ 100 mM; NaCl 500 mM; ditiotreitól 10 mM; pH 7,9), 1 µL de BSA 1 mg/mL, 5 µL de dNTP (para uma concentração final de 250 µM), 2 µL da enzima T4 DNA polimerase e 2 µL do fragmento Klenow da DNA polimerase. Adicionou-se 3 µg de DNA fragmentado e completou-se com água deionizada para um total de 50 µL finais. A reação foi incubada a temperatura ambiente por 30 minutos e depois a 75° C por 20 minutos para inativar as enzimas.

c) Desfosforilação

Adicionou-se a reação contendo 50 µL descrita acima, 35 µL de água estéril, 10 µL do tampão de defosforilação 10X (Tris-HCl 0,5 M; EDTA 1 mM; pH 8,5), 5 U de CIP (*Calf Intestinal phosphatase*), para um volume total de 100 µL. A reação foi incubada a 37° C por 60 minutos. Após esse período, adicionaram-se 100 µL de uma solução fenol: clorofórmio: álcool isoamílico na proporção de 25:24:1 e centrifugou-se por 3 minutos sob velocidade máxima. A fase aquosa superior foi transferida para um novo tubo e o DNA foi precipitado adicionando 10 µL de acetato de sódio 3 M, pH 5,2; 1 µL de glicogênio 20 mg/mL e 300 µL de etanol absoluto gelado. A solução foi armazenada durante 30 minutos em freezer -80° C e logo após, centrifugada a 13.000 rpm por 20 minutos a 4° C. O sobrenadante foi removido e o *pellet* de DNA lavado com 500 µL de etanol 80%. O DNA foi seco e ressuspenso em 20 µL de água estéril para uma concentração final de 50 a 100 ng/µL.

d) Ligação dos fragmentos tratados ao vetor pCR 4Blunt-TOPO

Para uma reação de ligação de 6 µL finais, adicionaram-se: 4 µL de DNA não diluído (aproximadamente 100 ng de DNA); 1 µL de solução salina (NaCl 1,2 M; MgCl₂ 0,06 M) e 1 µL do vetor pCR 4Blunt-TOPO. A reação foi incubada a temperatura ambiente *over-night* e precipitada adicionando-se 100 µL de água estéril e incubando-se a 37° C por 10 minutos. Após esse período foram adicionados 10 µL de acetato de sódio a 3 M, pH 5,2; 2 µL de glicogênio 20 mg/mL e 300 µL de etanol absoluto (-20° C). A reação foi armazenada durante 30 minutos em freezer -80° C e centrifugada a 13.000 rpm por 20 minutos a 4° C. O sobrenadante foi removido e o *pellet* de DNA foi lavado com 800 µL de etanol 80%. Após isso, o DNA foi seco e ressuspenso em 10 µL de água estéril.

2.6 Preparo de células eletrocompetentes

As células de *Escherichia coli* (Cepa XL1-Blue) foram preparadas de acordo com o seguinte protocolo: 10 mL de uma cultura saturada foi inoculada em 1 L de

meio LB (com metade da concentração de sal) e crescida até alcançar uma O.D 0,7 (Densidade óptica igual a 0,7) sob comprimento de onda de 600 nm com ajuda de um espectrofotômetro. As células foram coletadas por centrifugação (4000 rpm por 20 minutos a 4° C) e ressuspensas duas vezes em água estéril gelada (1L e 500 mL). Em seguida, foram ressuspensas em 20 mL de glicerol 10%, centrifugadas (sempre nas mesmas condições) e ressuspensas novamente em 2 mL de glicerol 10%, separadas em alíquotas de 40 µL e estocadas em freezer -80° C.

2.7 Eletroporação

Para a etapa de transformação, foram adicionados 1 µL da reação de ligação e 40µL de células para eletroporação *E. coli* (cepa *XL1 Blue*) em uma cuveta de eletroporação previamente resfriada no gelo. O eletroporador Gene Pulser (Biorad) foi ajustado às seguintes condições: capacitância: 25 µFD; resistência: 200 ; voltagem: 1,80 KV. Imediatamente após a eletroporação, adicionou-se 1 mL de meio LB na cuveta, transferindo cuidadosamente todo o preparado para um novo tubo e incubando-o a 37° C durante 1 hora. Um total de 100 µL de células resultantes da transformação foram plaqueadas em cada Placa de Petri contendo 20 mL de meio LB-ágar suplementado com 100 µg/mL do antibiótico ampicilina, 10 µL de IPTG (20%) e 40 µL de X-Gal (20 mg/mL). A incubação foi realizada durante 16 horas a 37° C.

2.8 Seleção de transformantes e cultura permanente

As colônias recombinantes (brancas) foram replicadas para uma placa estéril contendo 96 poços. Cada nova placa continha 96 colônias brancas, dispostas em 12 colunas (numeradas de 01 a 12) por oito fileiras (classificadas de A a H). Cada clone advindo da técnica de nebulização, contido nas placas recebeu a nomenclatura respeitando alguns critérios, como por exemplo: AD25F09-PA-A01.RM – onde respectivamente AD25F09 = clone BAC de *Arachis duranensis* placa 25 posição F09; PA = Placa A; A01 = posição do subclone na placa; RM = Primer Reverso M13.

Para as análises com enzimas de restrição, onde foram utilizados tubos, a nomenclatura respeitou os seguintes critérios, como por exemplo: AD25F09-HIII-

46.RM – onde o que difere é a identificação da enzima de restrição utilizada, como por exemplo *HindIII* – HIII e o número da amostra no tubo 46.

Em cada poço da placa estéril de 96 poços, foram adicionados 1,5 mL de meio LB contendo 1,5 µL do antibiótico ampicilina (100 mg/mL). As colônias foram inoculadas e os blocos foram selados com adesivo. Um pequeno orifício foi feito sobre cada poço para permitir a aeração. As placas foram incubadas a 37° C a 220 rpm, durante 16 horas.

Em outra placa de 96 poços destinada a conter a cultura permanente, foram adicionados 100 µL de glicerol 50% e 100 µL cultura de bactérias, para posteriormente serem congeladas em freezer -80° C.

2.9 Minipreparação de plasmídeos em tubos

Para extração de DNA plasmidial utilizou-se o protocolo de lise alcalina (Ahn *et al.*, 2000) com algumas modificações. As células foram coletadas por centrifugação a 13.000 rpm por 1 minuto. O sobrenadante foi descartado e o *pellet* foi ressuspenso em 200 µL de tampão (Tris-HCl 50 mM pH 8; EDTA 10 mM; 20 µg de RNase). Adicionaram-se 200 µL de tampão de lise (NaOH 200 mM; SDS 1%). A solução foi homogeneizada à temperatura ambiente e acrescida de 240 µL de tampão de neutralização (acetato de potássio a 3 M, pH 5,5). A solução foi homogeneizada novamente por 3 minutos à temperatura ambiente e centrifugada a 13.000 rpm por 1 minuto. O sobrenadante foi transferido para novo tubo, adicionando-se 200 µL de isopropanol e homogeneizando por 1 minuto à temperatura ambiente. A solução foi centrifugada a 13.000 rpm por 1 minuto e o sobrenadante descartado. O *pellet* foi lavado com 500 µL de etanol 70% gelado e posteriormente seco e ressuspenso em 50 µL de água *Milii-Q* estéril.

2.10 Minipreparação de plasmídeos em microplacas

O protocolo utilizado foi o de lise alcalina (Sambrook & Russel, 2001) com algumas modificações: Os blocos contendo as culturas foram centrifugados a 4.000 rpm por 6 minutos. O sobrenadante foi descartado e os blocos mantidos invertidos sobre papel absorvente. A cada poço foi adicionado 240 µL de solução GET

(Glucose 50 mM; EDTA 10 mM; Tris-HCl 25 mM pH 8) e posteriormente foram homogeneizados até a completa suspensão das células. Em seguida os blocos foram centrifugados novamente, descartando-se o sobrenadante e adicionando novamente 80 μ L da solução GET. Os blocos foram homogeneizados e 60 μ L das células foram transferidas para um microplaca com fundo arredondado contendo 2,5 μ L de RNase (20 mg/mL) e lisadas com 60 μ L de solução de lise (NaOH 200 mM; SDS 1%). As placas foram seladas, a solução homogeneizada por inversão e incubadas por 7 minutos a temperatura ambiente. As placas foram centrifugadas por alguns segundos (4.000 rpm). Em seguida foram adicionados 60 μ L de acetato de potássio 3 M, pH 5,2. As placas foram seladas, homogeneizadas por inversão e incubadas por 7 minutos à temperatura ambiente. As placas foram centrifugadas novamente por 6 minutos a 4.000 rpm e incubadas, sem o adesivo, em estufa a 90° C por 30 minutos. Após esse período, as placas foram incubadas em gelo por 10 minutos e centrifugadas a 4.000 rpm por 4 minutos. O sobrenadante foi filtrado em uma placa Milipore (MAGV N22) fixada, com fita adesiva, no topo de uma nova microplaca de polipropileno de fundo em V com poços de 250 μ L. O conjunto foi centrifugado a 4.000 rpm por 5,5 minutos. A precipitação do DNA foi realizada adicionando-se 90 μ L de isopropanol ao filtrado. A placa foi homogeneizada por inversão e centrifugada por 45 minutos a 4.000 rpm. O sobrenadante foi descartado e o precipitado foi lavado com 200 μ L de etanol 70% gelado. O *pellet* foi seco a temperatura ambiente por 60 minutos e foi ressuspenso em 30 μ L de água *Milli-Q* estéril. As placas foram armazenadas a 20° C.

2.11 Sequenciamento

As reações foram realizadas utilizando o Kit de sequenciamento BigDye Terminator cycle Sequencing (Applied Biosystems, CA, USA) em um sequenciador automático ABI377 (Applied Biosystems), seguindo as concentrações para amplificação de DNA plasmidial: 1,5 μ L de primer (2 μ M) complementares à sequência do vetor (T7 e T3/ M13F e M13R); 2 μ L de DNA (aproximadamente 200 ng/ μ L); e 2 μ L de premix em 10 μ L de reação final por amostra. As amplificações foram feitas utilizando um termociclador nas seguintes condições: 30 ciclos com uma etapa de desnaturação (96° C por 20 segundos); anelamento (50° C por 15

segundos) e de extensão (60° C por 60 segundos). Os produtos das amplificações foram purificados seguindo o protocolo do Kit utilizado.

2.12 Análise dos dados de sequenciamento

As sequências foram processadas utilizando-se o programa Staden Package (Staden *et al.*, 2003). A interface *pregap4* permitiu o tratamento inicial das sequências, removendo as sequências dos vetores através da ferramenta Cross_match e analisando a qualidade de cada base nucleotídica através de PHRED (Ewing *et al.*, 1998). Utilizando a ferramenta *Cap3* (Huang & Madan, 1999) dentro da interface do Staden, foi possível realizar a compilação, retificação, leitura e adição de sequências.

Para a identificação das ORFs, ou prováveis regiões codificadoras, utilizou-se o sistema de anotação automática do arroz, o RiceGAAS (*Rice Genome Automated Annotation System* - <http://ricegaas.dna.affrc.go.jp/>), (Sakata *et al.*, 2002), o qual é composto por programas de predição e análises de códigos de proteínas e estruturas de genes.

As sequências obtidas através da subclonagem do BAC AD25F09 foram comparadas com o banco de dados GeneBank, utilizando o programa BLAST (*Basic Local Alignment and Search Tool*) (Altschul *et al.*, 1997). Sequências GSS (*Genome Survey Sequence*) do gênero *Arachis* depositadas no Banco de Dados do NCBI (*National Center for Biotechnology Information* - (<http://www.ncbi.nlm.nih.gov/>) que contém leituras aleatórias de fragmentos genômicos, também foram utilizadas para a comparação. Outra estratégia utilizada baseou-se na comparação dessas sequências com ESTs de *Arachis*, podendo indicar genes putativos nas sequências anônimas.

3. RESULTADOS

3.1 Confirmação da presença do RGA

Através da técnica de hibridização por Southern blotting, foi possível confirmar a presença do RGA no BAC. A digestão do BAC AD25F09 com a enzima de restrição *HindIII* revelou uma banda de aproximadamente 4 Kb quando hibridizada com uma sonda feita a partir do RGA S1_A_36. O padrão de bandas identificando esta homologia é mostrado na Figura 5.

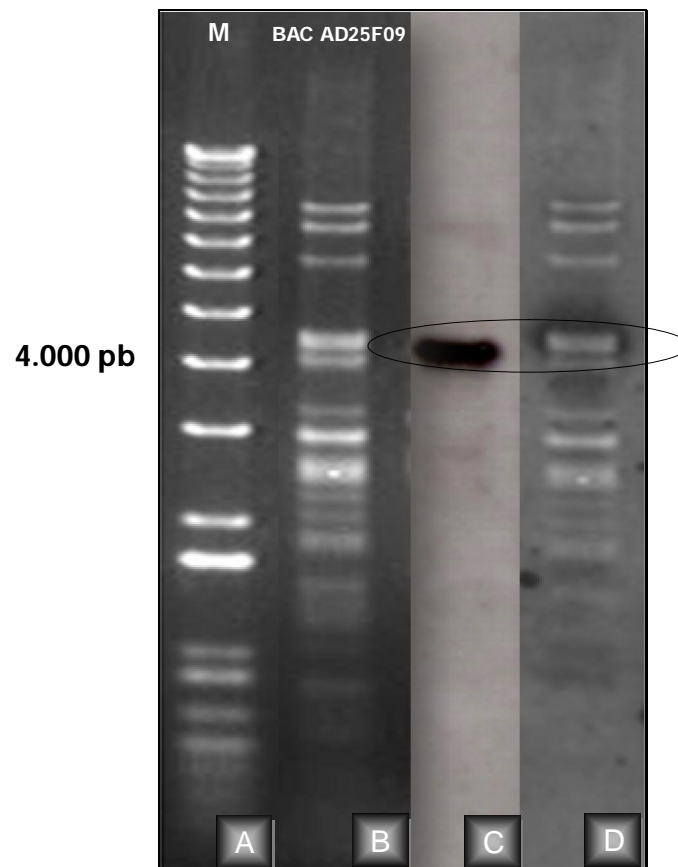


Figura 5: Marcador 1 Kb plus (A). Digestão do BAC AD25F09 com a enzima de restrição *HindIII* (B). *Southern Blotting* realizado com o BAC AD25F09 digerido e hibridizado com uma sonda construída a partir de um produto de PCR referente ao RGA S1_A_36 (C). Imagens sobrepostas (D).

3.2 Estratégias de subclonagem

A digestão do BAC AD25F09 com os três conjuntos enzimas de restrição produziu diferentes perfis, como esperado, em virtude dos diferentes sítios de restrição de cada enzima. A Figura 5B exemplifica um perfil de restrição realizado com a enzima *HindIII*.

O BAC AD25F09 digerido com a enzima de restrição *HindIII* e ligado em vetor pBS também linearizado com a mesma enzima, gerou um total de 305 colônias recombinantes (em um total de 412). Após a extração de DNA plasmidial, foram obtidas 352 seqüências (sequenciamento da maioria dos clones em apenas um sentido).

Quando o BAC foi digerido com a enzima de restrição *SauBAI* e o vetor com a enzima *BamHI*, foram obtidas 60 colônias e um total de 100 seqüências (20 clones sequenciados em um sentido e 40 em ambos).

Na digestão onde se utilizou a enzima *BamHI* para digerir tanto o BAC quanto o vetor, foram obtidas 66 colônias. Os plasmídeos recombinantes obtidos nessa estratégia foram sequenciados em uma só direção, produzindo 66 seqüências.

Um melhor resultado foi observado, quando a subclonagem foi realizada utilizando a mesma enzima de restrição tanto para o BAC quanto para o vetor. O perfil observado quando o BAC foi digerido com a enzima *HindIII* foi de fragmentos de vários tamanhos, enquanto que com a enzima *BamHI*, os fragmentos gerados foram em sua maioria grandes, sendo a ligação destes em vetor, mais difícil. Os fragmentos gerados com a enzima *SauBAI* foram, em contrapartida, muito pequenos e a ligação realizada em vetor linearizado com *BamHI*, mesmo estas sendo compatíveis, também não proporcionou uma quantidade satisfatória de colônias.

Este tipo de subclonagem gerou 431 colônias referentes aos três métodos de digestão e um total de 518 seqüências. Apesar da baixa quantidade de clones obtidos, foi de extrema importância a fragmentação do BAC com diferentes enzimas de restrição, pois a sobreposição de seqüências ajudou na montagem deste. Neste experimento, a extração do DNA plasmidial foi realizada através de duas formas: em tubos individuais e em microplacas de 96 poços. As duas estratégias foram satisfatórias, onde o método de extração individual proporcionou uma maior

quantidade de DNA, porém o método de extração em placa foi preferencialmente utilizado, em virtude da rapidez e uma maior homogeneidade entre amostras.

A fragmentação do BAC através da técnica de nebulização gerou fragmentos de tamanhos entre 400 pb e 4.000 pb de acordo com a Figura 6. Os fragmentos subclonados em plasmídeo deram origem a 1.536 colônias recombinantes (de um total de 1.854) distribuídas em 15 placas de 96 poços (sete placas foram sequenciadas em apenas um sentido e dez foram sequenciadas em ambos os sentidos) gerando 2.592 sequências. O número de colônias foi cinco vezes maior do que o obtido pela técnica a qual foram utilizadas enzimas de restrição (Tabela 1). Esta metodologia (nebulização) diferenciou-se da anterior, por não haver ligação entre extremidades compatíveis, e sim pela ligação entre fragmentos com terminais não-coesivos. Neste experimento, utilizou-se apenas o método de extração plasmidial de placas de 96 poços.

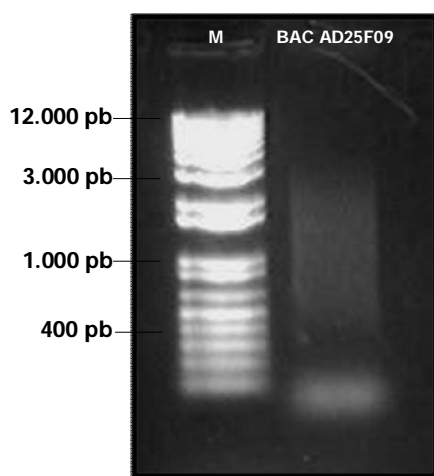


Figura 6: Gel de agarose 1% corado com brometo de etídeo, contendo o marcador molecular 1Kb plus (Invitrogen) e o BAC AD25F09 após a nebulização.

Tabela 1: Número de colônias recombinantes e sequências obtidas através das diferentes estratégias de subclonagem.

Subclonagem do BAC AD25F09		
Estratégia 1: Enzimas de restrição	Nº de colônias	Nº de sequências
BAC – <i>Hind</i> III / Vetor – <i>Hind</i> III	305	352
BAC – <i>Sau</i> 3AI / Vetor – <i>Bam</i> HI	60	100
BAC – <i>Bam</i> HI / Vetor – <i>Bam</i> HI	66	66
Estratégia 2: Nebulização	1.536	2.592
TOTAL	1.967	3.110

3.3 Confirmação da presença do inserto no vetor

Após a minipreparação de plasmídeos provenientes das duas estratégias de subclonagem, a confirmação da presença dos insertos no vetor foi realizada através de análise por enzimas de restrição, liberando o inserto do vetor, como exemplificado na Figura 7.

Apenas as amostras que comprovadamente apresentaram insertos foram enviadas para o sequenciamento de DNA.

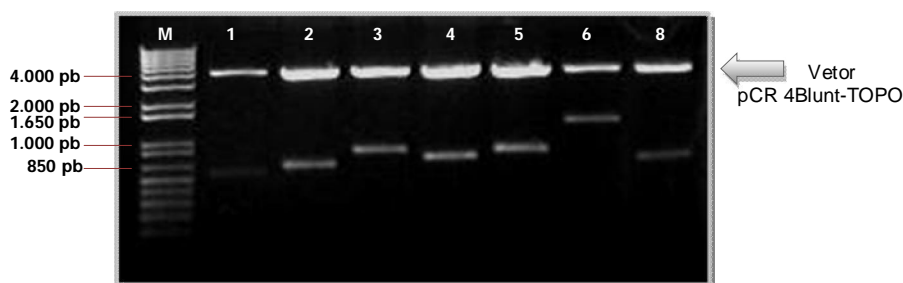


Figura 7: Gel de agarose 1% corado com brometo de etídeo referente à minipreparação de plasmídeos do BAC AD25F09 no vetor de clonagem pCR 4Blunt-TOPO (3.957 pb) (Invitrogen). O marcador molecular utilizado foi o 1Kb plus (Invitrogen)

3.4 Análise dos dados de sequenciamento

Ao reunir as duas metodologias de fragmentação do BAC, obteve-se um total de 3.110 sequências, que foram submetidas a análises pelo Programa Staden Package (Figura 8). Destas, 2.642 puderam ser analisadas pelo programa e 468 foram descartadas, principalmente, em virtude da má qualidade do sequenciamento. Um total de 1.108.565 bases foi agrupado pelo Programa, gerando 505 contigs de tamanhos variados. O BAC subclonado possuía um inserto de 110.000 pb e o sequenciamento proporcionou uma cobertura de 10 vezes o tamanho do inserto do BAC.

A soma dos nucleotídeos de todos os *contigs* foi 203.694 pb, ou seja, um valor aproximado do dobro do tamanho do inserto. Apenas os oito contigs de tamanhos mais expressivos, com valores mostrados na Tabela 2, somados, geraram 102.256 pb, um valor considerado próximo ao tamanho do inserto do BAC (Figura 9).

Tabela 2: Composição dos oito *contigs* selecionados da subclonagem do BAC AD25F09.

	Nome do <i>Contig</i>	Tamanho (pb)	Nº de seqüências agrupadas
<i>Contig</i> 1	AD25F09-PO-F10.FM (#2274)	27.567	499
<i>Contig</i> 2	AD25F09-HIII-P2F03.FM (#832)	21.664	420
<i>Contig</i> 3	AD25F09-PC-F09.RM (#24)	12.723	266
<i>Contig</i> 4	AD25F09-PH-B11.RM (#1615)	11.838	268
<i>Contig</i> 5	AD25F09-HIII-P2G03.RM (#902)	9.709	207
<i>Contig</i> 6	AD25F09-PL-B03.FM (#995)	9.467	211
<i>Contig</i> 7	AD25F09-PN-C10.RM (#2373)	5.789	97
<i>Contig</i> 8	AD25F09-PG-G04.RM (#887)	3.499	58
	TOTAL	102.256	2026



Figura 8: Janela da Interface GAP (Staden Package) mostrando um editor de *contig*. A Figura exemplifica a formação de uma seqüência consenso a partir das seqüências obtidas através das duas metodologias de subclonagem. O círculo grande mostra seqüências advindas de placas confeccionadas pela fragmentação do BAC através de nebulização e o círculo pequeno mostra uma seqüência advinda da fragmentação do BAC pela enzima de restrição *Sau*3AI.

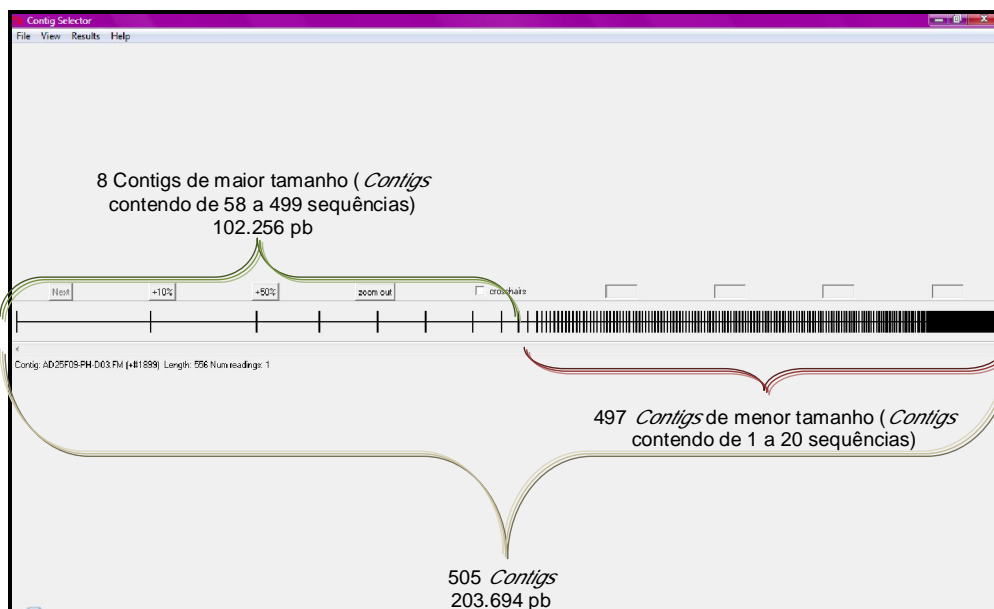


Figura 9: Janela da Interface GAP (Staden Package) mostrando todos os *contigs* produzidos a partir da subclonagem do BAC AD25F09. Em marrom, a escala de formação dos 505 *contigs*, do maior até o menor. Em verde, os *contigs* selecionados para anotação. Em vermelhos, os *contigs* não analisados em virtude do pequeno tamanho.

Os oito *contigs* de maior tamanho foram submetidos à análise através do Programa RiceGAAS, um sistema de anotação automatizada do genoma do arroz, do qual integra programas para predição e análise de estrutura dos genes e codificação das proteínas, como AutoPredgeneset, Blast, RepeatMasker, Genscan, RiceHMM, dentre outros. A análise gerou oito gráficos, que foram adaptados neste trabalho, somando-se várias outras informações, como o conteúdo A/T das sequências, identificação de marcadores SSR e sequências repetitivas. Para a identificação de genes putativos, também foram acopladas estratégias baseadas na comparação das sequências dos contigs gerados com bancos de dados do GenBank (NCBI), como de GSS do gênero *Arachis* (*Taxid*: 3817), que contém leituras aleatórias de fragmentos genômicos, além de comparações com bancos de ESTs.

Para os Gráficos referentes a cada *contig* (Gráficos 2 – 9), a linha preta representa a sequência de DNA do BAC com números indicando a posição das bases a cada intervalo específico, dependendo do tamanho de cada sequência. As setas verdes representam as regiões codificadoras preditas automaticamente pelo Programa AutoPredgeneset (RiceGAAS). Nestas regiões codificantes representadas pelas setas verdes, também foi possível realizar a comparação com sequências de nucleotídeos depositadas no GenBank, (blastn) ou proteínas (blastp), identificando similaridades e domínios conservados. As setas azuis correspondem ao resultado da análise através do Genescan utilizando-se a matriz padrão de *Arabidopsis thaliana*, arroz e monocotiledôneas. As setas vermelhas representam as regiões que apresentaram similaridade com proteínas depositadas no GenBank através do blastx.

A predição da função, para ser consistente, foi primeiramente baseada na grande similaridade da ORF com a proteína presente no banco de dados (também foi levado em conta o tamanho da sequência). Tal fato pode ser observado pelo número correspondente ao *E-value*, onde quanto menor foi esse valor, maior foi a probabilidade de similaridade real entre as sequências comparadas (uma vez que tal valor corresponde à probabilidade do alinhamento ter ocorrido ao acaso).

Como as anotações foram baseadas somente na sequência, ou seja, não foram realizados experimentos bioquímicos, a atribuição da função foi sempre hipotética (*Putative*) para cada ORF.

Casos existiram em que uma ORF, apesar de inicialmente ser considerada (computacionalmente) como sendo uma determinada proteína, não foi assim descrita e anotada, principalmente em vista das evidências e considerações realizadas, como a baixa similaridade com as proteínas dos bancos de dados, além de uma baixa porcentagem de cobertura e *E-value*. Dessa forma a proteína foi caracterizada como hipotética (exemplo: Gráfico 2 – ORF 1).

3.4.1 Contig 1 (AD25F09-PO-F10.FM)

O maior dos contigs gerados na subclonagem do BAC AD15F09, denominado *Contig 1*, possui 27.567 bases, resultante da junção de 499 sequências. A composição nucleotídica da sequência, ou seja, a frequência de bases é: A (32,69%), C (17,02%), G (18,24%) e T (32,06%).

De acordo com o Gráfico 2 e a Tabela 3, a análise preliminar do conteúdo biológico das sequências geradas através do sistema de anotação automática do arroz (RiceGAAS) identificou quatro regiões codificadoras de proteínas, sendo todas elas localizadas na fita senso. Destas quatro regiões, uma, a ORF 1, possui 2.980 pb sendo composta por 3 éxons e 2 íntrons, foi identificada como proteína hipotética.

A ORF 2 possui 4.048 pb e é composta por 6 éxons e 4 íntrons. A predição da função dessa região foi para uma poliproteína de cópia, ou seja, uma proteína essencial relacionada a elementos retrotransponíveis. Domínios conservados como a transcriptase reversa e a *RNAse H* foram identificados, evidenciando a identificação de um retrotransposon LTR da família Ty1-Copia, principalmente em virtude da organização dos domínios encontrados.

A ORF 3 possui 7.921 pb, sendo composta por 10 éxons e 8 íntrons. A ORF 4 possui 4.351 pb, com 5 éxons intercaladas por 4 íntrons. Essas ORFs (3 e 4), correspondem às regiões que codificam proteínas relacionadas à resistência. Alguns domínios conservados foram encontrados em ambas as regiões como, por exemplo, os domínios TIR, NB-ARC e LRR. Setas demonstrando sequências repetitivas foram desenhadas de acordo com análises feitas no Programa Staden Package. Elas mostram regiões que se repetem em tandem, situadas aonde foram encontrados os domínios TIR, NBS e LRR nos dois contigs, reforçando a idéia de que os genes de resistência podem estar ocorrendo com famílias em tandem.

As matrizes de *Arabidopsis* e monocotiledôneas (Programa Genscan) e arroz (RiceHMM) identificaram éxons bastantes similares a todas as regiões correspondentes às ORFs.

Em análises realizadas a partir de comparações com o banco de dados GSS do NCBI, um total de 111 hits ($E\text{-value} < e^{-107}$) com o gênero *Arachis* foi encontrado. Destes, 81 situavam-se no início da sequência, em um local onde não havia indícios para predição de genes. Nove hits praticamente cobriram a segunda região codificadora de proteína predita para este contig. Três hits situavam-se distribuídos pelo resto da sequência, principalmente nas regiões correspondentes às ORFs 3 e 4.

Um total de 18 hits ($E\text{-value} < 4e^{-24}$) com ESTs do gênero *Arachis* foram encontrados na sequência. Sete deles encontravam-se no início, em um local onde não havia indícios para predição de genes, assim como a maioria dos *hits* com sequências GSS. Todos os 11 *hits* restantes situavam-se em regiões codificadoras de proteínas.

3.4.2 Contig 2 (AD25F09-HIII-P2F03.FM)

O segundo maior contig gerado na subclonagem do BAC AD25F09, denominado *Contig 2*, possui 21.664 bases, resultante da junção de 420 sequências. A composição nucleotídica da sequência é: A (31,31%), C (18,51%), G (17,15%) e T (33,02%).

A análise preliminar do conteúdo biológico das sequências geradas através do sistema de anotação automática identificou cinco regiões codificadoras de proteínas, sendo duas delas localizadas na fita senso e três na fita antisenso (Gráfico 3 e Tabela 3). Destas cinco regiões, três foram identificadas como proteínas hipotéticas (ORFs 1, 3 e 5).

A ORF 1 possui 779 pb compondo-se 2 éxons e 1 íntron. A ORF 2, com 4.472 pb, possui 5 éxons e 3 íntrons e sua função foi predita como uma proteína ativadora de elemento transponível. A ORF 3 possui 1.319 pb, sendo composta de 3 éxons e 2 íntrons. Em relação à ORF 4, com 3.664 pb e 3 éxons e 2 íntrons, houve a identificação de uma putativa transcriptase reversa. A ORF 5 possui 731 pb sem íntrons.

As matrizes de *Arabidopsis* e monocotiledôneas (Programa Genscan) e arroz (RiceHMM) identificaram éxons bastantes similares a todas as regiões correspondentes às ORFs.

Em análises realizadas a partir de comparações com o banco de dados GSS do NCBI, um total de 116 *hits* ($E\text{-value} < 1E^{-126}$) com o gênero *Arachis* foi encontrado. Todos eles situavam-se em locais onde foram preditas ORFs.

Um total de 14 *hits* ($E\text{-value} < 6E^{-26}$) com ESTs do gênero *Arachis* foram encontrados na sequência, sendo que a maioria encontrava-se em regiões codificadoras de proteínas.

3.4.3 Contig 3 (AD25F09-PC-F09.RM)

O terceiro maior contig gerado na subclonagem do BAC AD25F09, denominado *Contig 3*, possui 12.723 bases, resultante da junção de 266 sequências. A composição nucleotídica é: A (34,50%), C (15,71%), G (14,73%) e T (35,05%).

A análise preliminar do conteúdo biológico das sequências geradas através do sistema de anotação automática identificou três regiões codificadoras de proteínas, sendo duas localizadas na fita senso e uma na fita antisenso (Gráfico 4 e Tabela 3). Destas regiões, duas delas (ORFs 2 e 3) foram identificadas como proteínas hipotéticas.

A ORF 1 possui 5.185 pb compondo-se 4 éxons e 3 íntron. A predição da função corresponde a uma região que codifica uma proteína relacionada à resistência. Alguns domínios conservados foram encontrados como, por exemplo, os domínios TIR, NB-ARC e LRR. As matrizes de *Arabidopsis* e monocotiledôneas (Programa Genscan) e arroz (RiceHMM) identificaram éxons bastantes similares a região correspondente a ORF 1. A ORF 2, com 187 pb, não possui íntrons. A ORF 3 possui 571 pb contendo dois éxons e um íntron.

Em análises realizadas a partir de comparações com o banco de dados GSS do NCBI, um total de quatro *hits* ($E\text{-value} < 7e^{-83}$) com o gênero *Arachis* foi encontrado. Apenas um *hit* localizava-se no meio da ORF 1, dois hits na ORF 2 e o quarto *hit* em uma região não-codificadora, entre as ORFs 2 e 3. Nenhum *hit* com ESTs foi identificado na sequência correspondente a esse agrupamento.

3.4.4 Contig 4 (AD25F09-PH-B11.RM)

O quarto maior contig gerado na subclonagem do BAC AD25F09, denominado *Contig 4*, possui 11.838 bases, resultante da junção de 268 sequências. A composição nucleotídica da sequência é: A (32,39%), C (14,23%), G (17,37%) e T (36,02%).

A análise preliminar do conteúdo biológico das sequências geradas através do sistema de anotação automática identificou uma região codificadora de proteína localizada na fita senso (Gráfico 5 e Tabela 3). Esta possui 6.907 pb compondo-se 8 éxons e 7 íntrons. De acordo com análises esta região codifica uma proteína relacionada à resistência. Alguns domínios conservados foram encontrados nestas regiões como, por exemplo, os domínios TIR, NB-ARC e LRR.

As matrizes de *Arabidopsis* e monocotiledôneas (Programa Genscan) e arroz (RiceHMM) identificaram éxons bastantes similares a região correspondente às ORFs.

Em análises realizadas a partir de comparações com o banco de dados GSS do NCBI, apenas um *hit* ($E\text{-value} = 0.0$) com o gênero *Arachis* foi encontrado. Este estava localizado numa das extremidades da ORF. Três *hits* com ESTs ($E\text{-value} < 3e^{-47}$) do gênero *Arachis* foram identificados, sendo que dois deles encontravam-se no meio da ORF e um no final dela, exatamente no mesmo local onde foi identificado um dos *hits* com seqüências GSS.

3.4.5 Contig 5 (AD25F09-HIII-P2G03.RM)

O quinto maior contig gerado na subclonagem do BAC AD25F09, denominado *Contig 5*, possui 9.709 bases, resultante da junção de 207 sequências. A composição nucleotídica da sequência é: A (33,08%), C (16,36%), G (16,64%) e T (33,92%).

A análise preliminar do conteúdo biológico das sequências geradas através do sistema de anotação automática identificou três regiões codificadoras de proteínas, sendo duas na fita senso e uma na fita antisenso (Gráfico 6 e Tabela 3). Destas três regiões, uma foi identificada como proteína hipotética (ORF 1).

A ORF 1 possui 2.449 pb, compondo-se por 2 éxons e 1 íntron. A ORF 2, com 887 pb sem íntrons foi identificada como uma putativa proteína ligante de DNA. A ORF 3 possui 485 pb sem íntrons e de acordo com análises, esta região corresponde a uma proteína relacionada à resistência. Um domínio conservado foi encontrado nesta região, sendo ele o domínio TIR. Setas representando sequências repetitivas demonstram regiões relacionadas, porém não codificantes, sendo estas situadas em regiões anteriores às ORFs.

As matrizes de *Arabidopsis* e monocotiledôneas (Programa Genscan) e arroz (RiceHMM) identificaram éxons bastantes similares a região correspondente às ORFs.

Em análises realizadas a partir de comparações com o banco de dados GSS, dois *hits* ($E\text{-value} < 2e^{-151}$) com o gênero *Arachis* foram encontrados em regiões codificadoras. Treze *hits* com ESTs ($E\text{-value} < 1e^{-09}$) do gênero *Arachis* foram identificados, sendo que sete deles encontravam-se em regiões codificadoras.

3.4.6 Contig 6 (AD25F09-PL-B03.FM)

O sexto maior contig gerado na subclonagem do BAC AD25F09, denominado *Contig 6*, possui 9.467 bases, resultante da junção de 211 sequências. A composição nucleotídica da sequência é: A (33,24%), C (17,06%), G (19,87%) e T (29,83%).

A análise preliminar do conteúdo biológico das sequências geradas através do sistema de anotação automática identificou duas regiões codificadoras de proteínas localizadas na fita senso (Gráfico 7 e Tabela 3). Destas duas regiões, uma foi identificada como proteína hipotética (ORF 1).

A ORF 1 possui 824 pb sem íntrons. A ORF 2, com 4.397 pb, possui 5 éxons e 4 íntrons e sua função foi predita como uma putativa poliproteína *pol*, encontrada em retrotransposons. Em outra análise realizada através do Blastx, essa mesma região, além de outra anterior à ORF 1 foram bastante similares à proteína Integrase, também pertencente à classe dos retrotransposons LTR.

As matrizes de *Arabidopsis* e monocotiledôneas (Programa Genscan) e arroz (RiceHMM) identificaram éxons similares a região correspondente às ORFs.

Em análises realizadas a partir de comparações com o banco de dados GSS do NCBI, 29 *hits* ($E\text{-value} < 8e^{-12}$) com o gênero *Arachis* foi encontrado, onde 15 situavam-se na segunda região codificadora e quatro na região similar a uma putativa proteína Integrase, de acordo com o Blastx. Um *hit* ($E\text{-value} < 2e^{-28}$) com um EST do gênero *Arachis* foi localizado na extremidade da ORF 2.

3.4.7 Contig 7 (AD25F09-PN-C10.RM)

O sétimo maior contig gerado na subclonagem do BAC AD25F09, denominado *Contig 7*, possui 5.789 bases, resultante da junção de 97 sequências. A composição nucleotídica da sequência é: A (27,12%), C (18,48%), G (20,19%) e T (34,20%).

A análise preliminar do conteúdo biológico das sequências geradas através do sistema de anotação automática identificou duas regiões codificadoras de proteínas, sendo uma na fita senso e uma na fita antisenso (Gráfico 8 e Tabela 3). Estas duas regiões foram identificadas como proteínas hipotéticas, sendo que a ORF 1 possui 1.334 pb, sem íntrons e a ORF 2 possui 1320 pb, compostas por 3 éxons e 2 íntrons.

As matrizes de *Arabidopsis* e monocotiledôneas (Programa Genscan) e arroz (RiceHMM) identificaram éxons bastante similares a região correspondente a ORF 1. Em análises realizadas a partir de comparações com o banco de dados GSS do NCBI, 102 *hits* ($E\text{-value} < 7e^{-110}$) com o gênero *Arachis* foram encontrados. A distribuição destes *hits* cobriu praticamente toda sequência desde o início da ORF 1 até o final da ORF 2, indicando uma região de alta repetitividade. Apenas um *hit* ($E\text{-value} < 1e^{-47}$) com um EST do gênero *Arachis*, com aproximadamente 100 pb, foi localizado em uma das extremidades da ORF 2.

3.4.8 Contig 8 (AD25F09-PG-G04.RM)

O oitavo maior contig gerado na subclonagem do BAC AD25F09, denominado *Contig 8*, possui 3.499 bases, resultante da junção de 58 sequências. A composição nucleotídica da sequência é: A (28,58%), C (17,06%), G (20,95%) e T (33,41%).

A análise preliminar do conteúdo biológico das sequências geradas através do sistema de anotação automática não identificou regiões codificadoras de proteínas (Gráfico 9 e Tabela 3), porém, análises realizadas comparando a sequência com o banco de dados de proteínas não-redundantes através do Blastx, identificou uma região de aproximadamente 1.700 pb com similaridade a uma proteína de resistência da família TIR-NBS-LRR de *Medicago truncatula* ($E\text{-value} < 7e^{-64}$). As matrizes de *Arabidopsis* e monocotiledôneas (Programa Genscan) identificaram éxons bastantes similares a essa região. A partir de comparações com o banco de dados de ESTs quatro *hits* ($E\text{-value} < 6e^{-29}$) com ESTs do gênero *Arachis* foram identificados e o maior deles, com aproximadamente 500 pb, localizava-se em uma das extremidades desta região identificada como similar a uma proteína de resistência. Comparando-se a sequência com o banco de dados de sequências GSS do NCBI, 10 *hits* ($E\text{-value} < 4e^{-45}$) com o gênero *Arachis* foram encontrados. A distribuição destes *hits* deu-se em regiões não codificadoras.

A partir de análises feitas para os oito maiores contigs, um total de 20 possíveis regiões codificadoras de proteínas foi identificado neste trabalho, como pode ser observado no Gráfico 1 e Tabela 3. A disposição destas possíveis regiões nos *contigs* pode ser visualizada na Figura 10.

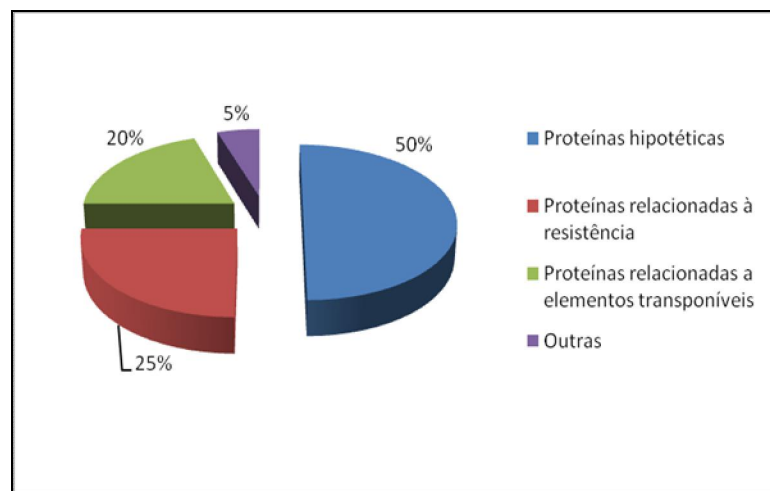


Gráfico 1 : Porcentagem de ORFs identificadas no BAC AD25F09.

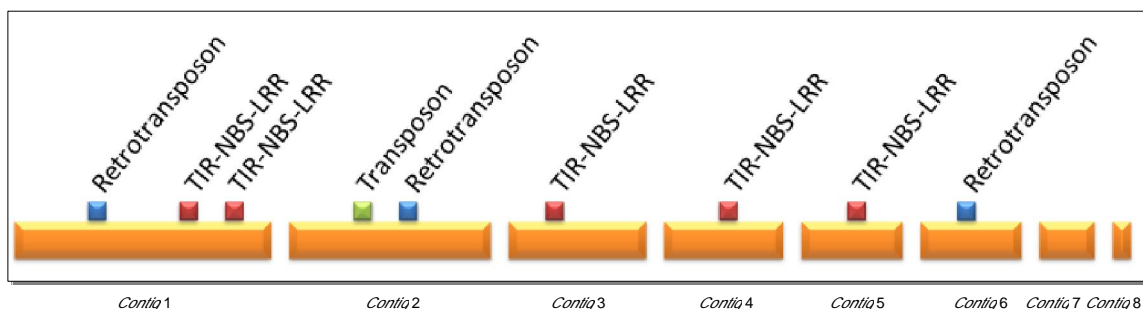


Figura 10: Disposição das possíveis regiões codificadoras de proteínas nos *contigs* montados a partir do BAC AD25F09.

Tabela 3: Lista de ORFs preditas para cada *contig*, através de análises feitas pelo Programa RiceGAAS.

Contig 1 - AD25F09-PO-F10.FM			
	ORF 1	Código(s) de Acesso	E-value
Predição da Função	Hypothetical protein	-	-
Blastn	AdH1157-CPY2-M16TF AdH1 <i>Arachis duranensis</i> genomic clone	FI285460	3e-98
Blastp	Hypothetical protein	-	-
	ORF 2	Código(s) de Acesso	E-value
Predição da Função	Putative copia polyprotein (<i>Drosophila simulans</i>) retrotransposon copia (fragments)	gi 345511 pir PC1232	5e-59
Blastn	AHT1B5RGH003-M17TP AHT1BS-Rearray <i>Arachis hypogaea</i> genomic clone	FI503053	0.0
Blastp	Putative copia-type pol polyprotein-like [<i>Arabidopsis thaliana</i>]	gi 8777569 dbj BAA97087.1	e-169
Blastx 1	retrotransposon protein, putative, Ty1-copia subclass [<i>Oryza sativa</i> (japonica cultivar-group)]	gi 78708224 gb ABB47199.1	e-134
	ORF 3	Código(s) de Acesso	E-value
Predição da Função	TMV resistance protein N	gi 46577339 sp Q40392.1	e-123
Blastn	<i>Arachis hypogaea</i> isolate PLTRP1H11 resistance protein PLTR gene	AY747357	0.0
Blastp	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	0.0
Blastx 2	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	5e-52
Blastx 3	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	6e-66
Blastx 4	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	1e-57
	ORF 4	Código(s) de Acesso	E-value
Predição da Função	TMV resistance protein N	gi 46577339 sp Q40392.1	e-115
Blastn	<i>Arachis stenosperma</i> clone S1_A_36 RGA	AY157808	0.0
Blastp	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	0.0
Blastx 5	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	2e-55
Blastx 6	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	0.0

Contig 2 - AD25F09-HIII-P2F03.FM			
	ORF 1	Código(s) de Acesso	E-value
Predição da Função	Hypothetical protein	-	-
Blastn	AdH1165-CPY2-J21TR AdH1 <i>Arachis duranensis</i> genomic clone	FI302164	2e-63
Blastp	No hit	-	-
Blastx 1	Hypothetical protein	-	-
Blastx 2	Hypothetical protein	-	-
Blastx 3	AC073433_4 transposase, putative [<i>Arabidopsis thaliana</i>]	gi 12321083 gb AAG50652.1	2e-99
	ORF 2	Código(s) de Acesso	E-value
Predição da Função	hAT family dimerisation domain containing protein [<i>Oryza sativa</i> (japonica cultivar-group)]	gi 77552015 gb ABA94812.1	3e-92
Blastn	AdH1163-CPY2-E12TF AdH1 <i>Arachis duranensis</i> genomic clone	FI304541	0.0
Blastp	AC073433_4 transposase, putative [<i>Arabidopsis thaliana</i>]	gi 12321083 gb AAG50652.1	e-119
Blastx 4	Hypothetical protein	-	-
Blastx 5	AC073433_4 transposase, putative [<i>Arabidopsis thaliana</i>]	gi 12321083 gb AAG50652.1	4e-19
	ORF 3	Código(s) de Acesso	E-value
Predição da Função	Hypothetical protein	-	-
Blastn	703_4_14168661_16654_43576_035 <i>Arachis batizocoi</i> genomic clone	DX508553	1e-21
Blastp	No hit	-	-
	ORF 4	Código(s) de Acesso	E-value
Predição da Função	Putative RNA-directed DNA polymerase (Reverse transcriptase); Zinc finger, CCHC-type; Peptidase asparic, tactive site; Retrotransposon gag protein [<i>Medicago truncatula</i>]	gi 87162498 gb ABD28293.1	1e-04
Blastn	AdH1029-CPY2-N23TF AdH1 <i>Arachis duranensis</i> genomic clone	FI281874	e-159
Blastp	Putative RNA-directed DNA polymerase (Reverse transcriptase); Zinc finger, CCHC-type; Peptidase asparic, tactive site; Retrotransposon gag protein [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	1e-04
Blastx 6	Putative reverse transcriptase [<i>Arabidopsis thaliana</i>]	gi 4263052 gb AAD15321.1	0.4
Blastx 7	Hypothetical protein Osl_04802 [<i>Oryza sativa</i> Indica Group]	gi 125528728 gb EAY76842.1	0.003
	ORF 5	Código(s) de Acesso	E-value
Predição da Função	Hypothetical protein	-	-
Blastn	AdH1181-CPY2-K18TF AdH1 <i>Arachis duranensis</i> genomic clone	FI310420	0.0
Blastp	Hypothetical protein Osl_04802 [<i>Oryza sativa</i> Indica Group]	gi 125528728 gb EAY76842.1	4e-05
Contig 3 - AD25F09-PC-F09.RM			
	ORF 1	Código(s) de Acesso	E-value
Predição da Função	TMV resistance protein N	gi 46577339 sp Q40392.1	e-139
Blastn	AdH1168-CPY2-N15TR AdH1 <i>Arachis duranensis</i> genomic clone	FI312554	0.0

Blastp	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	0.0
Blastx 1	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	e-104
ORF 2		Código(s) de Acesso	E-value
Predição da Função	Hypothetical protein	-	-
Blastn	<i>Arachis batizocoi</i> methylation unfiltered library (LibID764) <i>Arachis batizocoi</i> genomic clone	DX509051	2e-11
Blastp	No hit	-	-
Blastx 2	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	1e-47
ORF 3		Código(s) de Acesso	E-value
Predição da Função	Hypothetical protein	-	-
Blastn	AdH1026-CPY2-G18TR AdH1 <i>Arachis duranensis</i> genomic clone	F1287593	0.001
Blastp	Hypothetical protein	-	-
Blastx 3	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	7e-12
Contig 4 - AD25F09-PH-B11.RM			
ORF 1		Código(s) de Acesso	E-value
Predição da Função	TMV resistance protein N	gi 46577339 sp Q40392.1	e-103
Blastn	<i>Arachis stenosperma</i> clone S1_A_168 resistance protein gene, partial cds.	(AY157809)	e-151
Blastp	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	e-158
Blastx 1	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	e-112
Blastx 2	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	3e-47
Blastx 3	TIR [Medicago truncatula]	gi 87162908 gb ABD28703.1	e-116
Contig 5 - AD25F09-HIII-P2G03.RM			
ORF 1		Código(s) de Acesso	E-value
Predição da Função	Hypothetical protein	-	-
Blastn	C20L_79_G12_E002.g1 USDA-Tifton Peanut Library C20L <i>Arachis hypogaea</i> cDNA clone C20L-079_G12 5', mRNA sequence	ES756475	3e-21
Blastp	No hit	-	-
Blastx 1	WD40-like [<i>Medicago truncatula</i>]	gi 87162909 gb ABD28704.1	1e-07
Blastx 2	WD40-like [<i>Medicago truncatula</i>]	gi 87162909 gb ABD28704.1	1e-07
Blastx 3	WD40-like [<i>Medicago truncatula</i>]	gi 87162909 gb ABD28704.1	6e-13
Blastx 4	WD40-like [<i>Medicago truncatula</i>]	gi 87162909 gb ABD28704.1	6e-13
ORF 2		Código(s) de Acesso	E-value
Predição da Função	Putative DNA binding [<i>Arabidopsis thaliana</i>]	gi 15221318 ref NP_172083.1	1e-09
Blastn	No hit	-	-
Blastp	Putative DNA binding [<i>Arabidopsis thaliana</i>]	gi 15221318 ref NP_172083.1	1e-09
Blastx 5	Hypothetical protein	-	-
Blastx 6	Hypothetical protein	-	-
Blastx 7	Hypothetical protein	-	-
Blastx 8	Putative DNA binding [<i>Arabidopsis thaliana</i>]	gi 15240931 ref NP_198666.1	2e-04
ORF 3		Código(s) de Acesso	E-value
Predição da Função	TMV resistance protein N	gi 46577339 sp Q40392.1	6e-34
Blastn	SI_S017 SI, cDNA forward subtractive library from peanut leaves inoculated with <i>Cercosporidium personatum</i> <i>Arachis hypogaea</i> subsp. fastigiata cDNA	EL966749	e-128

Blastp	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	1e-52
Blastx 9	Hypothetical protein	-	-
Blastx 10	WD40-like [<i>Medicago truncatula</i>]	gi 87162909 gb ABD28704.1	0.006
Blastx 11	Hypothetical protein	-	-
Blastx 12	TIR [<i>Medicago truncatula</i>]	gi 87162908 gb ABD28703.1	9e-51
Contig 6 - AD25F09-PL-B03.FM			
ORF 1		Código(s) de Acesso	E-value
Predição da Função	Hypothetical protein	-	-
Blastn	AdH1171-CPY2-E18TR AdH1 <i>Arachis duranensis</i> genomic clone	FI311144	1e-77
Blastp	Hypothetical protein [<i>Vitis vinifera</i>]	gi 147784422 emb CAN74951.1	4e-12
Blastx 1	Putative Integrase, Zinc Finger [<i>Medicago truncatula</i>]	gi 87240724 gb ABD32582.1	0.0
Blastx 2	Hypothetical protein [<i>Vitis vinifera</i>]	gi 147834092 emb CAN64335.1	4e-09
ORF 2		Código(s) de Acesso	E-value
Predição da Função	gag-pol polyprotein [<i>Glycine max</i>]	gi 29423279 gb AAO73527.1	e-117
Blastn	AdH1191-CPY2-K10TF AdH1 <i>Arachis duranensis</i> genomic clone	FI305248	0.0
Blastp	Putative Integrase, catalytic region; Zinc finger, CCHC-type; Peptidase aspartic, catalytic [<i>Medicago truncatula</i>]	gi 87240724 gb ABD32582.1	e-143
Blastx 3	Putative Integrase, catalytic region; Zinc finger, CCHC-type; Peptidase aspartic, catalytic [<i>Medicago truncatula</i>]	gi 87240724 gb ABD32582.1	e-156
Contig 7 - AD25F09-PN-C10.RM			
ORF 1		Código(s) de Acesso	E-value
Predição da Função	Hypothetical protein	-	-
Blastn	764_6_14176532_5489_43635_011 <i>Arachis batizocoi</i> methylation unfiltered library (LibID764) <i>Arachis batizocoi</i> genomic	DX510941	0.0
Blastp	Hypothetical protein	-	-
Blastx 1	Hypothetical protein	-	-
ORF 2		Código(s) de Acesso	E-value
Predição da Função	Hypothetical protein	-	-
Blastn	643_2_14152159_16654_43223_092 <i>Arachis duranensis</i> methylation filtered library (LibID643) <i>Arachis duranensis</i> genomic clone	DX511165	4e-36
Blastp	No hit	-	-
Blastx 2	Hypothetical protein	-	-

De acordo com análises feitas a partir de comparações com sequências depositadas no banco de dados (NCBI), foram identificados ao todo 54 *hits* com ESTs do gênero *Arachis* e 51,8% destes estavam associados a regiões codificantes (Gráfico 10). Em relação às comparações com seqüência GSS, foram identificados 365 *hits* com fragmentos GSS do gênero *Arachis*, dos quais 70,4 % estavam associadas a regiões codificantes situadas nas sequências produzidas (Gráfico 11).

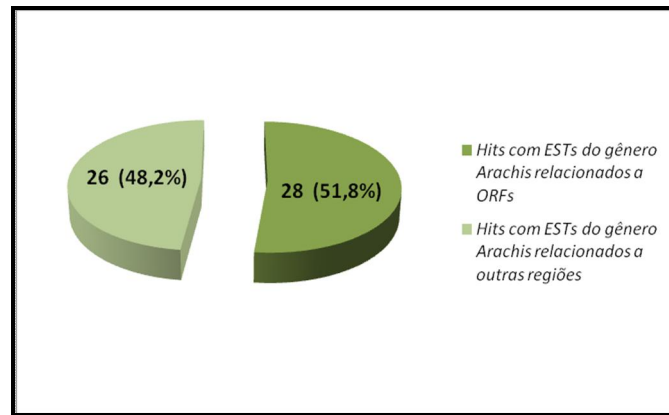


Gráfico 10: Comparação das sequências dos *contigs* gerados a partir da subclonagem do BAC AD25F09, com bancos de dados de ESTs depositados no GenBank (NCBI).

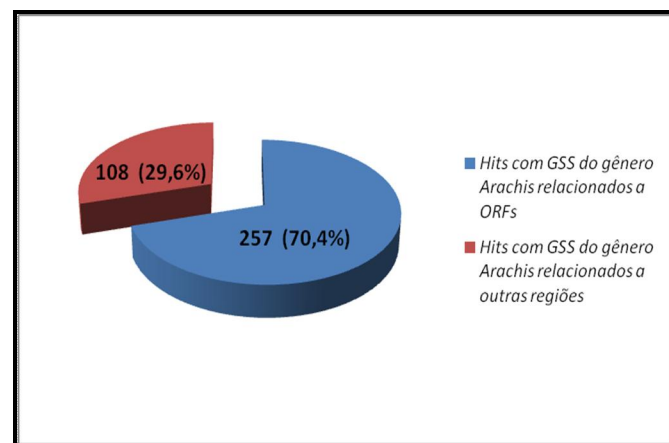


Gráfico 11: Comparação das sequências dos *contigs* gerados a partir da subclonagem do BAC AD25F09, com bancos de dados de sequências GSS depositadas no GenBank (NCBI).

4. DISCUSSÃO

A identificação de análogos a genes de resistência (RGAs) no genoma de plantas é bastante importante, pois estes são fortes indícios da presença de genes ou grupos de genes de resistência. Até hoje, vários marcadores RGAs ligados a genes de resistência já foram identificados em diversas culturas como batata, tomate, aveia, cenoura, maçã, alface e soja (Van der Linden *et al.*, 2004; Calenge *et al.*, 2005; Syed *et al.*, 2006; Hayes & Saghai-Marroof, 2000), enriquecendo mapas genéticos e elucidando os processos envolvidos na defesa de plantas contra diversas pragas.

Em *Arachis*, alguns RGAs foram identificados, a partir de *primers* construídos para motivos conservados da região NBS (Bertioli *et al.*, 2003). Alguns desses RGAs foram mapeados no mapa genético de referência (Leal-Bertioli *et al.*, 2009). O RGA S1_A_36 (isolado a partir do genoma AA de *Arachis stenosperma*) foi mapeado no GL 4 está co-localizado com um QTL para a resistência ao fungo foliar *C. personatum*. O sequenciamento total de um clone BAC contendo este RGA pode proporcionar uma ferramenta ideal para estudos envolvendo genes e elementos repetidos contidos neste BAC. Para tal, a subclonagem, o sequenciamento, a remontagem e a caracterização da sequência do BAC AD25F09 foram realizadas neste estudo.

Quanto às estratégias de subclonagem utilizadas, observou-se que, apesar da fragmentação por enzimas de restrição tipo II gerar extremidades compatíveis durante a clivagem do DNA e do vetor, facilitando o anelamento das extremidades coesivas complementares através da ligação com a enzima *DNA Ligase*, esta produziu uma menor cobertura do BAC.

Três conjuntos de enzimas foram utilizados para a geração de diferentes perfis de fragmentação, no intuito de promover a montagem da sequência. Uma importante consequência da especificidade das enzimas de restrição escolhidas é que o número de clivagens feitas por cada uma delas no BAC é definido (em virtude da frequência com que os sítios de corte ocorrem), portanto, cada enzima gerou uma família única e limitada de fragmentos.

Com a geração de 518 sequências, a montagem do BAC foi dificultada, principalmente em virtude da redundância das sequências obtidas, clonagem

preferencial dos fragmentos de menor tamanho e desenvolvimento de subclones contendo sequências quiméricas. Consequentemente, apenas uma pequena cobertura do BAC pôde ser representada por esta estratégia.

Alternativamente, optou-se pela utilização de outra forma de fragmentação do BAC, a nebulização (método mecânico). Esta estratégia gerou fragmentos de tamanhos aleatórios que puderam representar essencialmente todo o BAC. O perfil de fragmentação escolhido foi de 400 a 4000 pb. Perfis onde foram observados fragmentos menores que 400 pb foram descartados, com o intuito de diminuir a clonagem preferencial por fragmentos pequenos.

O método de fragmentação mecânica utilizado difere do enzimático, principalmente por gerar fragmentos de DNA aos quais as extremidades precisam ser reparadas, utilizando as enzimas *T4 polimerase* e fragmento *Klenow*. Após o tratamento dos fragmentos que tornou as extremidades abruptas, estes foram ligados ao vetor pCR 4Blunt-TOPO, através da técnica proposta pelo *TOPO[®] Shotgun Subcloning Kit*, que utiliza propriedades da topoisomerase (Shuman, 1994).

Este método de ligação, apesar de mais difícil quando comparado com o enzimático, foi mais eficiente, otimizando o tempo e produzindo cinco vezes mais colônias do que o método que utilizou a restrição enzimática, gerando dessa forma 2.592 sequências.

Ao reunir as duas metodologias de fragmentação do BAC, obteve-se 3110 sequências que foram agrupadas em 505 contigs, na tentativa de remontar o BAC inteiro. O grande número de sequências obtidas através da subclonagem foi bastante útil para a análise, pois uma grande parte do BAC pôde ser remontada utilizando 102.256 pb (referentes aos oito maiores contigs), ou seja, metade do número de pares de base obtidos neste trabalho que foi de 203.694 pb.

Os oito maiores *contigs* com tamanhos variando entre 3 e 28 Kb foram utilizados neste estudo para predição de possíveis genes encontrados no BAC. Juntos, esses oito contigs formaram 102.256 pb, o que é aproximadamente o tamanho real do inserto clonado no BAC, ou seja, 110.000 pb.

O sequenciamento do clone BAC, através da técnica de *shotgun*, apresentou duas fases: a primeira correspondeu à obtenção das sequências dos subclones e posterior formação de *contigs*, a segunda, dirigiu-se à montagem da sequência, onde foi necessário corrigir os vários tipos de anomalias que uma montagem pode

apresentar (sequências contaminantes, sequências de vetores não removidas, sequências quiméricas dentre outras). Além disso, as regiões de baixa qualidade foram identificadas e, quando necessário, a edição foi feita manualmente.

Os espaços ocasionais encontrados em virtude de alguns fragmentos de DNA que não conseguiram ser clonados serão preenchidos através da técnica de *Primer Walking*, onde se utiliza o final de uma sequência clonada para a confecção de um primer para sequenciar fragmentos adjacentes que não foram clonados. A amplificação dessas regiões onde, hoje, localizam-se os espaços, juntamente com os 497 *contigs* construídos, sem dúvida, auxiliará no fechamento total da sequência do BAC AD25F09.

O processo de identificação dos *frames* de leitura na sequência de DNA, seguido da identificação de possíveis funções desempenhadas pelos produtos gênicos, conhecido como processo de anotação gênica, são fundamentais na compreensão inicial dos genes. A obtenção desses dados é extremamente auxiliada pelo uso de programas de computadores, porém a utilização de critério por parte dos pesquisadores, também é um fator importante no processo de anotação.

Com auxílio do Programa de predição RiceGAAS, foram identificadas 20 ORFs distribuídas na sequência de oito *contigs* referentes à subclonagem e sequenciamento do BAC AD25F09. Exatamente 50% dessas regiões foram identificadas como proteínas hipotéticas, 25% com proteínas relacionadas à resistência, 20% com proteínas relacionadas a elementos transponíveis e 5% relacionadas a outros tipos de proteínas.

Algumas interfaces compiladas ao programa de predição RiceGAAS, como os programas Genscan e RiceHMM, dentre outros, identificaram éxons em matrizes de genomas já sequenciados como o de *Arabidopsis* e arroz respectivamente, bastantes similares a regiões correspondentes às ORFs identificadas neste trabalho. Além da identificação dessas regiões codificadoras de proteínas, outra análise adicional foi realizada, através da comparação das sequências dos *contigs* com bancos de dados de ESTs depositados no GenBank (NCBI). Este tipo de comparação também forneceu evidências na ocorrência de regiões gênicas, pelo menos no que diz respeito a sua estrutura, pois as sequências produzidas através de sequenciamentos de ESTs representam a expressão de genes estágio-específicos. Também foram utilizadas neste estudo sequências GSS, com o intuito

de comparar as sequências produzidas com fragmentos aleatórios de DNA genômico depositadas em bancos de dados sob diversas formas, conseguindo, dessa forma, capturar trechos de regiões codificantes únicas ou de baixa expressão. Essas ferramentas são poderosas para estudos de mapeamento e para confirmação de anotação *in silico* de genomas.

De acordo com as análises para predição de genes (Programa Predgeneset, integrado ao programa de anotação automatizada RiceGAAS) realizadas neste trabalho, foram encontradas cinco regiões distribuídas pelos contigs, similares ao gene N (GenBank ref. gi|46577339|; sp|Q40392.1|) (E-value < e^{-34}) que codifica a proteína N relacionada à resistência ao patógeno viral TMV (*Tobacco Mosaic Virus*).

O produto do gene N interage com a replicase do TMV segundo o modelo clássico gene-a-gene descrito por Flor (1971) (Whitham *et al.*, 1996). O gene N é dominante e confere resposta de hipersensibilidade, conforme descrito por Holmes (1938). Seguida da resposta de hipersensibilidade, ocorre à morte celular, que aparentemente funciona como fator delimitante ao vírus, evitando sua proliferação além do sítio de infecção (Erickson *et al.*, 1999).

A sequência de aminoácidos correspondentes à proteína N codificada por esse gene apresenta três domínios: um domínio de ligação a nucleotídeos (NBS), um domínio com uma região rica em leucina (LRR), que provavelmente relaciona-se com a interação proteína-proteína, além do domínio amino-terminal (TIR), que possui similaridade com a proteína *Toll* e com a interleucina-1, que são relatadas como participantes da interação proteína-proteína na via de transdução de sinais em insetos e animais, respectivamente (Whitham *et al.*, 1994; Dinesh-Kumar & Baker, 2000; Dinesh-Kumar *et al.*, 2000).

Esses domínios, também foram observados na sequência do BAC a partir de comparações com sequências depositadas no Banco de Dados do NCBI utilizando as ferramentas Blastx e Blastp. Os resultados do Blastx mostraram alta similaridade com uma proteína (TIR) de resistência de *Medicago truncatula* (GenBank ref. gb|ABD28703.1|; gi|87162908|), espécie também pertencente à família das leguminosas. Análises pelo Blastn e sequências GSS mostram, em contrapartida, alta similaridade com fragmentos de DNA do gênero *Arachis*, aumentando a confiança na montagem correta do BAC.

Setas representando elementos repetitivos, encontrados com o auxílio do Programa Staden Package, identificaram claramente duas ORFs ligadas a proteínas de resistência no Contig 1 (Gráfico 2), identificando possíveis regiões em tandem, como a porção TIR, a porção NBS e a porção LRR, indicando que genes de resistência podem estar na forma de *clusters* (agrupamentos). Outro exemplo localizado no Contig 4 (Gráfico 5) identificou uma grande região que se repetia no início da ORF 1 e no final da sequência deste mesmo contig, em um local com alta similaridade a uma proteína de resistência, indicando a presença de dois genes contendo sequências que podem ser domínios em tandem. O sequenciamento total do BAC poderá auxiliar na caracterização completa desses análogos a genes de resistência e suas relações. As demais setas identificando regiões repetitivas não estavam relacionadas a regiões de resistência.

Elementos transponíveis foram encontrados em todos os organismos estudados até o momento e são importantes constituintes do genoma vegetal. A principal característica dos elementos transponíveis é a capacidade de se moverem dentro do genoma, inserindo-se em novos sítios próximos ou até mesmo dentro de sequências gênicas, o que pode causar mutações, alterações de estrutura e função de genes, rearranjos cromossômicos, mudanças na regulação gênica, podendo dessa forma servir como fonte de diversidade (Kumar & Bennetzen, 1999). Existem duas classes principais de elementos transponíveis: (Classe I) os retrotransposons e (Classe II) os transposons. Um retrotransposon pode ser definido como um elemento genético que se move no interior da célula via um intermediário de mRNA, seguida de transcrição reversa e inserção da cópia de cDNA em um novo sítio do genoma. Transposons são sequência móveis que utilizam a enzima transposase para se moverem de um local para outro no genoma.

Foram identificadas neste trabalho, quatro ORFs relacionadas a elementos de transposição, sendo uma correspondente a um transposon e três a retrotransposons, todas elas com E-value $< 10 e^{-04}$.

A partir de comparações com o banco de dados GenBank e predição de função através do programa AutoPredgeneset, os retrotransposons identificados pertenciam ao grupo LTR do tipo Ty1-Copia, distinto de outro grupo principalmente em virtude da ordem dos genes. A identificação de sequências repetitivas flanqueando a ORF (representadas por setas flanqueando a ORF 2 situada no

Contig 1 – Gráfico 2) revela o fato destas possivelmente serem as porções LTR do elemento, portanto um elemento inteiro pode ter sido caracterizado. Comparações com ESTs revelam que estes podem estar transcricionalmente pouco ativos neste BAC.

De acordo com os modelos de cromossomo vegetal propostos, a maior parte do genoma é constituída por sequências repetitivas, sendo que os retrotransposons LTR podem constituir cerca de 50% (Schmidt & Heslop-Harrison, 1998; Kumar & Bennetzen, 1999), enquanto os genes perfazem cerca de 5% de todo o genoma, o que torna importante o estudo desses elementos na composição e função de cada organismo.

A existência de elementos transponíveis dentro de genes pode acarretar na perda da função gênica. O BAC subclonado possui vários genes de resistência, porém, intercalados por elementos transponíveis. A co-localização do RGA contido no BAC com um QTL, revela o fato de que esses genes possivelmente podem estar ativos.

Diversos trabalhos utilizando a técnica de fragmentação de genomas ou vetores do tipo BAC visando o sequenciamento total de uma determinada região, têm auxiliado no conhecimento mais detalhado da estrutura e organização genética de diversos organismos. Além de uma análise mais detalhada, esses estudos visam principalmente o desenvolvimento de marcadores, busca de RGAs, busca de SSR derivados de ESTs através do sequenciamento das extremidades do BAC, alinhamentos entre mapas genéticos e caracterização de sequências.

Pesquisas que visam à caracterização de uma sequência específica de interesse, além do desenvolvimento de marcadores moleculares, têm tido grande êxito quando combinadas estratégias de sequenciamento de extremidades de BACs e sequenciamento do genoma através de *shotgun* (Van Wijk *et al.*, 2007; Liu *et al.*, 2009). Essas estratégias, juntas, podem proporcionar, ainda, a anotação ou re-anotação de mapas físicos, como é o caso de um mapa desenvolvido para soja (*Glycine max*) em um trabalho realizado recentemente por Saini *et al.* (2008).

Em Lourenço (2004) através da fragmentação randômica de DNA (*shotgun*), um método de sonicação (fragmentação mecânica), obteve-se 0,5% do genoma de *Eucalyptus grandis* (640 milhões de pb) sequenciado. Neste trabalho foram identificados 310 elementos repetitivos, 986 microssatélites, além da identificação de

vários genes putativos comparando-se com bancos de dados de ESTs e banco de dados não-redundante de proteínas do NCBI. Outro exemplo foi um estudo feito por Gentzmittel *et al.*, (2002), um EST homólogo a *Helianthus annuus* L. (girassol) (HaELP1) foi obtido a partir de sequências consenso para os genes AtELP de *Arabidopsis thaliana*. Três BACs positivos para uma hibridização feita com uma sonda construída com HaELP1 foram identificados. A subclonagem e sequenciamento de um BAC demonstraram a presença de dois diferentes locos putativamente homólogos aos genes que codificam proteínas transmembrânicas com EGF e domínios de integrina, além de auxiliarem estudos relacionados ao mapeamento físico do genoma de Girassol.

Hibridização *in situ* com fluorescência (FISH) utilizando os BACs positivos como sondas, permitiram a localização cromossomal de insertos de DNA homólogos a genes humanos em duas espécies caninas (Klukowska *et al.*, 2004). A utilização de microssatélites associados a genes pode acelerar a identificação de genes candidatos para características fenotípicas e estudos de ligação (Klukowska *et al.*, 2004).

Para o melhoramento genético, a utilização de genes oriundos de espécies selvagens nem sempre é acompanhada de ganhos agronômicos, uma vez que a introdução de genes utilizando hibridizações e repetidos retrocruzamentos pode ocasionar o arraste gênico, ou seja, vários genes para características indesejáveis poderiam ser incorporados em conjunto com os genes de interesse. Porém, em um estudo de melhoramento, realizado com sucesso por Holmes (1938), relatou-se que a resistência de *Nicotiana glutinosa* ao TMV era controlada pelo gene N. A resistência ao TMV foi incorporada a *N. tabacum* L. utilizando-se a hibridização interespecífica entre *N. glutinosa* (n = 24) e *N. tabacum* (n = 48) pela espécie anfidiplóide *Nicotiana digluta* e repetidos retrocruzamentos, originando a cultivares Samsoun (Holmes, 1938; Marathe *et al.*, 2002).

Em relação à incorporação de genes de interesse baseada na transformação genética, até hoje, poucos estudos foram realizados com o amendoim. Em 2001, Rohini & Sankara transformaram o amendoim com o gene para quitinase de tabaco. Genes para quitinase de arroz e glucanase de alfafa também foram incorporados em amendoim (Chenault *et al.* 2005), assim como o gene para oxidase oxalato de cevada (Livingstone *et al.* 2005), o gene Cry1EC (Tiwari *et al.*, 2008), além do gene

humano Bc1-xL (Chu *et al.*, 2008). Essa forma de melhoramento, através da transformação, mostrou-se um método efetivo para incorporação de genes no genoma do amendoim.

Estudos com o gênero *Arachis*, também visam a introgressão de genes de resistência de espécies silvestres na espécie cultivada. O uso de seleção assistida por marcadores pode ser bastante útil nesse processo, uma vez que vários estudos, inclusive este, estão sendo realizados para a descoberta e caracterização de genes de interesse agrônomo a serem incorporados em amendoim.

Visando obter a sequência completa dos genes de resistência identificados neste estudo, pode-se desenvolver primers através de *Chromosome Walking*, ou seja, sequenciar as extremidades de uma sequência conhecida até completar totalmente os espaços não sequenciados. Desta forma, o sequenciamento do BAC AD25F09 que contém um RGA, utilizando a técnica de sequenciamento através da fragmentação randômica mostrou-se bastante eficiente, pois possibilitou acessar e descrever aspectos importantes, além de ter mostrado ser uma estratégia alternativa para se obter um panorama geral de parte do genoma de *Arachis*.

Este trabalho pode, ainda, tornar-se importante para a ancoragem do mapa genético com um futuro mapa físico a ser desenvolvido para *Arachis*. O mapeamento físico de *Arachis* possibilitará a localização cromossômica de genes de interesse, auxiliando no processo de melhoramento assistido, além de possibilitar a análise genômica comparativa entre *Arachis* e outras espécies, principalmente às leguminosas.

**CAPÍTULO II - DESENVOLVIMENTO E MAPEAMENTO DE MARCADORES
MICROSSATÉLITES A PARTIR DO BAC AD25F09**

1. INTRODUÇÃO

A identificação de um grande número de marcadores moleculares polimórficos constitui uma etapa inicial na geração de mapas de ligação, pois quanto mais saturado estiver o mapa, maior será a capacidade informativa deste, e conseqüentemente mais rapidamente os dados serão utilizados nos programas de melhoramento (Carneiro & Vieira, 2002). Dentre os marcadores moleculares mais utilizados nos estudos de mapeamento genético estão os marcadores microssatélites ou SSR (*simple sequence repeats*), caracterizados por possuírem repetições em *tandem* de pequenos motivos de sequência com um a seis nucleotídeos, sendo encontrados amplamente distribuídos pelo genoma da maior parte dos eucariotos, embora também presente em procariotos (Litt & Luty, 1989). Acredita-se que o principal mecanismo por trás do surgimento e amplificação destas sequências nos genomas seja o deslize ou o mau-pareamento de motivos durante a replicação. As fitas de DNA separam-se e reassociam-se de forma incorreta, o que poderia gerar cópias de trechos de DNA (alelos) com diferentes tamanhos ou números de repetições de um determinado motivo no próximo ciclo de replicação por meio de inserção ou deleção de uma unidade de repetição (Schlotterer & Tautz, 1992). Em virtude disso, esses marcadores apresentam uma alta taxa de mutação, resultando em uma ampla variação no número de unidades repetidas, o que faz com que marcadores baseados em microssatélites sejam altamente informativos e amplamente utilizados em programas de melhoramento de plantas, em mapeamento genético e em identificação de indivíduos (Goldstein & Schlotterer, 1999).

Evidências demonstram que a distribuição genômica dos microssatélites não é aleatória. Presumivelmente, eles afetam a organização da cromatina, regulação da atividade gênica, recombinação, replicação do DNA, ciclo celular e a maquinaria de reparo (Li *et al.*, 2002). Além disso, tem sido mostrado na literatura que um grande número de microssatélites está localizado em regiões de transcrição do genoma, incluindo ESTs (Morgante *et al.*, 2002).

Segundo Ferreira & Grattapaglia (1998) devido à expressão codominante e o multialelismo, os marcadores microssatélites são os marcadores que possuem o mais elevado conteúdo de informação de polimorfismo (PIC). Em virtude disso, estes

marcadores são muito utilizados em estudos de genética de populações de espécies silvestres, devido a sua confiabilidade, praticidade operacional, robustez e serem mais informativos geneticamente. Outra vantagem desses marcadores é que ocorre a conservação de sítios microssatélites entre espécies relacionadas, tornando possível, em alguns casos, a transferência destes marcadores entre espécies ou mesmo gênero utilizando *primers* heterólogos (Moore *et al.*, 1991)

Nem todos os locos microssatélites são aproveitáveis para o mapeamento genético. A seleção de locos informativos é baseada nas características destes locos e nos resultados que estes microssatélites permitirão obter na identificação de segregação nas progênes resultantes de cruzamentos controlados. Dentre os principais parâmetros utilizados nesta seleção estão: o tipo de repetição (di-, tri-, tetra-, ou pentanucleotídicos), o número de repetições, o número de um mesmo loco no genoma, o número de alelos ao loco, a heterozigosidade e outros (Brondani *et al.*, 1998; Garner, 2002). Regiões microssatélites maiores seriam mais úteis para análise genética possibilitando maior informação por loco, principalmente em estudos envolvendo genética de populações.

A baixa variabilidade genética do amendoim, em relação às características agronômicas de interesse, tem dificultado os avanços no melhoramento molecular da cultura, porém marcadores microssatélites vêm mostrando bastante eficiência em relação às análises feitas com outros marcadores, detectando algum polimorfismo na espécie cultivada do gênero *Arachis* e mais de 900 SSRs foram publicados para o amendoim até hoje (Hopkins *et al.*, 1999; He *et al.*, 2003; Krishna *et al.*, 2004; Ferguson *et al.*, 2004; Moretzsohn *et al.*, 2004; Gimenes *et al.*, 2007; Tang *et al.*, 2007; Cuc *et al.*, 2008). Além disso, a taxa de transferibilidade de microssatélites de *A. hypogaea* para espécies silvestres da seção *Arachis* tem sido relativamente alta (Moretzsohn *et al.*, 2005; Gimenes *et al.*, 2007), o que permite seu uso para a análise das relações genéticas entre essas espécies.

Em 2005, Moretzsohn e colaboradores publicaram um mapa de ligação baseado em uma população de 93 indivíduos F₂ obtidos do cruzamento entre duas espécies silvestres de genoma AA, *A. duranensis* e *A. stenosperma* e um total de 170 marcadores microssatélites foi mapeado. Este mapa vem sendo aprimorado e, atualmente possui aproximadamente 188 marcadores microssatélites. Recentemente foi construído um mapa genético para o amendoim cultivado também

baseado em microssatélites, do qual 135 marcadores já puderam ser mapeados (Varshney *et al.*, 2008).

A subclonagem e sequenciamento do BAC AD25F09 que contém o RGA S1_A_36 co-localizado com um QTL para *C. personatum* mapeado no mapa de referência em *Arachis*, além de ter sido utilizada para a caracterização deste BAC, foi também utilizada para o desenvolvimento de marcadores microssatélites. O objetivo deste estudo consiste no mapeamento desses marcadores SSR além de identificar aqueles mais próximos ao RGA em questão, visando à seleção assistida por marcadores.

2. MATERIAL E MÉTODOS

2.1 Material vegetal e população de mapeamento

Os experimentos foram conduzidos no Laboratório de Biotecnologia (Pós-graduação da Universidade Católica de Brasília) e Embrapa Recursos Genéticos e Biotecnologia. As plantas utilizadas neste trabalho foram fornecidas pelo Banco Ativo de Germoplasma (BAG) de Espécies de *Arachis* situado na Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF.

Todos os procedimentos realizados neste Capítulo, com intuito de gerar marcadores microsatélites a partir do BAC AD25F09, estão exemplificados na Figura 1.

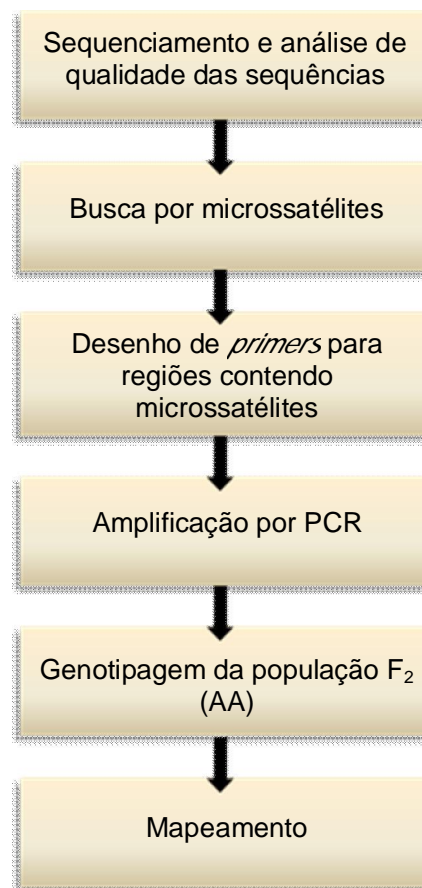


Figura 1: Esquema de geração de marcadores microsatélites a partir BAC AD25F09.

A população de mapeamento utilizada foi obtida na Embrapa Recursos Genéticos e Biotecnologia (Moretzsohn *et al*, 2005) através do cruzamento entre duas espécies silvestres: *A. duranensis* K7988 e *A. stenosperma* V10309. As duas espécies possuem o genoma AA, sendo que *A. stenosperma* possui resistência ao nematóide 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 um híbrido F₁, que foi autofecundado para gerar a população de mapeamento F₂ com 93 indivíduos.

2.2 Extração do DNA genômico de *Arachis* e desenho de *primers* para microsatélites

Para os experimentos realizados neste trabalho, utilizou-se o DNA genômico dos parentais e F₁ da população de mapeamento AA (Moretzsohn *et al*, 2005); os parentais da população de mapeamento BB (Moretzsohn *et al*, 2009) e quatro indivíduos tetraplóides, sendo dois deles sintéticos (Tabela 1). Para a genotipagem foi utilizado também os DNAs da população F₂ de 93 indivíduos da população A.

Tabela 1: Lista de espécies do gênero *Arachis* selecionadas para extração de DNA genômico para utilização em PCR. A Tabela mostra o tipo de genoma referente a cada espécie.

Espécies	Genoma
<i>Arachis duranensis</i> (V14167)	AA
<i>Arachis duranensis</i> (K7988)	AA
<i>Arachis stenosperma</i> (V10309)	AA
F1 - <i>Arachis duranensis</i> (K7988) x <i>Arachis stenosperma</i> (V10309)	AA
<i>Arachis magna</i> (KG30097)	BB
<i>Arachis ipaënsis</i> (KG30076)	BB
<i>Arachis hypogaea</i> cv. IAC-Runner 886	AABB
<i>Arachis hypogaea</i> cv. IAC-Caiapó	AABB
(<i>Arachis ipaënsis</i> (KG30076) x <i>Arachis duranensis</i> (V14167)) ^c	AABB
(<i>Arachis gregoryi</i> (V6389) x <i>Arachis linearifolia</i> (V9401)) ^c	AABB

O DNA total foi extraído de folhas jovens (recém surgidas e antes da expansão foliar), para evitar a extração simultânea de compostos fenólicos, polissacarídeos e proteínas presentes em folhas adultas. Um protocolo de extração

de DNA baseado em CTAB publicado por Ferreira & Grattapaglia, (1998) foi modificado para uso em *Arachis* sp. Aproximadamente 200 mg de tecido vegetal foram macerados (com o auxílio de *beads*) juntamente com 700 µL de 2x CTAB e 2 µL e 2-mercaptoetanol em cada tubo de microcentrífuga de 2 mL. As amostras foram incubadas a 65° C por 60 minutos. Após esse período foram adicionados 700µL de clorofórmio-isoamílico (24:1) para cada amostra e misturou-se até formar uma emulsão. As amostras foram centrifugadas a 14.000 rpm por 5 minutos e a fase aquosa foi transferida para dois tubos novos de 1,5 mL. Foram adicionados 600 µL de tampão 1x CTAB e os tubos foram agitados lentamente e centrifugados a 14.000 rpm por 1 minuto. O sobrenadante foi descartado e o precipitado dos dois tubos foi ressuspenso em 300 µL de NaCl 1,2 M. O volume dos dois tubos foi transferido para um único tubo, totalizando 600 µL. As amostras foram centrifugadas a 12.000 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 13.200 rpm por 2 minutos e o sobrenadante foi descartado. O *pellet* foi lavado duas vezes com 500 µL de etanol 70% e posteriormente seco e ressuspenso em 100 µL de água estéril com 0,01 mg/mL de RNase-A. As amostras foram quantificadas em gel de agarose 1,5% utilizando um marcador de peso molecular (*High Mass Ladder* - Invitrogen). As amostras foram diluídas para 5 ng/µL e mantidas a -20° para realização do experimento.

A busca por sequências contendo microssatélites no BAC AD25F09 foi realizada utilizando o algoritmo *Tandem Repeat Occurrence Locator* (TROLL). Este programa é baseado no algoritmo de *Aho Corasick* que utiliza uma estratégia denominada “*dictionary approach*”, onde as sequências repetitivas são conhecidas a priori, isto é, procura-se por regiões microssatélites utilizando-se um dicionário de motivos. O programa TROLL utiliza este algoritmo para procurar padrões pré-selecionados em uma sequência-texto e, acrescentado a isto, procura por estas regiões em tandem, tornando possível a identificação de microssatélites (Castelo *et al.*, 2002).

Depois de selecionadas as regiões contendo os microssatélites, o Programa *Primer3* (Rozen & Skaletsky, 2000), foi utilizado visando o desenvolvimento dos pares de *primers* para estas regiões. Este software identifica sequências de oligonucleotídeos flanqueando as sequências microssatélites identificadas pelo

TROLL obedecendo alguns parâmetros, como: tamanho dos oligonucleotídeos, temperatura de anelamento, diferença de temperatura de anelamento máxima entre os pares de *primers* dentre outras.

2.3 Amplificação por PCR

A reação de PCR foi feita com todos os indivíduos listados na Tabela 1. Para um volume de 25 μL de reação de PCR, foram utilizados: 2,5 μL (10%) de tampão para PCR sem MgCl_2 ; 0,75 μL de MgCl_2 50 mM; 0,5 μL de cada primer a 10 μM ; 0,5 μL de dNTPs a 10 mM, 0,5 μL de *Taq* Polymerase (*Taq Platinum* Invitrogen) e 1 μL de DNA (20 ng/ μL). O programa utilizado foi: 5 minutos a 94° C; 30 ciclos de 1 minuto a 94° C, 1 minuto a 50-56° C (dependendo do primer), 1 minuto a 72° C; 7 minutos a 72° C.

Os produtos de PCR foram inicialmente analisados em gel de agarose 1% corados com brometo de etídeo e visualizados sobre luz ultravioleta. Com a confirmação do funcionamento correto da reação, estes produtos de PCR foram posteriormente analisados em gel de poliacrilamida corado com nitrato de prata, de acordo com o protocolo descrito por Creste *et al.* (2001). O tamanho dos alelos foi estimado por comparação com marcador DNA padrão de 10 pb (Gibco), e a presença de polimorfismo para cada par de *primer* foi anotada.

2.4 Mapeamento de microssatélites

Cada par de *primers* que produziu polimorfismo entre as duas plantas parentais foi utilizada para genotipar toda a população F_2 (93 indivíduos) juntamente com seus parentais e a planta F_1 . As análises foram realizadas atribuindo-se 1 para a presença da banda e 0 para ausência da banda, em cada um dos 96 indivíduos analisados. O teste de qui-quadrado (X^2) foi utilizado para testar a hipótese nula de segregação 3:1 destes marcadores.

As análises de ligação foram realizadas através do programa Mapmaker, versão 2.0 (Lander *et al.*, 1987), como descritas por Moretzsohn *et al.* (2005). Uma matriz de genotipagem composta pelos marcadores polimórficos neste trabalho foi inserida na matriz contendo mais 716 marcadores polimórficos genotipados

anteriormente (José, 2006; Moretzsohn *et al.*, 2006; Alves, 2008). Um LOD escore mínimo de 3.0 e uma fração de recombinação máxima (r) de 0,35 foram utilizados como parâmetros para estabelecimento dos grupos de ligação usando a função “group”. O comando “try” foi utilizado para determinar a posição exata dos novos marcadores dentro de cada grupo. As frações de recombinação foram convertidas em distancia de mapa, Centimorgans (cM), usando a função de mapeamento de Kosambi, que leva em consideração o coeficiente de Interferência (I).

3. RESULTADOS

3.1 Extração de DNA e qualidade das amostras

O protocolo de extração de DNA desenvolvido por Ferreira & Grattapaglia (1998), modificado neste trabalho, atendeu às necessidades requeridas para os experimentos, entretanto houve degradação em algumas amostras, mostrando que devem existir compostos que atuam neste processo nas folhas de *Arachis* sp. Todos os DNAs dos indivíduos exemplificados na Tabela 1 foram quantificados com o auxílio de um espectrofotômetro (Nanodrop ND1000) e a Tabela 2 mostra os valores das concentrações obtidas neste trabalho.

A Figura 2 mostra um gel de agarose 1% corado com brometo de etídeo, contendo amostras de DNA genômico de algumas espécies do gênero *Arachis* para uso em PCR. A análise de qualidade do DNA foi feita através de gel de agarose, aliada a análise de quantificação com auxílio de um espectrofotômetro (Nanodrop ND1000), permite a avaliação da integridade da molécula de DNA. O marcador High Mass (Promega) também foi utilizado como controle da quantidade de DNA das amostras no gel.

Tabela 2: Quantificação de DNA genômico de espécies de *Arachis* em Nanodrop.

Espécies	Quantificação (ng/ μ L)
1- <i>Arachis duranensis</i> (V14167)	480
2- <i>Arachis duranensis</i> (K7988)	470
3- <i>Arachis stenosperma</i> (V10309)	250
4- F1 - <i>Arachis duranensis</i> (K7988) x <i>Arachis stenosperma</i> (V10309)	340
5- <i>Arachis magna</i> (KG30097)	900
6- <i>Arachis ipaënsis</i> (KG30076)	720
7- <i>Arachis hypogaea</i> cv. IAC-Runner 886	320
8- <i>Arachis hypogaea</i> cv. IAC-Caiapó	380
9- (<i>Arachis ipaënsis</i> (KG30076) x <i>Arachis duranensis</i> (V14167)) ^c	980
10- (<i>Arachis gregoryi</i> (V6389) x <i>Arachis linearifolia</i> (V9401)) ^c	1200

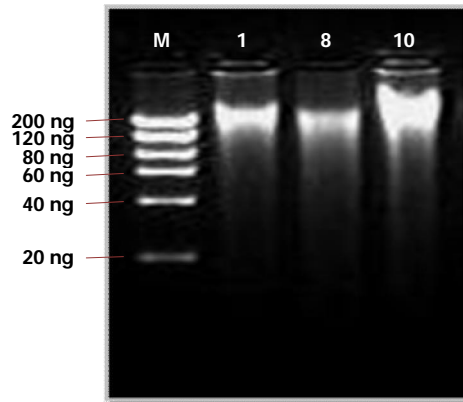


Figura 2: Gel de agarose 1% corado com brometo de etídeo referente à quantificação de três amostras de DNA (Amostras 1, 8 e 10) junto com o marcador molecular High Mass Ladder (Invitrogen).

3.2 Desenho de *primers* para regiões contendo microssatélites e amplificação destas sequências

Muitas regiões contendo microssatélites com repetições mono foram identificadas (denominadas poli(n)), porém apenas três delas foram utilizadas para confecção de *primers*. Sete microssatélites compostos por di-nucleotídeos variando entre repetições de cinco a 24 vezes foram encontrados. Um microssatélite tri e dois compostos por tetra-nucleotídeos, todos eles contendo quatro repetições, também foram identificados. A Tabela 3 mostra a sequência dos 13 pares de *primers* obtidos neste trabalho, o SSR amplificado por eles e o tamanho esperado dos produtos.

Para identificar marcadores microssatélites polimórficos associados ao RGA, os 13 pares de *primers* foram testados com os genótipos parentais e F₁ da população A; genótipos parentais da população B; e ainda genótipos tetraplóides, de acordo com a Tabela 1.

Uma análise preliminar para constatar a especificidade e o funcionamento dos primers na PCR, assim como as temperaturas ideais de anelamento para cada primer e ainda o perfil de polimorfismo observado foram realizadas. Em gel de agarose 2% corado com Brometo de etídeo, no entanto, a resolução de fragmentos em géis de agarose não permitiu distinguir possíveis heterozigotos com precisão ou alelos próximos em tamanho. Desta forma visando identificar possíveis polimorfismos, houve a necessidade de uma segunda etapa de análise realizada em

géis de poliacrilamida a 4% em condições desnaturantes com todos os pares de *primers* e os dez indivíduos listados na Tabela 2.

Tabela 3: *Primers* utilizados para amplificação de microssatélites.

Nome do <i>Primer</i>	<i>Primers</i> 5' - 3'	SSR/poli(n)	Produto (pb)
AD25F09-PB-H01	<i>Primer</i> FWD - TTTTCTGGGTAATCATTTCATGC <i>Primer</i> REV - TCTCTCAATTCTTTTGTAAGTGTGA	(TA) ⁷	364
AD25F09-HIII-P1E08	<i>Primer</i> FWD - CTCTTCGATTGCTTTGAGGAGT <i>Primer</i> REV - TACCAAACCCCTTCACTACCCAG	(GT) ⁹	347
AD25F09-PC-H08	<i>Primer</i> FWD - GGGTCAGATTTATCGGTGAAAA <i>Primer</i> REV - ATGAGAAGAAGAAGCGGCATAG	(TC) ⁹	363
AD25F09-PA-D03	<i>Primer</i> FWD - ACTAAAATGGACATCAACTTAGA <i>Primer</i> REV - TTGAGAGTACTAGAATAAGAGTACGA	(TA) ¹⁵	302
AD25F09-PD-A10	<i>Primer</i> FWD - CATTTAATTCAAACCATCAGCG <i>Primer</i> REV - GCTGTTCCTTGGCTTTTCTTTA	(AT) ²⁴	351
AD25F09-HIII-34	<i>Primer</i> FWD - ACTCATGGGTTAGAGTCCCTCA <i>Primer</i> REV - CAAGATCGTAGTAAAGGGTCGG	(TC) ⁵	160
AD25F09-HIII-82	<i>Primer</i> FWD - TGGAACTTACAAGAATGGCAA <i>Primer</i> REV - CCAACCCTGTTAGTAGTAGGCG	(A) ¹⁰	194
AD25F09-HIII-100	<i>Primer</i> FWD - TTGGGGTATTTTATCCCCTTTT <i>Primer</i> REV - AGTTCATTAGCCCCTTACCCA	(T) ⁹ (T) ⁹	383
AD25F09-S1A36-HIII-51	<i>Primer</i> FWD - TTGATTCTGTACAACCTCCGGTT <i>Primer</i> REV - GATCAATCTTCTTTCCATGGCT	(T) ⁸	346
AD25F09-HIII-37	<i>Primer</i> FWD - GAAGCAAGGGAAGAACAAGCTA <i>Primer</i> REV - AAACCCAACATGCAATTACCTC	(TCA) ⁴	394
AD25F09-Sau-27	<i>Primer</i> FWD - GAAGAAGAGGCTTACCTACCA <i>Primer</i> REV - TGTTCTTTCCGCAGTGTATG	(TAAT) ⁴	232
AD25F09-Sau-05	<i>Primer</i> FWD - TTTGTTGCCCCATTTAGAAACT <i>Primer</i> REV - ACTCTTATTTGGCATTAGCCCC	(GAAA) ⁴	139
AD25F09-HIII-P6H06	<i>Primer</i> FWD - ACCAACCTAGGAAACACAGTCCCGT <i>Primer</i> REV - AATTCAAACCATCAGCGATAAA	(TA) ²⁴	195

3.3 Análise dos marcadores microssatélites e mapeamento

Sete pares de *primers* (AD25F09-PC-H08; AD25F09-PA-D03; AD25F09-HIII-100; AD25F09-S1A36-HIII-51; AD25F09-HIII-37; AD25F09-Sau-27 e AD25F09-Sau-05) apresentaram um resultado aparentemente monomórfico em todos os dez indivíduos, como exemplificado nas Figuras 3A e 3B.

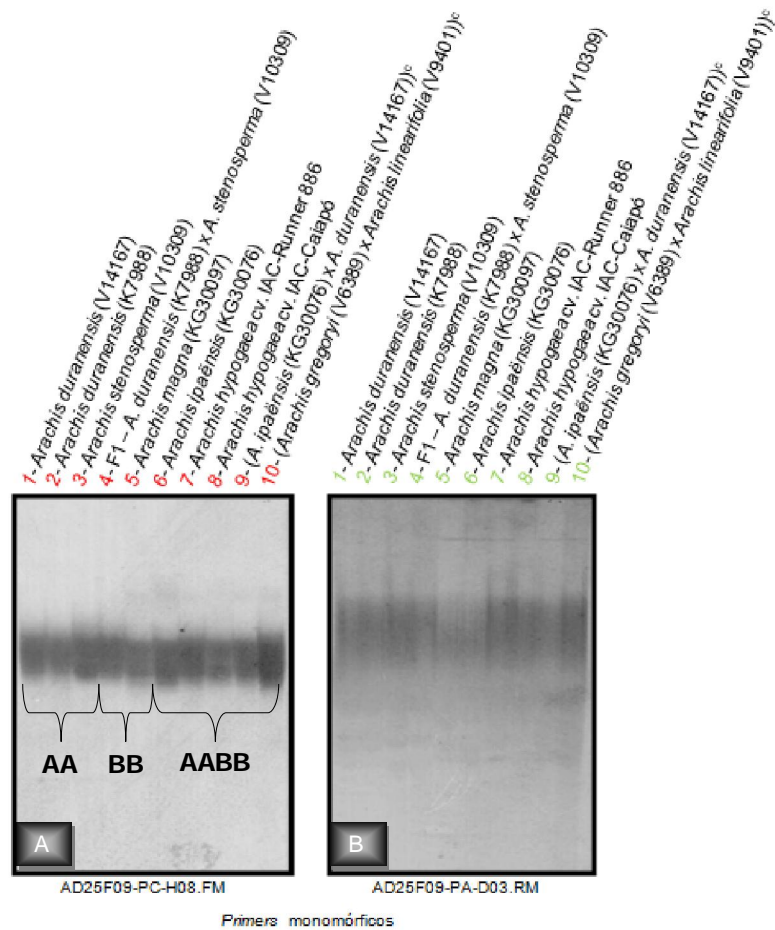


Figura 3: Géis de poliacrilamida a 4%. (A) Primer AD25F09-PC-H08; (B) Primer AD25F09-PA-D03.

Seis pares de *primers* (AD25F09-HIII-82; AD25F09-HIII-34; AD25F09-HIII-P6H06; AD25F09-PB-H01; AD25F09-PD-A10; AD25F09-HIII-P1E08) identificaram um padrão polimórfico, como exemplificado nas Figuras 4A e 4B respectivamente.

Destes seis pares de *primers*, cinco amplificaram gerando um padrão polimórfico de presença ou ausência de bandas em um parental ou outro da população A (*A. duranensis* K7988 ou *A. stenosperma* V10309), enquanto que o indivíduo F₁ sempre amplificava de forma semelhante a um dos parentais (Figura 5A-E). Estes cinco pares de *primers* foram utilizados para genotipar a população F₂ de 93 indivíduos, com o intuito de mapear essas regiões polimórficas no mapa genético de referência.

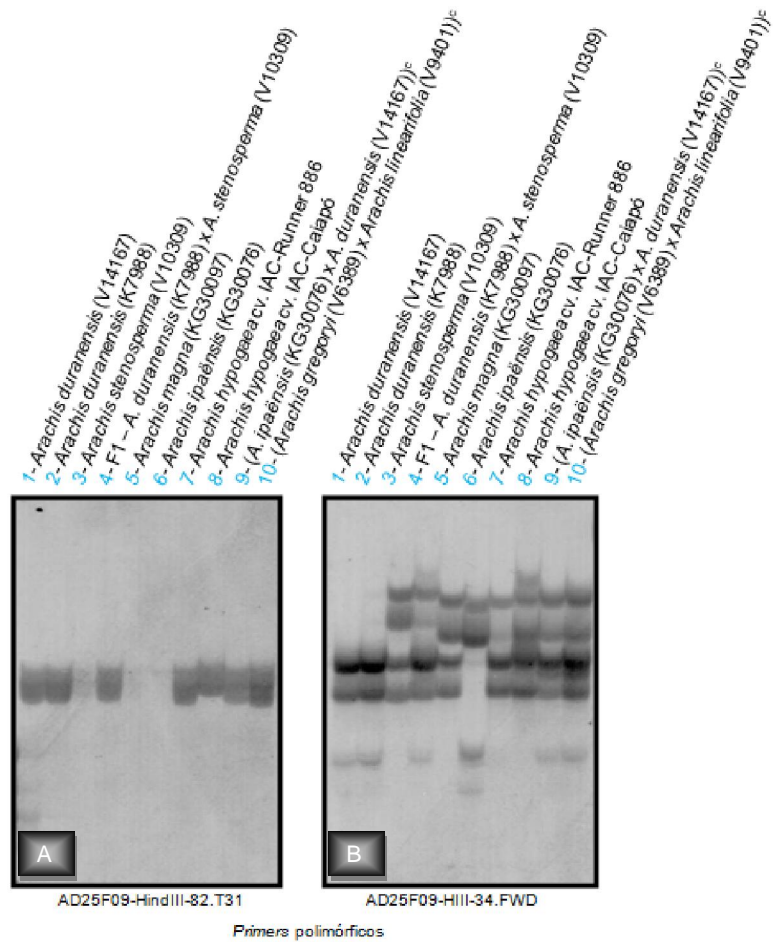


Figura 4: Géis de poliacrilamida a 4%. (A) Primer AD25F09-HIII-82; (B) Primer AD25F09-HIII-34.

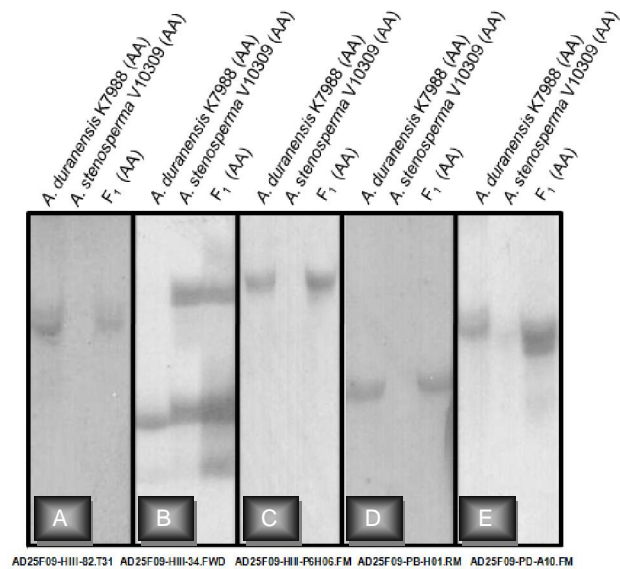


Figura 5: Géis de poliacrilamida a 4% mostrando a amplificação com cinco pares de *primers* nos parentais da população A e F1 resultante deste cruzamento. (A) AD25F09-HIII-82; (B) AD25F09-HIII-34; (C) AD25F09-HIII-P6H06; (D) AD25F09-PB-H01; (E) AD25F09-PD-A10.

Observou-se um perfil de amplificação distinto entre os primers testados dependendo dos parentais e da população testada, Os parentais da população B não amplificaram com o par de primers AD25F09-HIII-82, porém este apresentou um polimorfismo com perfil dominante na população A e um perfil de polimorfismo de tamanho nas duas diferentes cultivares de amendoim.

Quatro pares de *primers*, dos cinco genotipados, juntamente com o outro par AD25F09-HIII-P1E08, apresentaram polimorfismos de tamanho para diferentes cultivares de *Arachis hypogaea* como demonstrado na Figura 6 A-E.

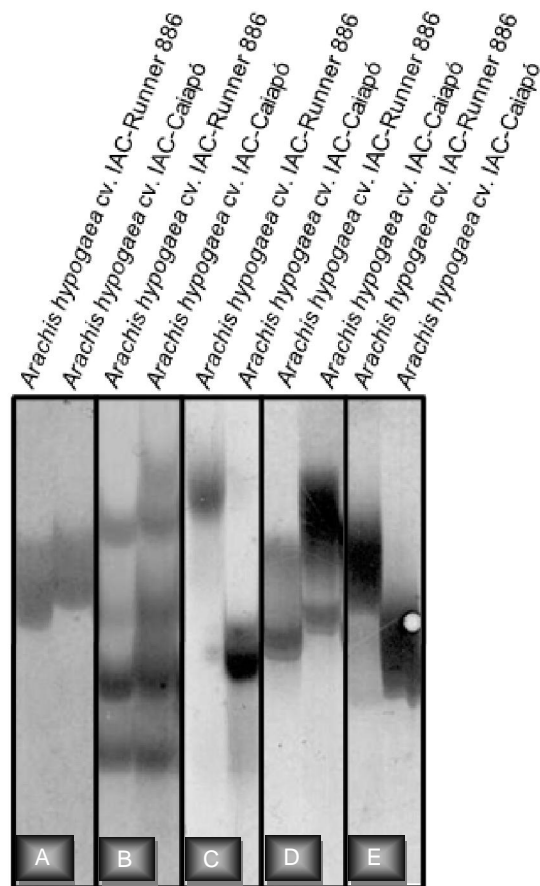


Figura 6: Géis de poli-acrilamida a 4% mostrando a amplificação com cinco pares de *primers* nas duas cultivares de *A. hypogaea* (IAC-Runner 886 e IAC-Caiapó). (A) AD25F09-HIII-82; (B) AD25F09-HIII-34; (C) AD25F09-PB-H01; (D) AD25F09-PD-A10; (E) AD25F09-HIII-P1E08.

Os cinco pares de *primers* polimórficos para população A, genotipados na população F₂, foram posteriormente analisados com auxílio do programa MapMaker, juntamente com os outros marcadores já mapeados nessa população. Utilizando um LOD score mínimo de 3,0 e uma fração de recombinação máxima de

0,35, os cinco marcadores puderam ser mapeados de forma agrupada no Grupo de Ligação 4, como mostrado na Figura 7.

Grupo de Ligação 4

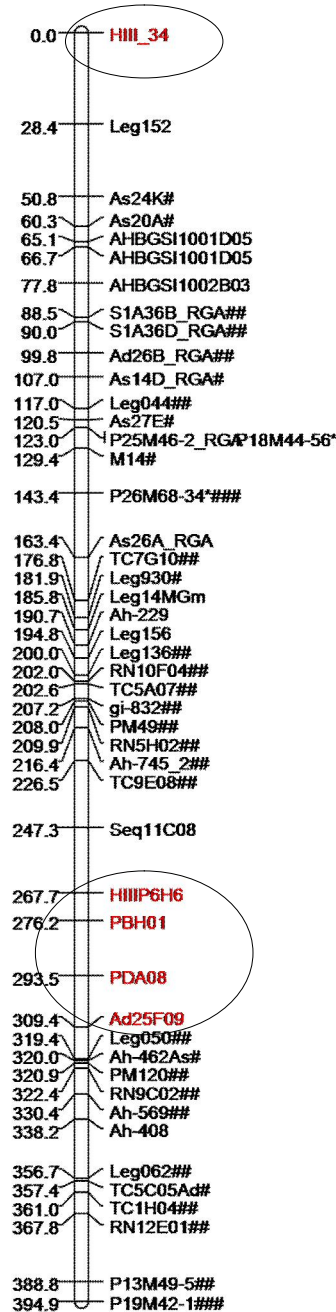


Figura 7: Grupo de Ligação 4 pertencente ao Mapa Genético de *Arachis* (Moretzsohn *et al.*, 2005). Circulados estão destacados os marcadores SSRs desenvolvidos neste estudo.

Os marcadores SSR (e poli(n)) identificados neste trabalho, estavam presentes em seis dos oito contigs gerados (Tabela 4). Dois dos marcadores mapeados localizavam-se em regiões que codificam proteínas da família TIR-NBS-LRR relacionadas à resistência (em vermelho). Estes marcadores são mostrados como círculos nos Gráficos 2-9, onde a cor verde significou polimorfismo e cinza ausência de polimorfismo.

Tabela 4: Localização de microssatélites nos *contigs* do BAC AD25F09 e resultado da amplificação em alguns genótipos do gênero *Arachis*.

Nome do <i>Primer</i>	Localização do SSR	Resultado da amplificação
AD25F09-Sau-05	Contig 1	Monomórfico para todos os indivíduos testados
<u>AD25F09-HIII-34</u>		Polimórfico para pop A e <i>A. hypogaea</i>
<u>AD25F09-HIII-82</u>		Polimórfico para pop A e <i>A. hypogaea</i>
AD25F09-PB-H01	Contig 2	Polimórfico para pop A e <i>A. hypogaea</i>
AD25F09-PD-A10		Polimórfico para pop A
AD25F09-S1A36-HIII-51		Monomórfico para todos os indivíduos testados
AD25F09-HIII-P6H06		Polimórfico para pop A e <i>A. hypogaea</i>
AD25F09-Sau-27	Contig 3	Monomórfico para todos os indivíduos testados
AD25F09-PA-D03		Monomórfico para todos os indivíduos testados
AD25F09-PC-H08		Monomórfico para todos os indivíduos testados
AD25F09-HIII-P1E08	Contig 5	Polimórfico para <i>A. hypogaea</i>
AD25F09-HIII-37	Contig 6	Monomórfico para todos os indivíduos testados
AD25F09-HIII-100	Contig 7	Monomórfico para todos os indivíduos testados

* Em negrito e sublinhado estão os marcadores associados a genes de resistência.

4. DISCUSSÃO

O presente estudo visou à busca de microssatélites em biblioteca de subclones do BAC AD25F09 provenientes de sequenciamento por fragmentação randômica. A partir dessas sequências foram identificados 13 microssatélites para os quais foi possível desenhar pares de *primers* a serem utilizados para mapeamento dos mesmos. Em virtude da baixa quantidade de SSRs encontrados, praticamente todos esses locos foram utilizados para confecção de *primers*, incluindo repetições mono, denominadas de poli(n).

Estes pares de *primers* foram testados em dez genótipos representantes do gênero *Arachis*, sendo eles: genótipos diplóides representantes da população A: *A. duranensis* (acesso V14167), *A. duranensis* (acesso K7988), *A. stenosperma* (acesso V10309) e F₁ resultante do cruzamento destes dois últimos; genótipos diplóides representantes da população B: *Arachis magna* (acesso KG30097) e *A. ipaënsis* (acesso KG30076); genótipos tetraplóides (AABB) de dois acessos do amendoim: *A. hypogaea* cv. IAC-Runner 886 e *A. hypogaea* cv. IAC-Caiapó; e dois anfidiplóides sintéticos: *A. ipaënsis* (acesso KG30076) x *A. duranensis* (acesso V14167))^c e (*A. gregoryi* (acesso V6389) x *A. linearifolia* (acesso V9401))^c, estes últimos importantes para a análise comparativa com os genótipos tetraplóides verdadeiros. O mapeamento dos genótipos *A. duranensis* e *A. ipaënsis* são também de interesse, pois estes são os prováveis parentais do amendoim (Halward *et al.*, 1991; Kochert *et al.*, 1996; Seijo *et al.*, 2004; 2007). Os parentais selvagens possuem um maior polimorfismo molecular e são ricos em alelos que conferem características de interesse, como genes de resistência a doenças e fatores abióticos (Nelson *et al.*, 1990; Stalker & Simpson, 1995; Burow *et al.*, 2001; Kameswara *et al.*, 2003; Moretzsohn *et al.*, 2004).

Sete pares de *primers* desenvolvidos neste trabalho (AD25F09-PC-H08; AD25F09-PA-D03; AD25F09-HIII-100; AD25F09-S1A36-HIII-51; AD25F09-HIII-37; AD25F09-Sau-27; AD25F09-Sau-05) apresentaram um padrão monomórfico para os todos os genótipos representantes das populações A, B e tetraplóides que foram testados, não sendo utilizados neste trabalho para fins de genotipagem ou mapeamento molecular. Os anfidiplóides sintéticos testados apresentaram um perfil

monomórfico para todos os pares de *primers* desenvolvidos, o que não gerou informações relevantes.

Seis pares de *primers* apresentaram-se polimórficos. São eles: AD25F09-HIII-82; AD25F09-HIII-34; AD25F09-HIII-P6H06; AD25F09-PB-H01; AD25F09-PD-A10 e AD25F09-HIII-P1E08. Apesar dos marcadores microssatélites serem co-dominantes, os cinco primeiros marcadores citados apresentaram um comportamento dominante quando testados na população A e F₁, onde houve a amplificação de alelos nulos em um dos parentais. Esse padrão dominante proporcionou a genotipagem da população F₂ composta por 93 indivíduos e com o auxílio do Programa MapMaker, os cinco microssatélites puderam ser mapeados de forma agrupada na região inferior do GL 4 do mapa genético de referência AA (Moretzsohn *et al.*, 2005), porém distantes do RGA S1_A_36 (mapeado na região superior, junto com outros RGAs).

O mapeamento agrupado desses marcadores ressalta a confiabilidade desses resultados, uma vez que foram obtidos de um BAC com inserto de 110.000 pb e situavam-se fisicamente próximos uns dos outros (através de análises feitas pelo sequenciamento do BAC realizado no Capítulo I). Contudo, o mapeamento distante geneticamente do RGA que os originou é compatível com o fato de que os RGAs pertencem a uma família multigênica, estando presentes em múltiplas cópias em diferentes locais do genoma, o que ressalta também, a necessidade da comparação do mapa genético com um mapa físico (em construção), com intuito de visualizar as distâncias físicas verdadeiras em pares de base que eles foram posicionados, em relação aos outros marcadores e QTLs mapeados.

Os marcadores SSR (e poli(n)) identificados neste trabalho, estavam presentes em seis dos oito contigs gerados, representando dessa forma, a maior parte da sequência do BAC. Dois dos marcadores mapeados localizavam-se em regiões que codificam proteínas da família TIR-NBS-LRR relacionadas à resistência. A grande vantagem deste trabalho é que ao se mapear estes microssatélites, também podem ser mapeados possíveis genes de resistência (RGAs) no mapa genético. A utilização de marcadores microssatélites exibe vantagens adicionais à utilização de RGAs no processo de seleção assistida, como alta reprodutibilidade, alto conteúdo informativo, por ser baseado em PCR, e principalmente por ser um marcador transferível entre espécies relacionadas. A identificação de marcadores mais robustos e transferíveis associados a locos que conferem resistência

possibilitará a seleção assistida nos programas de melhoramento genético, o que acelerará o processo de introgressão dos genes de resistência no amendoim.

Adicionalmente, quatro pares de primers que serviram para o mapeamento de SSRs juntamente com o par de primer AD25F09-HIII-P1E08, apresentaram polimorfismos de tamanho quando testados nos dois acessos selecionados de *A. hypogaea*. Esse resultado foi bastante interessante, pois a detecção de polimorfismo nesta espécie é bastante difícil devido à estreita base genética da espécie o que conseqüentemente, reduz a probabilidade de seleção assistida por marcadores. A utilização de microssatélites tem demonstrado ser mais eficiente que os demais marcadores moleculares na detecção de polimorfismo e para análises das relações e variações genéticas em várias espécies de plantas, incluído estudos realizados com o amendoim (Hopkins *et al.*, 1999; He *et al.*, 2003; Krishna *et al.*, 2004; Ferguson *et al.*, 2004; Moretzsohn *et al.*, 2005; Gimenes *et al.*, 2007; Tang *et al.*, 2007; Cuc *et al.*, 2008). Por conseguinte, as análises feitas neste trabalho, juntamente com testes futuros utilizando os pares para SSR em mais acessos de *A. hypogaea*, proporcionarão mais informação e marcadores polimórficos robustos para a espécie cultivada, acelerando os estudos genéticos e o melhoramento da espécie.

Futuramente serão necessárias análises com os SSRs desenvolvidos neste trabalho com a população F₂ composta por 93 indivíduos, utilizada para a construção do mapa B, proporcionando novos dados que serão importantes no estudo de sintonia entre esses dois genomas. Uma vez que esses marcadores foram desenvolvidos a partir de uma sequência contendo partes de genes de resistência associados a um QTL de resistência à cercosporiose, com o alinhamento entre mapas A e B, poderíamos identificar regiões associadas à resistência a cercosporiose também no genoma B.

Os programas de melhoramento de *A. hypogaea* têm conduzido a avanços significativos. Entretanto, a inclusão de genes oriundos de espécies selvagens relacionadas possibilitaria o aumento de variabilidade genética para várias características agrônômicas de interesse, principalmente a resistência a doenças que afetam o amendoim. Com o desenvolvimento de mapas para o gênero *Arachis* utilizando os genomas A e B, que são evolutivamente próximos, é esperado que haja sintonia entre os marcadores e os genes em ambos os genomas. Desta forma, uma vez alinhados os mapas, será possível transferir informações entre eles,

acrescentando assim, dados importantes aos dois mapas. Alguns estudos utilizando marcadores SNPs e SSRs foram realizados recentemente comparando os mapas A, B e espécie cultivada de amendoim (Alves, 2008; Moretzsohn *et al.*, 2009) e de *Arachis* com mapas de leguminosas (Bertioli *et al.*, 2009).

A localização de clones BAC em mapas genéticos, através da identificação de SSRs, por exemplo, possibilitarão a interação entre mapas físicos e genéticos. Marcadores microssatélites identificados a partir do sequenciamento de extremidades de BACs, por exemplo, permitem também estudos acerca da correlação entre mapas genéticos e físicos, como já observado em algumas culturas como a soja (Shultz *et al.*, 2007). Portanto, BACs são excelentes fontes de marcadores microssatélites como demonstrado em diversos trabalhos (Qi *et al.*, 2001; Lichtenzveig *et al.*, 2005; Joobeur *et al.*, 2006).

O desenvolvimento de marcadores microssatélites a partir de sequências do tipo GSS, ESTs e RGAs também tem sido foco de estudos (Cheng *et al.*, 2009; Senthilvel *et al.*, 2008; Sledge *et al.*, 2005; Yi *et al.*, 2006; Di Gaspero *et al.*, 2007).

Em um trabalho realizado por Claverie *et al.* (2004), um clone BAC isolado de uma biblioteca construída para *Prunus cerasifera* L., continha o gene *Ma* para resistência contra nematóides das galhas (*Meloidogyne* spp). Visando a clonagem deste gene, foi realizado o mapeamento de dois marcadores SSR flanqueando o gene.

Enquanto os métodos de melhoramento clássico têm tido grande sucesso ao longo dos anos, a engenharia genética tem o potencial de superar muitas de suas limitações. A identificação de marcadores associados a locos que conferem resistência permitirão a seleção assistida por marcadores nos programas de melhoramento genético do amendoim cultivado, o que acelerará o processo de introgressão dos genes de resistência pelo método de retrocruzamento. Uma vez que serão mapeados genes de resistência tanto em espécies possuidoras de genoma A como em espécies de genoma B, plantas AA com genes de resistência poderão ser cruzadas com plantas BB também resistentes, obtendo-se plantas AB. Estas, após poliploidização, poderão ser cruzadas com *A. hypogaea*, para introgressão desses genes.

CONCLUSÕES GERAIS

- Ambas as metodologias de subclonagem do BAC AD25F09 utilizadas neste trabalho (enzimas de restrição e nebulização) foram satisfatórias e garantiram uma grande porcentagem do BAC sequenciada.
- A análise do conteúdo gênico da sequência do BAC contendo o RGA identificou cinco regiões relacionadas aos genes de resistência pertencentes à família TIR-NBS-LRR, além de um transposon e três retrotransposons LTR do tipo Ty1-copia.
- Cinco marcadores microssatélites puderam ser mapeados de forma agrupada (pois se encontravam no mesmo BAC) no Grupo de Ligação 4 do mapa genético de referência de *Arachis*. O mapeamento destes marcadores derivados do BAC foi uma etapa inicial de mapeamento de outros possíveis genes de resistência (RGAs) à pragas. Mesmo com a identificação de marcadores SSR a partir do RGA, se faz necessária a identificação de outros marcadores mais próximos ao RGA S1_A_36 para a utilização segura no processo de seleção assistida.
- Com a montagem da sequência do BAC, constatou-se que dois dos marcadores mapeados, situavam-se fisicamente em regiões relacionadas a genes de resistência.
- O fato dos microssatélites mapearem em algumas regiões associadas a genes de resistência no mesmo grupo de ligação onde estão mapeados QTLs envolvidos na determinação da resistência à cercosporiose reforça a hipótese de que os genes de resistência à doenças em plantas devem estar organizados em blocos (*clusters*) no genoma.
- Seria importante a comparação do mapa genético com um mapa físico.
- Para S.A.M. será necessário o mapeamento de marcadores mais proximamente ligados aos RGAs.
- A identificação dos vários genes relacionados com a resistência ao fungo foliar *C. personatum* além de outras doenças é fundamental para auxiliar os programas de melhoramento. Assim, espera-se no futuro

que a seleção de genótipos resistentes seja auxiliada pelas informações moleculares, permitindo também a piramidação destes genes, e como consequência a diminuição dos custos para o agricultor com inseticidas e os danos causados ao meio ambiente pela aplicação destes defensivos agrícolas.

PERSPECTIVAS FUTURAS

- Os pares de *primers* para microssatélites desenvolvidos neste trabalho, também serão testados em mais acessos de *A. hypogaea*, o que proporcionará mais informação, além de marcadores polimórficos para a espécie cultivada, acelerando os estudos genéticos e o melhoramento do amendoim.
- Com intuito de completar a sequência total do BAC AD25F09, primers serão desenvolvidos para aplicação da técnica de “Primer Walking”.
- Novos microssatélites encontrados no BAC darão origem a novos pares de *primers*. Além disso, os pares de *primers* desenvolvidos neste trabalho também serão testados na população B visando à obtenção de marcadores a serem posicionados no mapa B de *Arachis*.
- Outros 17 BACs contendo RGAs são objetos de estudos. O BAC Ipa123-A17 (clone da biblioteca BAC de *Arachis ipaënsis*) já foi subclonado pela técnica de nebulização e 1.056 sequências já foram obtidas.
- Três pares de *Primers* para RGAs, mapeados na parte superior do grupo de Ligação 4 (As-14D-1; As-26A-1; P25M46-2) do mapa genético para população A de *Arachis*, foram construídos. Os produtos de PCR já foram clonados em plasmídeos e espera-se que estes, possam servir como sonda para realizar um novo screening nas bibliotecas BAC, com intuito de selecionar outros BACs que contenham RGAs, e então, subcloná-los, caracterizar as sequências e gerar mais marcadores.

REFERÊNCIAS BIBLIOGRÁFICAS

- Aarts MG, te Lintel Hekkert B, 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. 1998 Abr; 11(4):251-258.
- Ahn SC, Baek BS, Oh T, Song CS, Chatterjee B. Rapid mini-scale plasmid isolation for DNA sequencing and restriction mapping. BioTechniques. 2000 Set; 29(3):466-468.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997 Set 1; 25(17):3389-3402.
- Alves DMT, Pereira RW, Leal-Bertioli SCM, Moretzsohn MC, Guimarães PM, Bertioli DJ. Development and use of single nucleotide polymorphism markers for candidate resistance genes in wild peanuts (*Arachis* spp). Genet Mol Res. 2008;7(3):631-42.
- Aubourg S, Rouzé P. Genome Annotation. Plant Physiol. Biochem. 2001;(39):181-193.
- Baker B, Zambryski P, Staskawicz B, Dinesh-Kumar SP. Signaling in plant-microbe interactions. Science. 1997 Maio 2; 276(5313):726-733
- Banerjee D, Zhang X, Bent AF. The Leucine-rich repeat domain can determine effective interaction between RPS2 and other host factors in *Arabidopsis* RPS2-mediated disease resistance. Genetics. 2001;(158):439-450.
- Bell KS, Avrova AO, et al. Sample sequencing of a selected region of the genome of *Erwinia carotovora* subsp. *Atroseptica* reveals candidate phytopathogenicity genes and allows comparison with *Escherichia coli*. Microbiology. 2002;(148):1367-1378.
- Bertioli DJ, Leal-Bertioli SCM, Lion MB, Santos VL, Pappas G, Cannon SB, et al. A large scale analysis of resistance gene homologues in *Arachis*. Mol Genet Genomics. 2003 Out;270(1):34-45.
- Bertioli DJ, Moretzsohn MC, Madsen LH, Sandal N, Leal-Bertioli SCM, Guimarães PM, Hougaard BK, Fredslund J, Schauser L, Nielsen AM, Sato S, Tabata S, Cannon SB, Stougaard J. An analysis of synteny of *Arachis* with *Lotus* and *Medicago* sheds new light on the structure, stability and evolution of legume genomes. BMC Genomics. 2009;10:45.
- Bolonhezi D. Colheita e pós-colheita do amendoim. In: Santos, R. C. Ed. O agronegócio do amendoim no Brasil. 2005;(4):247-261.

- Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 1980 Maio;32(3):314-331.
- Brondani RPV, Brondani C, Tarchini R, Grattapaglia D. Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. *Theor Appl Genet.* 1998;(97):816-827.
- Bruggeman R, Druka A, Kudrna D, Kleinhofs A. Efficient cloning of resistance gene analogs from barley. *Barley Genetics Newsletter*, 2001;(31).
- 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.* 2001 Out;159(2):823-837.
- Calenge F, Van Der Linden CG, Van De Weg, Schouten HJ, Van Arkel, Denace EC, Durel CE. Resistance gene analogs identified through the NBS-profiling method map close to major genes and ATL for disease resistance in apple. *Theor Appl Genet.* 2005; (110):660-668.
- Cannon SB, Zhu H, Baumgarten AM, Spangler R, May G, Cook DR, Young ND. Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. *J Mol Evol.* 2002;(54):548-562.
- Carneiro MS, Vieira MAC. Mapas Genéticos em Plantas. *Bragantia*, Campinas. 2002;(61):89-100.
- Castelo AT, Martins W et al. TROLL-tanden repeat occurrence locator. *Bioinformatics.* 2002;(18):634-636.
- Chenault KD, Melouk HA, Payton ME. Field reaction to Sclerotinia blight among transgenic peanut lines containing antifungal genes. *Crop Sci.* 2005;(45):511-515.
- Cheng X, Xu J, Xia S, Gu J, Yang Y, Fu J, Qian X, Zhang S, Wu J, Liu K. Development and genetic mapping of microsatellite markers from genome survey sequences in *Brassica napus*. *Theor Appl Genet.* 2009;(118):1121-1131
- Chu Y, Deng EX, X. Faustinelli EY, Ozias-Akins P. Bcl-xL transformed peanut (*Arachis hypogaea* L.) exhibits paraquat tolerance. *Plant Cell Rep.* 2008;(27):85-92.
- Clément D, Lanaud C, Sabau X, Fouet O, Le Cunff L, Ruiz E, et al. Creation of BAC genomic resources for cocoa (*Theobroma cacao* L.) for physical mapping of RGA containing BAC clones. *Theor. Appl. Genet.* 2004 Maio;108(8):1627-1634.
- Collins NC, Webb CA, Seah S, Ellis JG, Hubert SH, Pryor A. The isolation and mapping of disease resistance gene analogs in maize. *Mol Plant-Microbe Interact.* 1998;(11):969-978.

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

CONAB – Companhia Nacional de Abastecimento. Acompanhamento da safra brasileira: grãos: sétimo levantamento. 2009 abril;41p. Disponível em: <<http://www.conab.org.br>> Acesso em 05 mai. 2009.

Cordero C, Skinner Z. Isolation from alfalfa of resistance gene analogues containing nucleotide binding sites. *Theor. Appl. Genet.* 2002 Jun;104(8):1283-1289.

Creste S, Neto AT, Figueira A. Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. *Plant Molecular Biology Reporter*. 2001;19(4):299-306.

Creste S, Tsai SM, Valls JFM, Gimenes MA, Lopes CR. Genetic characterization of Brazilian annual *Arachis* species from section *Arachis* and *Heterantheae* using RAPD markers. *Genetic Resources and Crop Evolution*. 2005;(52):1079-1086.

Cuc LM, Mace ES, Crouch JH, Quang VD, Long TD, Varshney RK. Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea*). *BMC Plant Biol.* 2008 ;855.

Danesh D, Peñuela S, Mudge J, Denny RL, Nordstrom H, Martinez JP, Young ND. A bacterial artificial chromosome library for soybean and identification of clones near a major cyst nematode resistance gene. *Theor. Appl. Genet.* 1998;(96):196-202.

Di Gaspero G, Cipriani G, Adam-Blondon A, Testolin R. Linkage maps of grapevine displaying the chromosomal locations of 420 microsatellite markers and 82 markers for R-gene candidates. *Theor. Appl. Genet.* 2007 Maio;114(7):1249-1263.

Dinesh-Kumar SP, Tham WH, Baker BJ. Structure-function analysis of the tobacco mosaic virus resistance gene N. *Proc. Natl. Acad. Sci. U.S.A.* 2000 Dez 19;97(26):14789-14794.

Dinesh-Kumar SP, Baker BJ. Alternatively spliced N resistance gene transcripts: their possible role in tobacco mosaic virus resistance. *Proc. Natl. Acad. Sci. U.S.A.* 2000 Feb 15;97(4):1908-1913.

Eiglmeier K, Wincker P, Cattolico L, Anthouard V, Holm I, Eckenberg R, et al. Comparative analysis of BAC and whole genome shotgun sequences from an *Anopheles gambiae* region related to Plasmodium encapsulation. *Insect Biochem. Mol. Biol.* 2005 Ago;35(8):799-814.

Elgar G, Clark MS, Meek S, Smith S, Warner S, Edwards YJ, et al. Generation and analysis of 25 Mb of genomic DNA from the pufferfish *Fugu rubripes* by sequence scanning. *Genome Res.* 1999 Out;9(10):960-971.

Erickson FL, Dinesh-Kumar SP, Holzberg S, Ustach CV, Dutton M, Handley V, Corr C, Baker BJ. Interactions between tobacco mosaic virus and tobacco N gene. *Philosophical Transactions of the Royal Society of London.* 1999;(354)653-658.

Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* 1998 Mar;8(3):175-185.

Fagundes, M.H. Sementes de amendoim: alguns comentários. 2002. Disponível em: <<http://www.conab.sp.gov>> Acesso em: 16 abr. 2008.

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;165p.

Fávero AP, Simpson CE, Valls JFM, Vello NA. Study of the evolution of cultivated peanut through crossability studies among *Arachis ipaënsis*, *A. duranensis*, and *A. hypogaea*. *Crop Science.* 2006;46(4):1546-1552.

Ferguson ME, Burow MD, Schulze SR, Bramel PJ, Paterson AH, Kresovich S, et al. Microsatellite identification and characterization in peanut (*A. hypogaea* L.). *Theor. Appl. Genet.* 2004 Abr ;108(6):1064-1070

Ferreira ME, Grattapaglia D. Introdução ao uso de marcadores RAPD e RFLP em análise genética. Brasília: Embrapa-Cenargen. 1998 ;(3)220.

Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science.* 1995 Jul 28;269(5223):496-512.

Flor, HH. Current status of gene-for-gene concept. *Annual Review of Phytopathology.* 1971;(9):275-296.

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* into *Arachis hypogaea*. *Genome.* 1996 Out;39(5):836-845.

Garcia GM, Stalker HT, Shroeder E, Lyerly JH, Kochert G. A RAPD-based linkage map of peanut based on a backcross population between the two diploid species *Arachis stenosperma* and *A. cardenasii*. *Peanut Science.* 2005;(32):1-8.

Garner TW. Genome size and microsatellites: the effect of nuclear size on amplification potential. *Genome.* 2002;45(1):212-215.

Gimenes MA, Hoshino AA, Barbosa AVG, Palmieri DA, Lopes CR. Characterization and transferability of microsatellite markers of the cultivated peanut (*Arachis hypogaea*). *BMC Plant Biology.* 2007 Fev 27;7(9).

Godoy IJ, Moraes SA, Pereira JCVNA, Martins ALM, Paulo EM. Cultivares de amendoim: potencial reprodutivo, resistência e resposta ao controle de manchas foliares. *Fitopatologia Brasileira.* 1994;19:288.

Godoy IJ, Moraes SA, Siqueira WJ, Pereira JCVNA, Martins ALM, Paulo EM. Produtividade, estabilidade e adaptabilidade de cultivares de amendoim em três

níveis de controle de doenças foliares. Pesquisa Agropecuária Brasileira. 1999; 34(7):1183-1191.

Godoy IJ, Moares SA, Kasai FS, Minoti D, Nomi AK, Makimoto P. Prevenção da aflatoxina no amendoim. Instituto Agronômico de Campinas. Boletim Técnico. 2001:22p.

Goldstein D, Schlotterer C. Microsatellites: Evolution and Applications. Ed Oxford: Oxford University Press, 1999.

Griffiths AJF et al. Introdução à genética. Rio de Janeiro: Guanabara. 2008.

Guimarães PM, José ACVFJ, Proite K, Bertioli DJ, Leal-Bertioli SCM. Desenvolvimento de marcadores moleculares para análogos a genes de resistência em *Arachis* spp. silvestres. Fitopatologia Brasileira. 2005;(30):663-667.

Guimarães PM, Garsmeur O, Proite K, Leal-Bertioli SCM, Seijo G, Chaine C, et al. BAC libraries construction from the ancestral diploid genomes of the allotetraploid cultivated peanut. BMC Plant Biol. 2008;8:14.

Haldane J. The combination of linkage values and the calculations of distances between the loci of linked factors. Journal of genetics. 1919;(8)299-309.

Halward TM, Stalker HT, Larue EA, Kochert G. Genetic variation detectable with molecular markers among unadapted germ-plasm resources of cultivated peanut and related wild species. Genome. 1991;(34)1013-1020.

Halward TM, Stalker HT, Kochert G. Development of an RFLP Linkage Map in Diploid Peanut Species. Theoretical and Applied Genetics. 1993;87(3):379-384.

Hammond-Kosack KE, Jones JDG. Plant Disease Resistance Genes. Annu. Rev. Plant Physiol. Plant Mol. Biol. 1997 Jun;(48)575-607.

Hammond-Kosack KE, Parker JE. Deciphering plant-pathogen communication: fresh perspectives for molecular resistance breeding. Curr Opin Biotechnol. 2003 Abr;14(2):177-93.

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. 2003;(84):79-88.

Hayes AJ, Saghai Maroof MA. Targeted resistance gene mapping in soybean using modified AFLPs. Theor Appl Genet. 2000;(100)1279-1283.

He G, Meng R, Newman M, Gao G, Pittman R, Prakash CS: Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). BMC Plant Biology. 2003;3(1):3.

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.* 2004 Nov;109(7):1426-1433.

Hilu KW, Stalker HT. Genetic relationships between peanut and wild species of *Arachis* sect *Arachis* (Fabaceae): Evidence from RAPDs. *Plant. Systematics and Evolution.* 1995;198(3-4):167-178.

Holbrook CC, Stalker HT. Peanut breeding and genetic resources. *Plant Breed. Rev.* 2003;(22):297–356.

Holmes FO. Inheritance of resistance to tobacco-mosaic disease in tobacco. *Phytopathology.* Saint Paul, 1938;(28)553-561.

Hopkins MS, Casa AM, Wang T, Mitchell SE, Dean RE, Kochert GD, Kresovich S. Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut. *Crop Science.* 1999;39(4):1243-1247.

Huang X, Madan A. CAP3: A DNA sequence assembly program. *Genome Res.* 1999 Set;9(9):868-877.

Joobeur T, Gusmini G, Zhang X, Levi A, Xu Y, Wehner TC, et al. Construction of a watermelon BAC library and identification of SSRs anchored to melon or Arabidopsis genomes. *Theor. Appl. Genet.* 2006 Maio ;112(8):1553-1562.

José ACVF. Desenvolvimento e Mapeamento de Marcadores Moleculares RGA em Espécies Silvestres de *Arachis*. Dissertação de Mestrado em Ciências Gneômicas e Biotecnologia – Universidade Católica de Brasília – DF; 2008.

Kameswara Rao N, Reddy LJ, Bramel PJ. Potential of wild species for genetic enhancement of some semi-arid food crops. *Genetic Resources and Crop Evolution* 2003;50(7):707-721.

Kanazin V, Marek LF, Shoemaker RC. Resistance gene analogs are conserved and clustered in soybean. *Proc. Natl. Acad. Sci. U.S.A.* 1996 Out 15;93(21):11746-11750.

Kawasaki S, Murakami Y. Genome Analysis of *Lotus japonicus*. *J. Plant. Res.* 2000; 113(1112):497-506.

Kelley JM, Field CE, Craven MB, Bocskai D, Kim UJ, Rounsley SD, et al. High throughput direct end sequencing of BAC clones. *Nucleic Acids Res.* 1999 Mar 15;27(6):1539-1546.

Klukowska J, Szczerbal I, Wengi-Piasecka A, Switonski M, Schelling C, Gmür A, et al. Characterization and mapping of canine microsatellites isolated from BAC clones harbouring DNA sequences homologous to seven human genes. *Anim. Genet.* 2004 Out ;35(5):404-407.

Knauff DA, Moore KM, Gorbet DW. Further studies on the inheritance of fatty acid composition in peanut. *Peanut Science.* 1993;(20)74-76.

Kobe B, Deisenhofer J. The leucine-rich repeat: a versatile binding motif. Trends Biochem. Sci. 1994 Oct;19(10):415-421

Kochert G, Halward T, Branch WD and Simpson CE. RFLP variability in peanut (*Arachis hypogaea*) cultivars and wild species. Theor. Appl. Genet. 1991;(81): 565-570.

Kochert G, Stalker HT, Gimenes M, Galgano L, Lopes CR, Moore K. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). American Journal of Botany. 1996; 83(10):1282-1291.

Kokalis-Burelle N, Rodríguez-Kábana R. Root-Knot nematodes. In: Kokalis-Burelle, N, Porter D.M, Rodríguez-Kábana R, Smith, D.H, Subrahmanyam P. Compendium of peanut diseases. 2.ed. New York: American Phytopathological Society, 1997;50.

Kosambi D. The estimation of map distance from recombination values. Annuaire of Eugenetics. 1944;(12):172-175.

Krapovickas A, Gregory WC. Taxonomia del genero *Arachis* (Leguminosae). Bonplandia (Argentina). 1994;8(1-4):1-186.

Krishna GK, Zhang J, Burow M, Pittman RN, Delikostadinov SG, Lu Y, Puppala N. Genetic diversity analysis in Valencia peanut (*Arachis hypogaea* L.) using microsatellite markers. Cell Mol Biol Lett. 2004;9(4A):685-97.

Kumar A, Bennetzen JL. Plant Retrotransposons. Annual Review of Genetics. 1999;(33)479-532.

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics. 1987;1(2):174-181.

Lavia GI. Karyotypes of *Arachis palustris* and *A. praecox* (Section *Arachis*), two species with basic chromosome number $x = 9$. Cytologia. 1998;63(2):177-181.

Leal-Bertioli SCM et al. Busca de resistência ao nematóide das galhas, *Meloidogyne* spp., e seqüências análogas a genes de resistência em acessos silvestres de *Arachis*. Boletim de Pesquisa e Desenvolvimento. Brasília: Embrapa Recursos Genéticos e Biotecnologia. 2000;(20)17p.

Leal-Bertioli SCM, Guedes I; Guimarães PM, Fávero A, Bertioli DJ. Isolamento de seqüências de *Arachis stenosperma* em resposta à infecção de *Puccinia arachidis* e *Cercosporidium personatum* por cDNA e AFLP. Boletim de Pesquisa e Desenvolvimento. Brasília: Embrapa Recursos Genéticos e Biotecnologia 2005;

Leal-Bertioli SCM, Moretzsohn MC, Bertioli DJ, Guimarães PM. Perspectivas da biotecnologia na busca de resistências em *Arachis* spp. In: Santos, R. C. Ed. O

agronegócio do amendoim no Brasil. Campina Grande, Embrapa Algodão. 2005; (10):423-451.

Leal-Bertioli SCM, José ACVC, Alves-Freitas DM, Moretzsohn MC, Guimarães PM, Nielsen S, Vidigal BS, Pereira RW, Pike J, Fávero A, Parniske M, Varshney R, David Bertioli DJ. Identification of candidate genome regions controlling disease resistance in *Arachis*. 2009.

Lee M. DNA markers and plant breeding programs. Adv. Agron. 1995;(55):263-344.

Lein JC, Asbach K, Tian Y, Schulte D, Li C, Koch G, Jung C, Cai D. Resistance gene analogues are clustered on chromosome 3 of sugar beet and cosegregate with QTL for rhizomania resistance. Genome. 2007 Jan;50(1):61-71.

Leister D, Ballvora A, Salamini F, Gebhardt C. A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. Nat. Genet. 1996 Dez;14(4):421-429.

Litt M, Luty JA. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am. J. Hum. Genet. 1989 Mar;44(3):397-401.

Liu Y, Qin X, Song XH, Jiang H, Shen Y, Durbin K, et al. Bos taurus genome assembly. BMC Genomics. 2009 Abr 24;10(1):180.

Livingstone DM, Hampton JL, Phipps PM, Grabau EA. Enhancing resistance to *Sclerotinia minor* in peanut by expressing a barley oxalate oxidase Gene. Plant Physiol. 2005;(137):1354-1362.

Lopez Y, Nadaf HL, Smith OD, Conneli JP, Reddy AS, Fritz AK. Isolation and characterization of the Delta(12)-fatty acid desaturase in peanut (*Arachis hypogaea* L.) and search for polymorphisms for the high oleate trait in Spanish market-type lines. Theoretical and Applied Genetics. 2000;(101)1131-1138.

Luscombe NM, Greenbaum D, Gerstein M. What is bioinformatics? A proposed definition and overview of the field. Methods Inf Med. 2001;40(4):346-358.

Martins W, de Sousa D, Proite K, Guimarães P, Moretzsohn M, Bertioli D. New softwares for automated microsatellite marker development. Nucleic Acids Res. 2006;34(4):31.

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. 1999 Nov;20(3):317-332.

Milla SR, Isleib TG, Stalker HT. Taxonomic relationships among *Arachis* sect. *Arachis* species as revealed by AFLP markers. Genome. 2005 Fev;48(1):1-11.

- Miyagi M, Humphry M, Ma ZY, Lambrides CJ, Bateson M, Liu CJ. Construction of bacterial artificial chromosome libraries and their application in developing PCR-based markers closely linked to a major locus conditioning bruchid resistance in mungbean (*Vigna radiata* L. Wilczek). *Theor. Appl. Genet.* 2004 Dez;110(1):151-156.
- Moraes SA, Godoy IJ, Martins SLM, Pereira JCVNA, Pedro Jr MJ. Epidemiologia da mancha preta (*Cercosporidium personatum*) em amendoim: resistência controle químico e progresso da doença. *Fitopatologia Brasileira.* 1994;19(4):532-540 (1994).
- Moretzsohn MDC, 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.* 2004 Jul 14;4:11.
- Moretzsohn MC, Leoi L, Proite K, Guimarães PM, Leal-Bertioli SCM, Gimenes MA, et al. A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (Fabaceae). *Theor Appl Genet.* 2005 Out;111(6):1060-71.
- Moretzsohn MC. Desenvolvimento e mapeamento de marcadores microssatélites e identificação de QTLs ligados à produtividade e à resistência à mancha preta em *Arachis* spp. Tese de doutorado, Universidade de Brasília, Brasília, 2006;134 p.
- Moretzsohn M, Barbosa A, Alves-Freitas D, Teixeira C, Leal-Bertioli S, Guimarães P, et al. A linkage map for the B-genome of *Arachis* (Fabaceae) and its synteny to the A-genome. *BMC Plant Biol.* 2009 Abr 7;9(1):40.
- Morgante, M, Hanafey M. et al. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nat Genet* 2002;30(2):194-200.
- Nam YW, Penmetsa RV, Endre G, Uribe P, Kim D, Cook DR. Construction of a bacterial artificial chromosome library of *Medicago truncatula* and identification of clones containing ethylene-response genes. *Theor. Appl. Genet.* 1999;(98):3-4.
- Nelson SC, Simpson CE, Starr JL. Resistance to *Meloidogyne arenaria* in *Arachis* spp. *Germplasm. J. Nematol.* 1989 Out;21(4S):654-660.
- Nelson SC, Starr JL, Simpson CE. Expressão de resistência to *Meloidogyne arenaria* in *Arachis batizocoi* and *A. cardenasii*. *Journal of Nematology.* 1990;(22):242-244.
- Nutter FW, Shoke FM. Management of Foliar Diseases Caused by Fungi. In: *Peanut Health Management*, eds. H.A. Melouk and F.M. Shokes. APS Press, St. Paul. 1995;65-73.
- Pertsemlidis A, Fondon JW. Having a BLAST with bioinformatics (and avoiding BLASTphemy) *Rev. Genome Biol.* 2001;2(10).
- Pezzopane JRM, Pedro Jr MJ, Moraes SA, Godoy IJ, Pereira JCVNA, Silveira LCP. Chuva e previsão de épocas de pulverização para controle das manchas foliares do amendoim. *Bragantia.* 1998;(57):285-295.

Proite K, Dias JGO, Guimarães PM, Bertioli DJ, Carneiro RG, Valls JFM, Leal-Bertioli SCM. Avaliação de resistência a fitonematóides e a fungos em diversas espécies de *Arachis* silvestres. In: VII Encontro Talento Estudantil. Embrapa Recursos Genéticos e Biotecnologia, Brasília, Anais do VII Encontro Talento Estudantil. 2002.

Proite K. Busca de Genes Envolvidos na Resistência de Amendoim Silvestre ao Nematóide das Galhas (*Meloydogyne arenaria*). Tese de doutorado, Universidade de Brasília, Brasília. 2007;220p.

Proite K, Leal-Bertioli SCM, Bertioli DJ, Moretzsohn MC, da Silva FR, Martins NF, et al. ESTs from a wild *Arachis* species for gene discovery and marker development. *BMC Plant Biol.* 2007;77.

Robledo G, Seijo G. Characterization of the *Arachis* (Leguminosae) D genome using fluorescence in situ hybridization (FISH) chromosome markers and total genome DNA hybridization. *Genetics and Molecular Biology.* 2008, 31(3):717-724.

Lourenço RT. Estrutura genômica de três mega bases de DNA genômico “sequenciamento por fragmentação randômica de DNA (shotgun)” de *Eucalyptus*. Conteúdo nucleotídico, seqüência repetitivas e genes”. Dissertação de Mestrado. Universidade Estadual de Campinas. 2004:138p.

Rommens CM, Kishore GM. Exploiting the full potential of disease-resistance genes for agricultural use. *Curr Opin Biotechnol.* 2000;(11):120-125.

Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 2000;132365-386.

Safár J, Noa-Carrazana JC, Vrána J, Bartos J, Alkhimova O, Sabau X, et al. Creation of a BAC resource to study the structure and evolution of the banana (*Musa balbisiana*) genome. *Genome.* 2004 Dez;47(6):1182-1191.

Saini N, Shultz J, Lightfoot DA. Re-annotation of the physical map of *Glycine max* for polyploid-like regions by BAC end sequence driven whole genome shotgun read assembly. *BMC Genomics.* 2008;9323.

Sakata K, Nagamura Y, Numa H, Antonio BA, Nagasaki H, Idonuma A, et al. RiceGAAS: an automated annotation system and database for rice genome sequence. *Nucleic Acids Res.* 2002 Jan 1;30(1):98-102.

Sambrook J, Russel DW. *Molecular cloning: a laboratory manual.* 3.ed. New York: Cold Spring Harbor Laboratory. 2001;(3)

Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain terminating inhibitors. *Proceedings National Academy of the Science.* 1977;(74):5463-5467.

Santos RC. Peanut crop: a viable alternative to Brazilian Northeast growers. *Ciência e Cultura.* 1995;(47):9-10.

- Santos RC, Godoy JI, Fávero AP. Melhoramento do amendoim. In: Santos, R. C. Ed. O agronegócio do amendoim no Brasil. Campina Grande: Embrapa Algodão. 2005;(4):123-192.
- Sato S, Isobe S, Asamizu E, Ohmido N, Kataoka R, Nakamura Y, et al. Comprehensive structural analysis of the genome of red clover (*Trifolium pratense* L.). DNA Res. 2005;12(5):301-364.
- Schlötterer C, Tautz D. Slippage synthesis of simple sequence DNA. Nucleic Acids Res. 1992 Jan 25;20(2):211-215.
- Schmidt T, Heslop-Harrison. Genomes, genes and junk: the large-scale organization of plant chromosome. 1998;(3)195-1999.
- Seijo JG, Lavia GI, Fernandez A, Krapovickas A, Ducasse D, Moscone EA: Physical mapping of the 5S and 18S-25S rRNA genes by FISH as evidence that *Arachis duranensis* and *A. ipaënsis* are the wild diploid progenitors of *A. hypogaea* (Leguminosae). American Journal of Botany. 2004;91(9):1294-1303.
- Seijo G, Lavia GI, Fernandez A, Krapovickas A, Ducasse DA, Bertoli DJ, Moscone EA: Genomic relationships between the cultivated peanut (*Arachis hypogaea*, Leguminosae) and its close relatives revealed by double GISH. American Journal of Botany. 2007;94(12):1963-1971.
- Senthilvel S, Jayashree B, Mahalakshmi V, Kumar PS, Nakka S, Nepolean T, et al. Development and mapping of simple sequence repeat markers for pearl millet from data mining of expressed sequence tags. BMC Plant Biol. 2008;8:119.
- Shen KA, Meyers BC, Islam-Faridi MN, Chin DB, Stelly DM, Michelmore RW. Resistance gene candidates identified by PCR with degenerate oligonucleotide primers map to clusters of resistance genes in lettuce. Mol. Plant Microbe Interact. 1998 Ago;11(8):815-823.
- Shultz JL, Kazi S, Bashir R, Afzal JA, Lightfoot DA. The development of BAC-end sequence-based microsatellite markers and placement in the physical and genetic maps of soybean. Theor. Appl. Genet. 2007 Abr;114(6):1081-1090.
- Shuman, S. (1994). Novel Approach to Molecular Cloning and Polynucleotide Synthesis Using Vaccinia DNA Topoisomerase. J. Biol. Chem. 269, 32678-32684.
- Simpson CE, Starr LJ. Registration of "Coan" peanut. Crop Science. 2001;(41):918.
- Simpson CE, Krapovickas A, Valls JFM. History of *Arachis* including evidence of *A. hypogaea* L. progenitors. Plant Science. 2001;(28):78-80.
- Sledge MK, Ray IM, Jiang G. An expressed sequence tag SSR map of tetraploid alfalfa (*Medicago sativa* L.). Theor. Appl. Genet. 2005 Set ;111(5):980-992.

- Soares JJ, Almeida RP, Santos RC, Santos JW, Silva CAD. Avaliação do nível de resistência de genótipos de amendoim à mancha angular causada por *Cercosporidium personatum*. Campina Grande: Embrapa Algodão. 1996. 46p
- Southern EM. Detection of sequences among DNA fragments separated by gel electrophoresis. *Journal of Mol. Biol.* 1975;(98):503-507.
- Staden R. The Staden sequence analysis package. *Molecular Biotechnology.* 1996;5(3):233-241.
- Stalker HT. A New Species in Section *Arachis* of Peanuts with a D-Genome. *American Journal of Botany.* 1991;78(5):630-637.
- Stalker H.T. Utilizing *Arachis* germplasm resources. In: NIGAM, S.N. Groundnut - A Global Perspective. Patancheru, ICRISAT. 1992:281-295.
- Stalker HT, Simpson CE. Germplasm resources in *Arachis*. In: Pattee, HE. *Advances in Peanut Science.* Edited by Pattee HE, Stalker HT. Stillwater, OK. American Peanut Research and Education Society. 1995:14-53.
- Stalker H.T. Peanut (*Arachis hypogaea* L.) *Field Crops Research.* 1997; 53:205-217.
- Stalker Ht, Mazingo LG. Molecular markers of *Arachis* and marker-assisted selection. *Peanut. Sci.* 2001;(28):117-122.
- Subrahmanyam P, McDonald D, Gibbons RW, Subba RPV. Components of resistance to *Puccinia arachidis* in peanuts. *Phytol.* 1983;(73):253-256.
- Subrahmanyam P, Moss JP, McDonald D, Subba RPV, Rao VR. Resistance to leaf spot caused by *Cercosporidium personatum* in wild *Arachis* species. *Plant Disease.* 1985;(69):951-954.
- Subramanian V, Gurtu S, Rao RCN, Nigam SN: Identification of DNA polymorphism in cultivated groundnut using random amplified polymorphic DNA (RAPD) assay. *Genome.* 2000;43(4):656-660.
- Takken FLW, Joosten MHQJ. Plant resistance genes: their structure, function and evolution. *European Journal of Plant Pathology.* 2000;(106):699-713.
- Tang R, Gao G, He L, Han Z, Shan S, Zhong R, Zhou C, Jiang J, Li Y, Zhuang W. Genetic diversity in cultivated groundnut based on SSR markers. *J. Genet. Genomics.* 2007;(34):449-459.
- Tanksley SD. Mapping polygenes. *Annu Rev Genet.* 1993;(27):205-33.
- Tao Q, Zhao H, Qiu L, Hong H. Construction of a full bacterial artificial chromosome library (BAC) library of *Oryza Sativa* Genome. *Cell Research.* 1994;(4):127-133.

- Timmerman-Vaghan GM, Frew TJ, Weeden NF. Characterization and linkage mapping of R-gene analogous DNA sequence in pea (*Pisum sativum* L.). *Theor. Appl. Genet.* 2000;(101):241-247.
- Tiwari S, Devesh K, Mishra, Singh A, Singh P. K, Tuli R. Expression of a synthetic cry1EC gene for resistance against *Spodoptera litura* in transgenic peanut (*Arachis hypogaea* L.) *Plant Cell Rep.* 2008;(27):1017–1025.
- USDA – United States Department of Agriculture Foreign Agricultural Service – Approved by the World Agricultural Outlook Board/USDA, Circular Series, novembro, 2008. Disponível em: <<http://www.usda.gov>> Acesso em: 25 nov. 2008.
- Valls JFM, Simpson CE. Taxonomy, natural distribution and attributes of *Arachis*. In: Kerridge, PC, Hardy B. (Ed.). *Biology and agronomy of forage Arachis*. Centro Internacional de Agricultura Tropical. 1994:1-18.
- Valls JFM. Diversidade genética no gênero *Arachis* e a origem do amendoim. In: Bandel G, Aguiar-Perecin MLR, Oliveira, G.C.X. (Ed.) *Encontro sobre temas de genética e melhoramento, genômica: uma abordagem em plantas*”. 2000;(17). Anais. Piracicaba: ESALQ. 2000:51.
- Valls JFM, Simpson CE: New species of *Arachis* L. (Leguminosae) from Brazil, Paraguay and Bolivia. *Bonplandia*. 2005;(14):35-64.
- Valls, JFM. Recursos Genéticos do Gênero *Arachis*. In: Santos, R. C. (Ed). *O agronegócio do amendoim no Brasil*. Campina Grande: Embrapa Algodão. 2005:45-70.
- Van der Biezen EA, Freddie CT, Kahn K, Parker JE, Jones JDG. *Arabidopsis* RPP4 is a member of the RPP5 multigene family of TIR-NB-LRR genes and confers downy mildew resistance through multiple signalling components. *Plant J.* 2002 Feb;29(4):439-51.
- Van der Linden CG, Wouters DCAE, Mihalka V, Kochieva EZ, Smulders MJM, Vosman B. Efficient targeting of plant disease resistance loci using NBS profiling. *Theor. Appl. Genet.* 2004 Jul;109(2):384-393.
- Van Wijk HJ, Harlizius B, Liefers SC, Buschbell H, Dibbits B, Groenen MAM. In silico identification and mapping of microsatellite markers on sus scrofa chromosome 4. *Anim. Biotechnol.* 2007;18(4):251-261.
- Vanhouten W, MacKenzie S. Construction and characterization of a common bean bacterial artificial chromosome library. *Plant Mol. Biol.* 1999 Ago;40(6):977-983.
- Varshney RK, Bertoli DJ, Moretzsohn MC, Vadez V, Krishnamurthy L, Aruna R, et al. The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.). *Theor Appl Genet.* 2009 Feb;118(4):729-39.

- Whitham S, Dinesh-Kumar SP, Choi D, Hehl R, Corr C, Baker B. The product of the tobacco mosaic virus resistance gene N: similarity to toll and the interleukin-1 receptor. *Cell*. 1994 Set 23;78(6):1101-1115.
- Whitham S, McCormick S, Baker B. The N gene of tobacco confers resistance to tobacco mosaic virus in transgenic tomato. *Proceedings of the National Academy of Sciences of the United States of America*. 1996;93(16):8776-8781.
- Wong RM, Wong KK et al. Sample sequencing of a *Salmonella typhimurium* LT2 lambda library: comparison to the *Escherichia coli* K12 genome. *FEMS Microbiol. Lett*. 1999;173(2):411-423.
- Yi G, Lee JM, Lee S, Choi D, Kim B. Exploitation of pepper EST-SSRs and an SSR-based linkage map. *Theor. Appl. Genet*. 2006 Dez ;114(1):113-130.
- Young ND, Weeden NF, Kochert G. Genome mapping in legumes (Family Fabaceae). In: Paterson AH (Ed.). *Genome mapping in plants*. R.G. Landes, Austin, TX. 1996:212-227.
- Young ND. The genetic architecture of resistance. *Curr. Opin. Plant Biol*. 2000 Ago;3(4):285-290.
- 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 U S A*. 1996 Out 15;93(21):11751-6.
- Yüksel B, Paterson AH. Construction and characterization of a peanut HindIII BAC library. *Theor Appl Genet*. 2005 Ago;111(4):630-9.
- Yüksel B, Estill JC, Schulze SR, Paterson AH. Organization and evolution of resistance gene analogs in peanut. *Mol Genet Genomics*. 2005 Out;274(3):248-63.
- Zhang HB, Martin GB, Tanksley SD, Wing RA. Map-based cloning in crop plants: tomato as a model system II. Isolation and characterization of a set of overlapping yeast artificial chromosomes encompassing the jointless locus. *Mol Gen Genet*. 1994 Set 28;244(6):613-21.
- Zhou T, Wang Y, Chen JQ, Araki H et al. Genome-wide identification of NBS genes in japonica rice reveals Significant expansion of divergent non-TIR NBS-LRR genes. *Mol Genet Genomics*. 2004;271:402-415.

ANEXO 1

Soluções utilizadas para Extração de DNA

Tampão CTAB 2X

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

Tampão CTAB 1X

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

Soluções utilizadas para confecção de gel de poliacrilamida

Gel 4% Desnaturante (1000 mL)

- 453 g de Uréia
- 534 mL de água destilada
- 50 mL de TBE 10X
- 125 mL de Acrilamida: bis (19:1) 40 % (solução comercial)

Preparo: Misturar em um béquer de vidro os reagentes na ordem em que aparecem e agitar com agitador magnético a temperatura de 55° C, até a total dissolução da uréia. Filtrar a solução em filtro Milli-Pore (20) e acrescentar 670 µL de TEMED. Estocar a 4° C.

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
- 10 µL de Ácido acético
- 6 µL de Bind Silane

Preparo: Lavar a placa com água e sabão e deixar secar. Espalhar a solução preparada e deixar secar por 5 minutos. Lavar uma vez com álcool.

Placa de Repel (Placa maior)

- 1 mL de Repel Silane (ou silicone)

Preparo: Lavar a placa com água e sabão e deixar secar. Espalhar a solução e deixar secar por 5 minutos. Lavar uma vez com água destilada.

Montagem do gel desnaturante

Colocar espaçadores apropriados na placa maior e colocar a placa menor por cima nivelando-as. Misturar 100 mL de acrilamida 4% com 666 µL de APS 10% (1 g de Persulfato de Amônio em 10 mL de água destilada). Verter esta solução preparada com auxílio de uma seringa e esperar 2 horas para a completa polimerização do gel.

Pré-corrída e corrída efetiva

Após a polimerização, colocar as placas na cuba de eletroforese e correr por 40 minutos a 90 W. Aplicar o gel e correr de 40 a 90 minutos, dependendo do tamanho dos fragmentos.

Coloração com Nitrato de Prata (Creste *et al.*, 2001)

Solução de Fixação (2 L)

- 200 ml de Etanol Absoluto
- 20 ml de Ácido Acético Glacial
- 1780 ml de água destilada

Preparo: Após a corrida, num recipiente adequado, colocar o gel e agitar lentamente por 10 minutos. Lavar o gel por um minuto em água destilada, agitando lentamente.

Solução de Oxidação

- 30 ml de Ácido Nítrico
- 1970 ml de água destilada

Preparo: Agitar o gel lentamente por 3 minutos. Lavar o gel por 1 minuto em água destilada agitando lentamente.

Solução de Nitrato de Prata

- 2g de Nitrato de Prata
- 2000 ml de água destilada

Preparo: Agitar o gel lentamente por 30 minutos. Lavar duas vezes, cada uma por 30 segundos, em água destilada, agitando lentamente.

Revelador

- 60 g de Carbonato de Sódio
- 1,08 ml de Formaldeído a 37% (adicionar na hora de usar)
- 2000 ml de água destilada

Preparo: Agitar o gel lentamente na solução (10 a 12° C) até que as bandas apareçam.

Solução Stop

- 100 ml de Ácido Acético Glacial
- 1900 ml de água destilada

Preparo: Agitar o gel lentamente por 5 minutos. Lavar o gel em água destilada, por um minuto agitando lentamente. Deixar o gel secar completamente.

Soluções utilizadas para *Southern Blotting*

Solução de Depuração

- HCl 250 mM

Solução de Desnaturação

- NaCl 1,5 M
- NaOH 0,5 M

Solução de Neutralização

- NaCl 1,5 M
- Tris-HCl 0,5 M

Solução Denhardtts 50X (100 mL)

- 1 g de BSA
- 1 g de Ficoll 400
- 1 g de PVP 40

Preparo: Dissolver em H₂O Milli-Q autoclavada.

Solução de Esperma de salmão 4 mg/mL (10 mL)

- 40 mg de espermidina
- Completar com água Milli-Q autoclavada

Solução de Pré-hibridização (25 mL)

- 6,25 mL de SSC 20X
- 2,5 mL de Denhardtts 50X
- 0,625 mL de SDS 20%
- 120 µL de esperma de salmão 4 mg/mL
- 15,626 mL de água Milli-Q autoclavada

Solução de Lavagem I (500 mL)

- SSC 1X
- SDS 0,1%

Utilizar a solução a 65° C

Solução de Lavagem II (500 mL)

- SSC 0,1X
- SDS 0,1%

ANEXO 2

Sequências dos oito maiores *contigs* gerados a partir da subclonagem do BAC AD25F09

Em cinza (claro) estão localizados os *primers Forward* e *Reverse*. Em cinza (escuro) encontram-se os SSRs e poli(n).

>AD25F09-PO-F10.FM

tgagaactagagtaaatggctccccatcatcaatcacttggacattagtactcaagaggtcctaagttacctttccaagccaagagtagtaaaa
tctactctaaaaccaaccaagcatctcatcaaacacttgggaaggcacaaaaagaaaagcatagtaaattgacaattgacaacaagaattaat
tctaaacaacatcaaatgtaaagaattagcaacaacaataaagggaacaagatctacatgaattacctcttatgtaatgaaagaaaatggaa
gaaacaatagtagacttacaacaaaacataagaacaacataaaaagagattacaccaaagaatagaagaagagtagaagtgaactataaggaatt
gagaagatagaagtagaagaagatgaattaaaacctagatctaaagaactaatcctaattcctaatactctagagagaagtgagagcttctc
tctctagaacctaatctaactactaaactaaactaatggttaactaacatctaaagtataatgattccccctcagtccttgactt
aaatagcatcagaattgacttggatttggccccaaggtcttcaaaaattcgctggccacgttttgctttaagtgaactaggtggcagcaacggtc
ggttggcgtactttgctggtggtggtccactatagctgtagcaactatggcaaaccttataatcgttgggaagccccggatgtactttccaac
cgaactggaacccatcatttggacctctgtagctcaagttatggctgtttaagtggaagaggtcggttgacagcttccggctccttcaatt
tcttcatgagtttctccaactttcatgcttcttctctcattcccttgatccaacttttggctcctaacaaccttaactcacttaacaacatataca
aggtatctaattgaaatcaaggggaatttagatttagctttttaagtcctaaaagatgttttcaactcttaagcacaataagggagaatataca
aaagcctaaactaactaccatcttcaactttttcacataactcatgcaaatctccttcaattcacaacacttatgacttctttatggactt
cctccccacactttaaccaagcatgttctcatgcttaagccaagagaaaagcaaggacatcaacatttattcaatgtaataaaactatatagcaa
cctaactatatagcaactaaatgcaaatgggttttactacttggtaaaaataaatcaatctccaagacatataatgacacaagtagggccaaga
tcatataacgattcatgagctcctaccaatgtaagtataaacatgaagttcacatagacttgcaagaagaacgctcgtaaacggcggaatcaag
gaattgagcatcgaacccctcaccgggagtgttgcactctagctcgtctgggtttgggggttaattctcctcaattctccccatcatgcttcc
aaaatttgtttctcattacaacatcaacaatatttctgcatgcatacataaattatcatgaggtcttttctttaggttgaatggggctaggggt
aaggtaggatgcatatttggctaagtgagctgaaatttgaattcttgataagcttaaacctccaccctaaccttatgacatctatacaattct
aaagcctaaactaactaccatcttcaactttttcacataactcatgcaaatctccttcaattcacaacacttatgacttctttatggactt
cactttgttttggggctatttgtccctttttatctttctttcttatataattttttctttccttttccatattattttctttctttt
tcttttggttttctctttttttcttcttatatacaaaaacctcaatgcatgaaggtttacatttgatcaatacatatgattacccaatt
ccaataatttcaatgaaaatcaaaaacctccttttctaccccaatgctccaaggttcccacacttaaatgatactcacacactagccta
agcctaaacaagatccaaataaggaattatttggtttttgccttcaaggtctgttaagtgtgtcaaaaataggaacaaggtgggttaactgaggt
caaatgggctaaacaaggaatatacaaggaagggcatttgggttaagtgagctaatgaaatgatggcctcaatcatabaaatgcatgaaacac
aaaataatggacataaagaatcaaacaaaatgaagattacattcatgaaagagaataatgcacacaagaagggaaaaataaggtggttatagga
gtaaccacaccattaggtcctaattcacttgccttggttctttagctcaaaaacctggtccaaaatacatcttccaagcaagttcaacaaa
aatttcaaatggtaggttgcctaaaacggtttcttggggaagaaatcatcacccctaaccaagtagctcctaataagaagaaagtggtaaaaa
atatgtacaatcttaactaacatgcaacctatcatgcatgcaataaactaacataacatttgggttgaaaaaagaaaaattgtaccatgggagat
cggtcgagcactcccacacttgaagattgcaacctctcagtgcactcgaagaaggaagcaaggtgagcgggttgcatacttatgtagctct
tcaaaaggttgcgtgagcttgggttggccttttgaaacctttctcttcccgatttgggtggcacaaccacaaaagaaaagaaagaaaag
aaaataaagcctatacaaaagatatacaaggaatagaacatagggggggcctaattgccaataaagagttagatctcctactacatggtgact
caacatgtaggaagaagaacaacataagcctaattggcattcaatgatcttgatgcaagagtaaaagttaaagcattgaaaaacatatgagcat
aagatcaaaaagaattgacataaaacaagattagtgataaagcaaacatctgggcccaacctaacctcaatgatgtaggaagcacaataaa
aggacaatgaggttaggaaggactaaagcactaccttgatagtgaacaaaaatttcaattaatgttaaacaggtcacaaagcaccagaataaa
gcaagagagctcctaataatgtagtaagaattcaacacattattgaaatgtcttagaaaaaaaacgaaaaacatgctacaaaaacttagat
gaaaaatgggaagtagaagtatgcaaatgtgataagcaaaagaaaaatgggaaggggagaaaaaaactcttttttttaccgagcgcgtgcgcg
tcatgacgcttaacgcgtagaggttggatattggttgaaggaagcgtacgcgcaacaggtgagccacgcgcgctcagagtaggcccagggcaata
tgcggcccaagctctggcacaactcctcggttaaaggtaccggggtgcagattgtgcaatcgacgagcgcgcgcgacaggtgctgcgctgcgctgcat
tggaacttaagatgattgctgctacgcgcaaggtgcgcgcaacgctgcactagacttgggtctgaggcccaattgctgcacagctccaggccta
actctcgggttttggctgggttgaatttttgctccacgcgctacgctagcagtcgctgctgctggttgggtcgaaaaatgctcaggtgctg
cgtacgcgcttaggtgagcgcaacgcgtggtgggttctctgttttcagaaaaattttctctatgcttttgcaccaattcaggcatccaaaacct
caagctactacaaaaacacctaaacctatttaacatgttaaaactaattaaactctaccaagatttcaaacatgaaatataaactaggt
ctaccaatgtcaaaaagataaaaatgaaaatgagaactcactcaaatcaatcacttatcaaaaactcaaaaagataggaag
agttaccattgggtgggtgctcctccactagcacttttattttcttctcaagtggacttatggggagctcctcatcaaggtggcttggct
tggattcatttggaaacttccaccaatgcttggctctcctattgtgcacaaagattcttttggattgagccaaggttggatggaggtcttcaca
agcttggggctccaaaagttagctctcttctgtgactcgggactcccacacttatcttcacacccgtctgaggttagatcatcatatagctc
catcgggttggtaacaagatgaattctctatgaaagtgcacaacaaatcctcctagaacctctatttggtaactattccaaaccttggattca
ctcttgaagtatcaaccaaaatgagccttgatttgaacgcaccaaccgaaacatcttctgcttacgctcactcccacaaagcatcctaaggtg
acctccgtttccaagcagccatctcaagtggaacaataaagcttaatgaaaatgaaacttaccactcaagtgagggagtagaggcaacctga
ggtaagaagcttctcaaaaaccacaagatcaaaaattatctgattctctagcctcagaaaaaaaagggaaaaaactctgagcacttgcagtgc
aaaaccagaaaaacagcagcagcagtagatgattcaaaaatcaataaaaaatccaaagtaaaaatccaatgactttaaaaataatgtagttaaact
ttactcattccaggtttcttctcgattgggtctcgagtttaaccattttttaatgaagaagttacttacttggaaagttaagtaaaaatgagttca
aatctgttttgaacccaacaaagcttagtacctttcaattgaaataaaatttatcaaaaactccaaatgagctaaatttggatttgggaactc
ctagctctccagaaaaaacgtagcttcttagttcacaacctaattctcaatttcaagagttaacaagattgtaaggttagatgcatagctgct
gaaatcgtttcttctcagatttgaccactaatataaaaaaatctacaactcccaactcctcaactcttaaaattctgaagttttagagaata

aagcaagttaatacaaaat tttataaaaaaaat tagtttgcgtccaaaact caactcgtagaagtcgcagcaaacacaaacaagt tgcctgcctgtt
gacct ttttttgctgcaggacagatttaacaacctaaact ttaaaaaat tttgccataaaat tggatatttaaaaaaaggtcccaaat tttttagtt
agttctttatattcccaagtttagcccaaactgggtctcacaatccgatcattacataat tagttacggcatatgcaatacaccacaca
caactctacatccttatacaaacaccatcctaatccatcataaataaataagagcacaccat taaactccacacccataaat tcca
atataaaatcccaacaatctattccaacaacat tacaaggtcaacttaatacaacaacaacaatccaactctccttggtctcctaac
ctaagttttcaacaacccgtaaatataaaactgcaaaaact taaatcatacctgggtgatcacttaattcaccacagtcagcctctcaacaca
aaatcacagccctccaagctcaatcaaacagccccaagcaagcttgcaccaacaagctcctaaagtaatccaatccaagctccaatgcat
aacacccctctaaactacacctaatacacacataatagtccaattcagttttctattaccaaaaaaagatggagctagggttaggggtg
ctttccatacccatatgctcaatagcttgagcccacaatccgggaagctaaactgaacctagaacatagaaatgcaacaagatcaccatagg
tttctaagttacaaaagaaagaggatagggattctgaagttataggaggctaccagtgaaatggttagatagaaaggtggagctcgacg
cgctgagcgcgtggcgcgaacgggtgccccgagcggagccgaacagaggagttaggtggatgaaaggttaggggttagggagcttctccc
ttcctccaattctggtttggcggtgcagcaactaaagaggaagaaggctgcccgttcccttataataggggtccgggtggacctcggg
tcggtttgggcccggattcaaccggttcggctcctccgggtccgattttgagccaaat ttcgaatttgatattaaaaatctcattttgatgagc
tctatcctatttttaatatagttttgcatttttagcttttctaatttaataacaatttataactaattaaactaccgattttagcggggtttac
atattcacctactgaatcataaactactaacgaaaatataactagaactgaaccacacatcaaccacttgatggattcaaaatattaaact
actttcacttggttcaattgcaactctgaacttgaacttaattatctccaacaacagagataactggttaaaactttcaaaagtttgatt
taaaaaatagaaaatgaaggaatcgcgaaaaacaaaaaaagataacgttagagaacgaacaaaaaata tagagatagaaaagtacatagat
cgaaaaatgtaaaaaat taaagaaatgaacaaaaactcctcaaaaaatgattatataattacgcataat tgaaaaagtgtt taaatagat
aacgatagggtaaaat taaat taaaaat aatataatata **ataataataataataata**aat tgtgtgagaggatgcttagaaagaat
agatgtccatattagcaat tcaatttctatggtgagaggaataagctcttatggtgaaggaataatagaggttgaaaaagcaagtttg
ttgctataaaaaactgtagagggcaaatggaatatacacagctcaataaaaact**ctcctcctcctcct**tttgctgcaatattctct
tcaactctcttattttatattgtttacatctccttattttatttttagttattatataacatggtatcagcaccgaagctcaacaaaat
ttaaagaagacttcaggttaacaaaat ttttattgtcaaaaactccttctcattgaataatgcttttgatataatcggaaacaaat ttttca
tggatctagatggttaaaatccatcttgattcaatggatcttggagatccat taaggctgaaaaat taatgcacccagaaatgataagccaaag
ccatgattttctctatctcctatctgacgtatgattgaaaaatgaatatccacat taaagatcctgcagatctgtggaagaccttgaaaa
aaggtacaactcatgtgatactcccaagcccaatggttagagaaaaactttctctacctccatgcctcgaatgctcctgagcttagcaatc
gagaaaaaggatttaaaaaatattatgagctaaat tcttgcctttttgctgctgaacgcaacaataagttgctctaaagaaatcatgaaacgcg
cccagctggcgcgtgccccat tttctgaagcaaatccggcaaat taaacccagaaagaggtaaatggcaagat tttgataatagaaaaat t
tgaaggaagaaat tttgtcacaagaaagatactcaccagaaggtgggataaagaaagaaacaatgggcaagataatcaat tggagtaaat
at ttcctgtggtggaagggccattggctcacgtacctgtgactcctcacaagcaccctagttgatctttatcaagcatccttgaaaaaggtg
caagggaaaggaataat ttttatttcaaatgatgaaaatccaccactcattaatgatctcaattctccaaggaatctgaaaggaatatt
gggtattgatcaatgatggaatagtttgataatgtagtggtttgtaagttatcagttgaaataat tttactgtgcactgactcctactcatt
tattattattcatttggtttttgaagaaaaatggcaaggaataatctggaagatatttgccctgccggatagtgcaagtttgcaactatct
taaaagtgatataattaccatctggtgcaaaagagagatgttaatactattat tggctcagacaaatgtagagaaggtccggaagag
ctataattttggttctctggaggaacaacat tttgataaaaaatgcaactatctagcaagctctcaaggaactgttgattctcaagata
tgcgcgaatggatcattcattatgaaacaatggaatgaggaataatgcaataattttatgtatcataactcattgataat taaataaaaaagt
aaaaagttaccctcactttcactcggattgattatatacaagattagtgtaattgaaatcacatgaccttgtaaacagaaagttactagccaa
atgaaatcacaactggcatgatagattaggttatccggaacaaacatgatgaggaggattatgaaaaactcctcaggaactcactcaagaag
cagaagattctaaactagtgaattttggtgctgcatgattcagaaggaagtaattttaaaggtcactcaccagtaaagattggatttgat
cccctgaaatcctcagaaggtccaggtgataatggtggaactatcactggaagatcttgatataatggtttagatattttatggtctgatag
atctcogagatggtcacatgctgcttatctctcogcaactggcgtttgocgagatctgggtcacaattatctgattaaaagcacaatt
ccaaaaatccaatcaagcaatctgcttgataatgctggtgaaat ttaactcccaagattttgatgcttattgattggcctaaatggaat
tgaacatccagtagcttattgtccacacaaaaatgggttagtgaactctctaatggaagcctccaat taaatgctccagttggaacaaat
aacaaatctcccaactcgggtttgagggcagctattttacatgcagcataactattcgtttgagggcaacagagttaccatcaattctcct
atgcaattagcttttggccagcagccaaatatttctcattaaagaatattcgggtgctgcgataatggtccccatgcaaccctcaatcgacca
aaatgggaccccagaaat tggggatataatggtgatatgattctcctcctatagtaagatattctgagatacaaacgggagatgatttaa
agccgggtttggtgatttgcattttgattgaaatcaaaaat tccagatctgaggggagagaataagctctcgaagaaagaaat taaatgca
tcatcgttgatgcat ttagatcctcgatcagggcaatgtaactagaagt tcaaaaatatacatttgcaagaaatgcaaatgaaatgctg
atgcat tttccgatacaagaggataaaccaatcttatatacaaacgaaaaatcccaactcgaattgattgcttagttgagcaagtagccac
tgaagcaaat tcaagctagaaggtgaggcgggaaaaagcccaactcgaat taaatgctccagttggaacaaatgcccacgaaacaaataca
cgccaaaagcgtggcagggctgctcggttccaaagacaaaaatcctcgaagagaaaaagaggtaaatattattctgttgaaaaaagacatagta
gagacacctgtagttgctcaaaatctgatatcaacgcagagaagcgttcaggtcactgaaat tggaaaaatgacagagatctcgaataat t
ctttacagggagagaatgggaccgaaatagacaatgtcaatgaaatattgcatataatgtagat taaatataatgcatgaaataagga
tcttgaagcaagatcagtcgaagaaatgtcaacaaagaaatgattggccaaatagaaagaaagccatgaaagctgaaatgagcactctgcaaaa
cgtgaagctcttggacctgtagctgtacacctaagagatgaaacactggtggaatacctaaatgggtattttgtaagaaacgaaatgagaaaatg
aagttgctgcctcaaaagccgactgtggcacaaggtttttcaaaagggcctggatagattatgaaagaaacgtattcccctgtagtggatgc
gataacattgctgtattttggtcagttacctgcataatcataagctgcatatgcatatgaatggaatggtggaatagcctattgtacggctcatta
gatcgggatattctatagaaagcctggaaggactaaatataatcaaacctcagtgaaatattgcaagggtataactcagttaaattgcaaa
gatcttbatatggtttaaagcatattggaagaaatggtgtaactcgtctactgagatctggccaaaaatggatcaagaatgatgatctg
tccatggttttcaaaagaaatctgactggtatcctataatgctggtgactcgtggaagatttaaatatcatgggactcctgaagagat
ccaaacattataaaaaactcaaaaagaaaggtttgagatgaaagacttgggaagaaactaaaat tttgtcagctgcagatcgagcatcaaaaa
atgggatctttatctcaaacgcacatacacagagaagatctgaaagaaatttataggaatgcaatcattaaagtaacccaatgattgt
aagattcttagatgtgaaaggaatcaactcctcaaaagagaaatgaagataccttggctcctggaagtaccatattcttagtccat tgg
gctaatgtagcttgcctaaatcaacgcagcctaacatatactcgggtgaaat taccagcaagatattgctcttccaaccaaagaaat
ggagtgaaatcaacaaaat tttcgataatggaacgggtgataaggtattttatctcttaggatacgaactagttggctatgcag
ataccggatctgtctgattcacataaaagggagatctcaaacaggaacactggtcacaatgattggtacaactatatacggaggtccacgaa
acagcagatagcagcaactcctcctaactcatgctgaaat tttggcagatcctggaagctagctgcggggtggtttttgctgagaggtctgat tcaa
tatattcggctcactgatgtagactgattgatcataagatagctccaaat tctcgtttgaagataatcacagcatgattgctcacttaaggggtg
gatacatcaaggtgtagaacaagcattttctccaaat tttctcactcatgatctcaaaaatcaagggcaaatgattgtccaaagat
ccattcaactgacaattcggcagatttatcacaaggtcactcccaaaactcctttgaaagattggtacatgagattgggattcggctgatt
cgagacattaaatgattgctcgcaagggggagacttatttgcgggtgggttcgtaaaaaatataatagagcaaaaaatattggttaagatga
gagattcggggtaagtaaat tttgtagaaaaataacttttataatggcagatagcaatacaaaaactaatgtagagaaatagaggtggagaa
caagttgtgagaccccaactgtcacaagaggaataatgattttgtttaaagttggaatgggtggttaaggttgcgaaggttttggga
gaacactggaatgccaagt ttttactcgtgaaagggcagatgaaatataagaaacaaaatgagaggaat tggggagagaaagataggttt
ttagcttat tagcttcaaaagttgaaagatctagaagaggtgaaaggtgaaaggtgaaaggtgaaaggtgaaaggtgaaaggtgaaaggt
taaggttaaggggtgaaaggtgaaaggtgagcgtgcttacaaggtggcttgaggtcctttttagttgtaaaacatgagagaggaatct

tctcacattcttcagcttctgggaacaagatgctcgtgggtgccccttactgcatgaaggagaggaagagcgtgctgccacaacggggggcgt
tatgaccatttggcgtctcgcaaaagctctgacccacctgaactcttacctctctcccttgcattcccaacagcacttggctctaaatagctc
atgacacaggttctcaatataaacatgtatcatgcaaaaattagatctcatgatacatgacagagcaacaatggcgtcatttgattcatgta
gccaccccacttagtgggacaagccttgggtggttcttggcaaaaattagatctcttcagctaaatgggtccccagatttgaagctta
actgcaatgtaaccaggaagaaattaaactctggggcaaaaacagacacatgaaggcgatataatatttggggataaactagatgctct
ttagggggtgctcaaaaatgcagacagggctcctgctcttatgaaaaatccggcatttggttttgatggagaatataatctaccgaagctctcctt
atgctacttttcaattttcatgctcctcagttagctctcaaaaatcaggttaattgattctattatggatctatctcgttggctcagtt
taagttgaaatattcaggccgatgacaaatcagtttatataaaatgggtattctcatggtttgatttttagtttggtaattcttttcttca
acttttggtaaaaataatattggttctgatagaggagagctggtgttatgcaaatctcgggaatgtgctctgggataaattgcatgctt
aaatttggtatgagaaaaaggggatggggaggaatataatgagcatttgagggttggatttgagtccagtggtgaggttttctatttt
tttaatatctgggaaaagcaactttaatctgggttctatggttaaatggtgcaacgaagttacagtgattttggtttatgggtggaggg
tttttaggaagatggcacttgagattggtatgaaactgagatctagcagaataaaggatatttgatgaaactagggattgccaattgacact
ccctagaggattctacacagttggaactaccaactcagggagagggattctctcagtttctaacagaattccacaacagaaaaatgcactatctc
taacttaactataacactagctattataggcagtgagtaataagagctcaacaacctatcagccagctggaccatgtaacaactctcaaaactg
acacttaactctacttaccacagcccaaccctgttagtagtagggcgtctctatatacagattgacaatgtctgaata. **ttttttttt**gttgg
ctgagcagcatagagtgagatttgcaagcataacgagacaatcagattcagttctacagagagtgagtgagaaaaatctacaagcttagcaaatg
tcaacttagttgcaattcttgaatttccagttgtctaatctctggggttatgctcttgatgtaacccatagaactttgatfactaatggtt
gatgatcaactttttaaatacacactcttttaagagttctacttgaagaagcttttaatatgctattttcttctttaaacagctatattt
tttcaatttgagacactcttcttccaactatgcaagctccaggtccagttctcagacaaaagaatacaccattggctcagcttcaactgaagtctc
ttgaattgttatcagaggatttgctcagctccaggtgctgctgctgatttttttctcttaattgaaaactctaatcttatggtgag
ttgaaatgaatttggtgcttaataatgggtgtaggatttttcttataaattatggtatgaaaaatataaaaaaaatccacttttggtaa
attactaaaagatattggtgtagtgcataaatacataaattgaaataggatataaggtatcttataaagtagaggtttggacaataatattt
tcaaaagataagataattgacagaaaaagacattcaactaggaattctcactcagctcaagctgataaattgagcaaatcttataaagt
ttcatggcattggaggaaatcattgcaatttttctcctcttgatgtaactcttcttaataaataatgtagacatttccaggaaaatgagata
aatgtataatgtagatgaggttctatgaaaggttaatgcagataatcaactgaaggtgctgaagccaatttatctcagtgcaataactc
aatagttaaagctgaagcttgggtagatataatctcaatattttcttctgcttagttctcactgagacatttttggttatcagctcaaa
ttacagaggcagttgatacttactttctcctgctcacatggtctgctcaagcttaaagaatggatacttctcctgcttactctgttaaga
agttctgaggtatgggtggattcgttgacatgactcagcagcaacgggtgcaaaaataatggtaaaaaatgtaataatgtagaaaaatagaat
tcttaactattctgataataataatgaggtgacacagcagatacaactgacatgataaataatgtagtaactatgtaggactggtaagcaag
ctagtcaaatggcattactatagattctcaactgtagaattgctcaactcagcagcaatgtagccctcttatttgggtctacattattg
ttagatgatttttataacattggtgaatataacatgcaaatcggagagagacattttaaagcaaacacagcaatagccagtaattattgcatg
tctcactgctcattgcaacttttcttttgtaaagggagtcatttagacatgatttggcgtctttttaaattccattctagcaaaaggttcaact
gtggcaagttgcaattctcctaagagactttacaggtgctttttgacttattccatcatgagttcagcacaacagccactgctcagcctgca
ggatattctaattagattactgtgggaaatataagcagccagaaatgtaactaaataaataatataatagcatabaaataggtttctgctta
ttacgtgtgcaactatgtaacacctacatttggtagaatacaagatttactggaagacattgaaagagaaactatcacaattgcaaatagaca
agaataatgaaactggattctcaactataatagaatagtagaacaagagtaacacatggtggccaccacacactattatctgtggaagtgaagc
agctaaaaataaaaaatgcaatttgggtatttggccatgattatggttaatttttctcactctcttctgcaaaaatagatacaatttttttaaaa
gaatgaaaaatgaaaaaaagagaaaaagaaaaagaaaaacggcaactaaagaatgctctatcttttatgtgtggtgataaataatgctca
aacgtccaaggtctatttctctttttaaatttttctccttattttaaatttctcatttgaagagtagccacatattttggcgttctcaataa
cttttggcaatttataaaaaaggttcaacttcttgaactgctcactctcctctcaatgttaacatattctcttcttcttcttctcaacaat
gtgcatctcatgacaaaaatataatgattacctctgctatctgctccatgcaagactcattcaagatttatggaagagaagactttgcaatg
attgaagatgattgagagtgaaatgtagatttcaacttctgcttggatacagattttctgctgcaaaagtataatgcccaggtgggtttcttca
cctgtagctcttgaaagctccgattggattatactcaaaaataaatggccactgctctcttattatcactgattgagcaagcaaac
ttggttatattctggtaaacctaggttttcatcacaaagtcattttgctgggtgctaatcccatgaaatagccgtagttgtgtggccagacgtgt
gcatataattttgaggagtgcaattttctctcacttttccatccaatttggttaacatttttttctgattttgtaaaaatagaaacagacc
atgttatataatatttggtagtctccttttacccttcttagcatttccacagctcaattcatttttgggtgactgggtttgtaagat
tgccatcaattcttggtaggataaactcttagttaaatttactggtgcaaaaaaaataaataaataaataaataaataaataaataaataa
agaaaaataagaagttgtaaaaatgattcattttaccatattttggtcttcttaacgtgattttcacaataatctctggtggacagaaacaag
tcaataaagcaaggtctttgcaaggttgaaggttccctatttggttctctgttagtgtagggcacttatttaaccaatcagattccac
aaactcttataatcagagtttccacacagcagcaatgactcttctctgttcttagtctacaagcaccagaaacacactgctcaag
ctgcagattcagttattccacagcgaactcactctgcttttggctgagatacacatctgaacttaagcattatggagctgctataggttaggt
ctctctagtcattctaccaccacagcattgagacatattcatggtcttgagtttaggggacaagacaccgcaaggttcaactgattcat
ctctatgctcactgcagaggaaggaatcaggaattttagagatgatagaaccttgagagaggtgaagttattcactgaactctcagag
caattgaagagctctatgtttgcagtcgttctgctcccaactatgctcctctggtgcttagatgagctccaaaagattggtgagtg
caagaacaacctggggtcagacattggtccagtgctctacgggtggtgacccctgtagtgtagggcaccaaaaaggaacctttgaggatgctttc
aggaacaagaagagagattgaggagagcagtgagaaggtcaagagatggagagatgctttaatacaagttgccatttactctggatgggact
ccaaaaatcagtgagttcactgctccatcagctcaattggtttatactgattatgtgcttacaataaataaagtaattgtcagcaagttaag
gtgattaaagattccaactttaagtagttgtcagcaagtgtagactgataactgagaggaccaaatgaccaactcttattatacctttgatata
ataaaactaaaaataaaactacttggagcaaggttaagcaattctcttgcataaattatgctctccacatgctatagtagttgaaagaaacat
gcaagtcattctatctcagccaataattgagaactttattatactctcctctatcctataataatctaaattggtttttcaaaaagaaaa
aatgtaactcttttggtagcgtttctattttagaccataaaaaaggtctctcttataataggcagggcaacacttgggaaagcattgctc
aacatgctacacaaggtcaattccaataatgcaactctgcatagagaaccttttgggtatggctcaaggggtggaggatgtaactacactcat
gtgattgggtgagtgatgtagatttacagggcatatggggcattgggtgggtggtggttaagacaactatgctagagcaactatgaaagccat
gaagatcaattccaacttagtttcttctgtcaatataagagacatgtagagacaatgggactctcaactcaaaaaatcttggtgaac
ataatcatgtgagtagatgacttccagtaatttgatgtaggaatgaggatcataaagaatcttcttatgcaacaagaaggtctctattgctct
tgcagatgtgaaatgagtaagccaattgggagaatttggcaggggaatcaagactgggttggccctggaagcagggtaagattgaaacacagagat
atgcatctacaagaacacatgaaggtggtgacacttatagggtagaattagataaaaagggaaagctctcgtttcttggttcaaaagcct
tcaaacgtgagttctgaaagaaggttatgggaaatgctccatgaaggtggtcaaaatatactgggtgcttccattggcacttaaggttttagg
tcttattctttatggaagaaatgctcagtgctggcattagtgctcacaataatgaaaggtgctccggatgcaaaaatctctgaaacactaaga
ataagttatgagatgttggattccatgcaaatggaataatttttagatagactgctttttaaagggaaagcctaaagacaaagattagatt
tattgaaagcgtgggttataatccaataatgaaattgattttttagttgtagaggtcttgggtacagtaaaagcaggaacatgattgtttaa
gaaaaagttgatggttggaaatgcatgattgcttcaagaaatggggagaaactttgctatcaagaatctcaaaatatacctagtaaacgt
agtcagttgggtctcagaggaacctgactctcatgcttacacaacaaaggtatggtccactcttcaacaattcaaggtgctctagcact
tattgattttcaactgctccataagactttctcatttttaattttctatacagggaaactgaaacaaatcaagcattagttctcaactcaat
aggaatgggacatactatggtcaacgttggagagacaaagcttttcaaaaatagtagcagctaaaaattctcaactttagtttttagagagct
catattcatatcaatttccaagtaacttcaaaagttctctgggaactttgtccactggagaccttggcccttgggttagcaagatagaa

agcagaagatgagaaaaggaagaagaggaagagctttgaattatgcagaactatcgatacaaatccacaaaaaattcttaacaatacacata
aatgttttaattttacattgaaatttgctgcaactacacaaaaatatttctttaaagcttatttttctttttttttttttcttcttcttct
tttagttgaatgaaaaataatctctctccaagaattttgtcacattatgtggttttttttcttttttttttttttttttttttttttttt
tctgttaaaaggttaaaaagaaggaaggaagaagaagaaggttttgaattacacataaatgctctcgctctatccgaaatttactgcaaa
aaatacaaaaaaattatgcagaactctacattacatttaattcaaacctcaggcgtataaaatagtaattttcaaaaaaacaataatattt
tgtaacacctaccatacttagtattatgcttaagtataagattgagatagtaaggtattacgacctctaagaatagtaataatataatata
atataatataatataatataatataatataatataatataatataatataatataatataatataatataatataatataatataatata
taaaaagcgtacgctcaggaaggattacttgctgtcaagaacctaatagaacatgataaaaaaataagcgaaggataaaagaaaagcca
aggaacagcdataaactagcctctgactcagcctcgaacactaaggctggcggaggatataatataatataatataatataatataatata
atatacaaaaatacaaaagtaaacctctatctatcctcaaccttaagaggagcagcgtacataaagtacttggagagtaagctacacata
tatacatatacaaaatagaacaaaaataatccaaggactactcgctttccaggatccagacgcctagcaggagcctctcgactgcactc
tgaaaaacaacaatataatggaatggaacccggagggtctcagcatgtaaaagtgcacgcgtataaaaaaaagggtcttgagaatgctac
aggcaactctagaactccattatacaaatatccaacttaattactaattagaagctataaaataggggtaggtattctaaactctacctaactac
tcaatttcaatcctaactcaacaccaaacctattttccatttccctccatccacacatgcaacgagaacaagcaaccacaacaggg
ttacgcacaagtaatagaacagatagtaaaaaatgaagataaacagatagcaggtgatataatcaatgaacaaaccaagaatgcatagca
atcaaaaacagcaaatgcatatgctgcactctgctcctatggctatgagggcctctgctgggtatccagcaccocccgcaagctcgaaaacc
ttagactatccccgctcgccactctcaagactctgcatagagtccacattcaatcatatataaatcactcaatgggggtatccataccg
gggaatttatacgtcgacggtcacccttacgactcaggttaacagagatcgagaatcaacctggaacacgtgggtggcagcagctttaa
cccaggaaaaactctacgatatcatcattataagcctctcaatagcctcatttcaatcatccagcctcagggcagcctctccctgg
tcccttttatcaaacctccttccacatttttcatcatcattcccttataatctatcttgattacccttttgggtcttgaccaaactattta
tcaaatctgctcccaacctcttaatactgaataatcttaaaactcaacgcaactcaacattcaatcaatcaacacagagatataactttta
actcttgaaccctagactcactaagactagggtggcacaacggccttaactcgcggggccggtcggtgacccggcaaaaagggcgggt
gggtgaaaaatttgggaccgcaaaagcaaaagcccgccctaaccccaacctttaaaccgtgggtctggcgtggggcaggctccctgg
cgggtcgaagctttttttgggtgaaccctattaccatgaaccagacccataaacctgaacacaccocggacccgactgcacgatccact
tcccaccagcgaacacagcaaaaacccctaaattgaatccccactcccccaacctaaattccagatttgcaagctgccacactctcactcctc
aattcctcgccagctactcagacacttagtctctctcagctcctcagcggctctcagcctctcactcctcactcaactcagctcctc
accgtcagttcctcaccgtcactcttagtctcctcactcggctcactcactcagctcctcagctcctcactccactcgtccatcga
cctcagctcgaacttagagttcctcctcaccocctgacggtcgtgagtcggagtcagttcctcaccgtcagattcctcggtcgtcgtggcctac
tgagtataagcctcactgctcgtctggcctcgtcgtcctcctcgtggtcctcggctcctgggtcttgagtttggtctgtgagtcgtgagttg
tccctcactcgtcgtggtctcactcctcggctcagcgcactccagctcctgggtcttggtcttggtcttggtcttggtcttggtcttggtctg
tatttggttgagtttgattatggttgatattgggttaaatctgatattgggtgagtttctatttacattttctatttacacactctctgat
cctcaaggtctcccaactctaacctttgctaattgctattactgaacacataattgtaactaatctcttggtttcaaaatatttaacacaagt
tgataatgatataacttgatgattgtaatttatagtggtgcttgctcaagtgcaagtgcatgggtggtggtatttatgatgacagaagaagggtg
tactaaactcagaagtagcaaaagatttggaacaaaataaaaccagaggcataatagatagtggttttgaggaagcacaatggatataagataatg
ttgttgaggatttcggaataaggttcaatgcatattccttatattttggatttggaacacttgctcttggtcacaacaaggtcactcactca
atggtctatataagcacaatggtttatggatccatcatgctcagcctcatgaggagaattctttatagggttgatctttaaagagaataaat
gattgttggttaaaactctgaaactctgatgcagattcagaattgtttcttggtttttggtggttgcataaaggttgagtttgattatggttg
ataatggttgatattggctgaaactctgaaactctgatgcgagattcagaatttggtttcttggttttttggtggttgcataaaggtgagtttgga
ttatggttgataatgggttgatattggctgaaactctgaaactctgatgcggattcagaacactattgctggttttttgattggttgctaaatgggt
gaaaacactagttcctataaaacatttacattacttttattcaggaaaataaaattcgtaaaactgtagtaaaccttggtactattggagttg
gtctcagggtcctcggctcaggttgatgaacctagtttgaaaaggctgagaccagcaacctccgagtttggaatttttcaaaaagctcag
tccagataaggttaggttagaacgctgctgagtgtaaggtgcaagaaaggtttaaagctggaggttaagcagattggcactctcactataaaa
cggcatctgtagattgtactaaaataagcattgaagataattggtcagactatagcagaattgcaactaaaatgggtcacttaagattgatt
caggagtggttagagatatgtttgctgggtatgttagttgctggggataagccttttaattggttggtgataggagatttagaaaatgggtgaa
atataatagttccactttgaaactctcctctaggaatcaggttaaggctgacatagtgaaagttcacaaaagagaagctgcgaaacttaaaaa
aatttagttccattccaaatagaatttgcttaacatctgatcttggaactcagcagtaaatgaggggtttatagttgactgcacattttg
ttgagagaactggaaaatcacagatgaaattcttaattttgctatagcctcctcctcacaagggatttgagttgcttcaaaaactcttat
gctttgactgagtggaagtgataaaaaagatttttccattactttggataatgctctctcaatgataacttggttgaacactgaaaagt
actttgattgctgaggttcattggttgatgggtgaaattctctcattgctcctcactcatttttaactcttatgtccaanaatggaa
tgaaaatattgggtgtagcaggtgataaagattagagagcttaataagttctgagaanaatctgaaagtgaaagttggttaagtttaaaatg
tgaagataatgagggactgagatagcactgcatatggttagatggtcctattaggtggaattcactttatgcaatgctgcaagtgctatt
ccttataagaaagcttttgaaattgataaagtaaaagaagctgggttaaaggagatttgctcctcattagatgagtgagagaagactgaaaaga
tatgtgattctctgttaaccttttcaagaaactaccaagtgatgtccggaactcttaccacaactcaacttgattttttacaagtttgga
aatcagcttatttttaagtaagttgaaagattgtagaggtcttataaggaactgggagaaaaaattgatttaagttcaagaaatattgg
gaagaaacaggttggttctcagcatttggggacagttcttgatcctagattaaacttaacactttggttcattgctataatgagattgatccta
ttagtgctaaagacaagtgaggttggaagagtaagttatacaagcttttgagggttatgaccaaaaatcctctacaactgtagagagttc
tcccactttcagagtaattttctcaagcaacattctcgtcaattgaaactcagcttattaaaaatggttggtgtaagtttttatactcctg
attatagtggtttctatttttcatgcttttgattcatatattgtttgtaattgcaaggaatttgatgtccgctaatcaagaagctgaagtga
aagtggaagaacccaactggatatttatttgagtgaggcaacattatttgcaatgatgcaatcattgatggttgcaatgggtggaagacaac
catcatcgtttccaacactatcactaaaggcagagattgttgagcattcctattactaccatggctcagaactcgtcattagcatgggt
ctcactctttgaaatagatagagagtcgtttggtgctcagataaattggaagcgggtgattgcaaccagaaattggaatggtggaatgatt
tggtatattatattatgacatacactattgagtgattttgaaacattatattgtaacttttttgctaatatatttaattactttttctagagga
agataaagatcaggaagataatgcaaaaggagaaggctattctcaggagttgggtccaatgatgatttgacttatataagatgaagatgaa
aattaaagtatttaatttaagtttaatttgcttgaagactattataattgcttttgactttggatttatatttttctatgggtctatggga
ctatttataagttataactatggttttgattttgataatggattatggttattttatggaaagactattatatttaattttgatttat
attgatttaaaatataagatattagatagctctataaaatataaaaaaaaataaataattaggcgggtctagcccgccagccgcccggc
cattaggcggggcgggttaggtttgggagcctgaatagcggggcggggcggggcggccagctgcagatggtgggtctcggcggggcgggg
caagctccccgctgacaccctaaactaagacactcctcagcactttattttctctgttttaataatacaaaaataaacctcggaattgcgca
aactttatgtccaggcgtctgctctgaaattagatttctcaacaaactaaagatcataatttttgattttctcagcctcaggaataaagaaaa
accgtgactactatgagtgcaataaaaccagaaaaacagcagtagctatgataatcaaaaattcaataaaaaatcaaatcaaatcaatgact
ttaaaatttaagtttaactttactcatcagtttctctcaatgaaatttgagtttaatacctttttaaagaagattacattacc
ggaagttaagtaaaaatgagttcaaaactctggtttgaaaccaaaagcttagtaccttcaattgaaataacttttatcaaaaatccaatt
agctaaatttgatttggaactcctagctcctccgaaacaagtagtctgggttcaaccacattgctattcaattataagatttaacagta
ttggaagttgctagcttagctgaaactctctttcctcagatttgaccacataatataaaaaatcacaactccaactcctcactcaat
aattctgaaagttagagaaaataaaacaagtaactcaattttataaaaaaataattctgctccaaaacgcaactcgtagaagtgcgagcaag
caaaaacagttgctgcctgtttgacctttctgctgctggacagattttacaacctaaatttaaaaatagccataaaattttataatcaaa

aaaggctcccaaat tttccagtttagttcccttatttcccatggttagcccaacttgggtctcattcaatccgatcattacataat tagttacg
gcataatgcaataccaccacaacaactctacatccttatttaacaaacccaactcctaattccatcataaaataatgtaagagcacaccat taat
aacttccacacctcataat tccaataataaaatccactaataaactctat tccaacaacattacaaggctaatcat taatacacacaacaattc
aactttatctatgggttctcttaacctaaagt tttcacacaacccgtaaatataaacatgccaacttaaacattccctggcgatcacttaatt
cacctaaaggcagctctcaacacaaaaatcacagccctccaagctcaat taaatagcccaaaagcaagccttgcccaacaaagctccaagta
atccaaattcaagtcacatgcataaacaccctcttaaacctacacctaataacataatataatggtccagttcagttttcttatatcaaaaaaag
attgagctagggttaggggttcttaccataccatagctcaatagcttgagcccaagttccggaagctaacctgaaacctagaacatagaa
attagacaagatttaccataggtttccaagttaccaagaagaggatagagattctgaaacttaaggaggcttactagtgaattatttgg
atagaaagtgagctcgacgcgctgagcgcatagccgcgaacgggtgcccgcgacggagccgggacggggagttatgggtgatacaaaagggtta
gggttgagagcttctctcctttctctaatctggtttggcgtgacgaacatcaatgaggaagaagaaggctgctgtttctttatattatg
ggctccggttggaccacgggtcgggttggggcccggttcaaccagttcgccctctcactccgat ttttagtcaaat tttcgaaat tgatatca
aaattctcgttttgacgagctctatccttttttaaatatagttttgacttttttagcttttctaataaaaaatgcaatttttaactaatttaata
ccgat ttttagcggggtttacatcctaccacctaattagaat tttgcctcaaaat tcaactcagttatctgaaaagagggtggtggtagttc
tccgcatatttggagtttcttgagcagcttaccacttccagcttcaacttttcaacttctctagtacataacttctactctgtctggctgcttg
gccatggatcttattatatactcaaccgaggttactcgaagcgtgagactcttctcaattgcattgattcaagtttcaatcgtgattgatt
attagaagtgactcctaaagttgagatggttaactggttaaggaatataaggataggttattctgatttccgatgattccgatctcaattct
cttcagaatctgacaagttcgtccttgccattactccatcttgactcactcactgatttcaacttgttaagtaaccttctcaagtttaggac
ttccttggttcatgtaactctttagtcggtttttatagtaagattggttactcgaatttgcgtaaacacttctctgctgctcggcttataaatt
cgggaactaaggtatttgatgctctagtttttatcaatcggtagcatttagcctctgtaagctccatctctcctgctcctcaaacatc
taatacgaacccgataattactcttcttacttccgatgtaaaatcttcaatgggtgacgat tccgaaatagtttctagtggtccccccttact
tcttcatgctttaaaacgccttaatacatgtccactttattcttatttctaaataacccaaaaacattttcttgattcatggtacgctctg
gcagctctcatggttaactcgaaaaactcttctcatgagtttcagacaacttcaactctgcttaactctgatttaaacacttaacactctctc
ggatctctctgattcaacaggttttagtactcttagtgcctttactcaactctattgctctcctaagcccttaactctacgatctctattt
cggcat ttcggttaggcatataaatctcttttaattctcttctgtttaccggagccttgatctcttagcggctccttattgctcactatgct
tccgcactctcatgataaccatggtgtaaacggtaaatagacttagtccgat tttcactctacactctcat tctcaacttagaacttcaac
tcatataactttctcctttgaaacttaactcaagaactccttataaactcttgcttggggcgtctactcaaacctcatgctgctcaca
gcacgcttctcaagtcccccaataatgcatggttagtaatttcttaaatggctcttattgcttaagttagagaagttggcataaaagatcaatc
ttctttccatgggttgaagacttccagcatatctggat tttcaactcacaacgggtgacccctgctgcaattatttagtcacagagaatcaaatcgg
tgaggattagggtttgaatttctcgtaatggtgctggtttacttctcatggttccaagcttggctttaaaggatgacactcaataatgca
tttaagggttatgctgacgcaaaatccaatcagagtttttttcaaaagaaagttatgctcaagggatgaacgaatagcatgaacgaggt
cataagaattcaaaaaaagtcttatatacagtttaaccggagttgtacagaatcaataggatgcatagaagaagaaacctagttgcaattcaga
aaagaagtgcataaaagcttaaatagaagaataaaggcatggtggatggtataggttctagttgggtgaaaggaatgcgagttcccaaac
gggcaatttccgcaccgcagatcactaggattcaaaaaattagtgatttatattaggatgagttcaagttaaatcaaaaagagacgattgt
acgagtagggaacagatcgaagaacaatgggtttttacagaagaaggtcaaggttaggacttgaatgtaggctacaagagaagggttagat
taagtctgtaagattgtagacgctctcgcgaagaacacctagttgagatgattctctcctccagtttcttcaactcggcttctgaaact
tactgattctaatataatcttcaacttctggtgcaacccaactgattcaaatcttgacactacgtctatgacaacactcagctcttctaaggt
ggaaattctagtacattttcaagaacggtcttagaaactctctgtagataggaattcagttcactcctcacttctcggatcaatttacaact
aaaggttgaacttctctattccttctccttgaaatcttaactcactggatttcaaacggtaactcggcatagctctcaatggttccaattc
cttagataaattcaaccggcttccacttcatctcggatccttctccaagaacctaaccttcatctaaagttagctaaagttagctcctc
ctcgaacaactagtaggcttctgatttaactcaggtttggaactcactcactcaaaaaacttgctgctcacaacttggatgctacatata
tgatacgaaggttagttaaanaactcaactgtaacttaaggttacctcgggtctgaagatcggatcgaactgattcttagcat tttctgtg
acagctcagagctattgctggttctcccgactgtaaacatcgcctagctggctcgtggaactctctgggagtagctctccataaact
tctgtaaaccttctgactctcaggttaacgttctcggatgacttctcaaaccaactcagggaactcactaaacttctctcagactcagc
agttgaagatgctctctcaagcccttctcactgtaaaccaactcactcgtataggaattcagttcactcctcacttctcggatgcaattcaga
acacaaattgctgaattcagcgggaatagtcagcaaccgacatattcttctgctcagttgcat taaactccaactcctttgcaatgcaatagct
gagcttaggattttgcttagtaagaattttataaaaaatagtcgctggtgtagataatgcttcaaaccaaaaaatcctctcgtgcaaacgt
ttggttgcacaagtaacaaaccccaataaaatgataaacccaaggttttaaacctcgggtgctcctcctcaaggaattgcaaggaaggtatg
cttatattggctatgaaaaaggtaaaat tgggggttttggaaatagggaagaatgacaaagtaattcaaatgacaactaaaataaaataat
aactataaaataaactctggaagatagagaactggaagctcactcacttagttatcttatcaattgtagagaaatggatttttctcccac
ttgttaaccttttaactatgaaggttaagttaagtgaatgtaattcaactcaagctcaggtccaggtcttctctggaaggttaggtattg
gatctcgaatttaattcttgaanaattcaaat tttcagttcaacaatgagtttgataactcaaggttaccataatcaaccaaagccaaaagg
aataaaactcacttgaatgaaataatttggatagaaccaagcatctataacataaagt tgaanaactcctcaaatatataaataaaata
tcaatttaaacatggctggacaataatcagcacaattgggcaacatacaataaacacaaaataaaataatcagagtaataaaagtagaagaga
aacataaattttgaaactgtaagaagaataactcaactcaataatagaanaactcaataactcaaaaactcaagagagagagaaacactctc
ctaaaaactacatcaaaaactaaaattgtaatttagagagcctctgaggtctctgcaagttccctgactttaaactctgtgtttctgggocgaaa
actgggtgaaatgcccggcccaaaatctctgcccagcatttttgtaattctgcaagttgcccagctcagcgcctccgcgtcatccaggtgacg
tcaactcagcgtttttctgtatcagcgcgtgcccgtctccacagcctcgcgtgcttttgcaagcttcaactcgcgcgctcagctcagcgcgca
tgcgctcactctgattttcttccatttccacgggtgcccgtgagccatgagaccgctgacttctcgtggtcatctcctcaattctctgttcc
ttccatttttccacgctcctcttattctcctaagccattcctgcctatgaaactgaaacacttaacacacagctcaaggtcagatggtta
atagagagagattagattagctaaattaaaccaaagaagcatggtttcaactatagaaactaaactaggaaggaattgtaaaatcatgcaaat
catatgataaagtgggtgaaaagcttgataaaaactcaattaaaggttaaaactcaataataaaactcaaaaatagtggtttattggggcagttcaactct
ttgacttttcaatgatttctccactttgcatgcttccctctcactcctttgactccattcttagccttttcaactctgaaatcactaacaacat
atcaaggcatctaataagaaatcaaggtgaaatgaaatttagtaatttaaggtctagaagcatggtttcacactaaagcacaataataggagaca
atcacaanaactgctatttcaatggataaaatgaggttaaaaggtataaaactctctaaaactcaatgcaagataaacctgcaaacgggttt
atcatttatcaatagcattgaagaataattttctgtagaattcctgcttgaactcactcaaggaatattctgttaactctatttgcaggtggt
atataactctgcccacttctgagaatctcctccaacatttaagctactctccacagactgacttctgggtatgctgaggtgataaag
aactctcctcaactctgaaaccacagatcagtgacaacccaactttgacggtttgtgtgattgaaatttagagtttagagtttagactttgt
cacatttagcctataaattagcatggttttgattctctccatattgtgctcaaggtgaaataaacactgctttttaaagcctttatattgatgaa
ttctaatctctcttgattccataagatgctctgaggttggtagtaattcaggtgaaataggctaggcatggaatcaaggaagcaagga
aggaagcatcaaggtgagagaagcacaanaagccaaagattgactcagggcatgacgcgtaacgcgcaaacggcctcgcgcgcacattgca
aagtaggcaaaaggcgcacgcgtaccgtgcccgcacgcccagtagtgccacatgacctcaatgcaacaaagctcaactcaatgcaacaaagcttggag
ggtttctgaacccattttggcgcacaatgctaaggaagaagaataaaaggatgaaggat taaggggaagaagaagaaaggggagaaatgaactagga
ataacacacacactaggattagtttttagtttagttcttagagagaagctcactctctcagaaattagaattaggaattaggttaggttaatt
ctctctcttagacttaggttttaattctgctttgattcattctcttatacaactcctcaactctctctctctctctctctctctctctctctct
ttaataattgtaattcttcat tttggtttggatagattggttctctttat tctcttcaataatgcaatttgagataaattcatgtgattggtga
tggtggtgattggttgatttaattctttgcaataattgtagattagattttatctgtgtgcaatttactatgctttctctctctctctccca

agtgtttgatgaaatgcttgggtggatttttagttagatatttggctcctctgggttaggttagtaatttagtaaaccttgagttatctaattcc
ttgtttagtgataattgggagagatgctaaattgggttggagtgactaaagctagtttttcttgggagtggttaggactttggctcaagt
caattcatccacttactttcctttatttagtaagggtaactaagtggttagcaatgaacaattctcatcacaatgagaaggataactaggat
aggacttctagttcctcacacctcacaagagccttttagttgttagcttactttctgcatcttatctccatgctcttatcaaaacccca
aaataactcacaacataacaagacactctcatgttaattcttaggggagaaacccgaggtccaataactctcggttataaaattttagggttt
gtactagtacacaaacttttggtaacgaaggattatgttgggttagagactatactttacaacgagatttcatagtgaaattctaaacc
gtcaaaaatccaatcgtaaaaatggcgcgctggccgggatgcaatgggtgttattgttattgggtattgtatataatgtgaaatgtaaatatc
ttgcttttggcttcatttgctagtgtagaattttatctcccttccaccatgaaattctcactttggctatgagtttgggttcaaatatgttgt
aggtagtgagagcttcaatgaggatgtgtatcaaggatggaacaatcaagggtgggagagccatagcatatgatcaattctcatggcaaca
cctccaccaatgcaactatgaagaagaccatctctatgatgcatatcaacctaatagctatgggaaatcactttgtgatttcaagaaccaccac
catatgctatgaaatcatatctcaactgagcggcaaccattctcacaagcctctcataccaagcacttctataatccaatccaatcaca
tgaccaaccatccaacatataattttatgacgatatgagcaaacactttagaaccaccaaacctatctgatcactccaagaaccact
caatacacacaatctccacaaccttaccaaaaggaaacccttcttattatgaaccttctctcaaaaataatgaaccttctaccaccacacaag
ccccaatagacgatctctcactctgttactcaaggacaagaagcctgaagcgggatacacttgagtttggaccaacttgactagggtagt
gtccaccttagcctaccgatgtttgaaactcaaggtacttccgtagccatattgtgaaaaacctaagaagagcaaggtatgaggagagattg
ggaaactcgggtggaaaatgaggaggctagggttatttagagca

>AD25F09-PC-F09.RM

aagcttggcgcctcctcaaatggaaaattataaattttattttagggaagaaattacgtttatcatgagtgcaatttcttaggtgctaac
cttttcaaaaataaatacagggagagaagggtgagatgatgttcatgttggctgtttcagttccagctttcatcaaatcttgaatgtcttgctta
cataccagcgggtcctgaactctcactgtgatactccataataatcgtaggaaataaacatttgggaagagataatctggcaagattgggt
caacaaaacagtcactctcaagcatagaagaaaaaatgggagaaatccaatcaccggcattgcaagtaagctcgggtgatgaattcttc
gcatgacagataaacgaggctcctctgtagcgcctgattttgattctcaactcggaaacagatgcaagtgcaatgggtttcagatgctgga
aatgtggctgtttatcctgttaccttctccctggaacaacctgatgtgtatttctccctgcaaccaactttgcatgagctcatcaccat
cttggctagctgagctggcgaaggacagcagcctttgatataacatcatgacatcggatgcatccaatgatttacaacggctgctgcaa
ttctctcagacttgatggagctcgggcaaaaactccagatcagggaactcattaggtccagatgctgagcctgggaaatctcaggatgctt
attgaaacctcaacacatttctccagataaaaagcaatgaggttagagtagaagagaatccaagaattggaaggagaatttggggtcag
aaacatccaagcggtagggacttcaaccagatattgggtgctgggaaggcaagtaagccttttgcattctgttaagccaactcagataggcc
aaccaaatctccaatcgtcgtgggtagctcgttatgctgtacaactcaaaagtaagaatggataactgtttcatgcatccccaatctggt
agttttctcaaacctattgcaactcatalagatctagttccttgagtgaaactcatctcaatttatctccaagcgtttcgagactccaacttga
ttaaattcaatccaagaagggtttttatgggtgtagagagattggggaatgtcattcagctctgaacagcatgaaagatcaaggtttcaagatt
ggagcccagaaagatcaggcgtttgctcagggttggctgtttgacagatttaagttactccaactttctagaactgaacaaatgaaata
tcagagcatctactcttggctgtaaatgttaattgtaataaaaacttgatagtgcttttgatgaaagggtttaccttttccatgcccacagct
gtacaattttgctaaagatacagatcaattccaacaagtctcatagtgcaactctgtaaggggtagagttccatggaacaaactcagctgcaa
aactctaaogtacaggaatattgcagagaatgggagatttcacgcgctcaaaaatgaggagcttagctgacatattgtgagaagctaaa
tcatccaatcagttcatccgcttggctatagaacaatgctcagctgtgctttagtttccatcacaagaacccgttgaaagctctatgc
tgattgaaacaaaaataaataaatacaactaaataaattggctccttagttacttaggtagttatataatctctatacactttataattgtgaa
aaagagtgggcaccctacctgttttggctaaagtagcagatcgtcttcaaccacacactgcttgccttagctatcagcatcctgggagatt
cttgatttacaataagtttggccattcttcaatcagatcatgcaattctcagagttactcccccatacttatctgtcctatgtgaccaatga
tctataatcaaaaatcaataccaatttccgcatgatgaccacaatatttaataatctttgtgcaataatctcttgactctcttgaagaaa
catgaaatcttagaaaaatgttctttccatggaatctcaaacatcattgcttatttcaatacatcaataatcggaaatgggaagattct
ttattttccaatagcactatgcaaacctcaatggatctgcaataaagataggaccccaactttaagtgcaaaaggaagaccgcaactata
attgactctctttggaacaaatcaaaaaccccttctgtaggttctggcagtttaaaagcttttgaacaaaagagattaaaggcttcaatttcc
actaatcttcaactcaaaagtcttctgcaccccttggctccttagcagcaacttggaactgtggtctctgggttgtaatttttctgctgcaaccaa
accaactctgctctccagctaaattctctaaattgcttttcatgatttacatcatcaagaacaagaagaactttttgagacgtaacaggtttg
aagttatgctctccatcactcgtataaaagagcagttgaactaatcccaatttgatcaagaagttggttttggatgtaggaatattctttt
ttctcacatcgctccttagctcagcaagaagcaagcaactcaaatctactcgaatagtttcaaaagacagctctagcaatagttgtcttac
ctatgcccgccttcccgtatgctatagcgaacatcaatcaatccaagcctatttgacaacactgttccactttgcaactttgcaatcaatccc
tataagattcttcaatgaaagatggcaactttggaatcagttttcagtatattggttagaagattcttccacaaggtgctcctctcattc
aggaagattattcttggtaaatctagtttagacaaaagacgagaataaaagaaaacacttttctttagaaaaaaatataaataagagaagat
aaacaaatgcaactcaaaaaatcttcttcttcaaccgctactggaaactgtggtgcaacaccagctccagctcaaatagattatct
ttcgtctcatttggcaaacctgagctcactctcaactgaccttattgtctaaaatccccatgagtttttagctcaagaaaaatcagttata
acgaaagcagttttaaattgaaatattacttgaaaatagtttaatacataatctatgtattatttatttgaataacatttttaattttata
cttgatattcttataacctcaatttagatgccatattataaactttataattagtttctcaaatcttaaatcaagatattctcaaacct
tgaaaaataaaaacaaaaacacatgcaacagtttttaggttaacatataaattagtttttaattttttagtttacttaaaaaatgaagtaaaa
agataactcaagaagaaaatgagactatgtagctaaaataagattgttggctatattggttagatataaaaaaaaaaaaaagctgaaaaagag
aaaacaaactaaagcattttgttcaactatctctgagctttgaaactcaataataataataaataaatttttactaagtcacctcccatttt
taggtaaagcacaatttaattttacttttataataaatactctcaataattcttattttaaacttagtctctatttttaggaaagatcac
ccaccaaaaatttgggtgggctagggtatctatcaaaatttatcccttaatttagttattatataaaaaatttaacaaaataattataaataat
atttttataattatctcaaatattcttaattttaaattttaaactcaaaactcaatctcccttaaatttaaacttaaaccttaaaaaataa
aatacaagtaaaataaataaaataaataaataaaataaatacaactcaaaaactcatttttcaggttcaaggtgggtataataattat
aataagagtaataaacacaataatttttggatgaattttggatgtattaaaaaatttttagcaatataataattttatattgttaaaaaat
tatttatataaaaaatataatgacgaatttaaatgaaattttggatgtagtatacaaaaatattttgtatgtgttatctattattttggatctac
tgaagagttatttatataaaaaatataaacacgaattttagatattgacaaaaatattttgtgattattatttatttttggatgtta
taaacataagtaaaataaataaaagaaatacaaacgaatcctgtagcaagaagaaatgttgagcaataaataatagcttttcaaaaatac
aatattactggatataaaataaaatttttttagtgacatttttcttaacttatatttctcgggataaactttcacaatataacaaaacata
tataatatttcaattctgtaggggattaaactaaatagcaaacataatattacaaaattttataaattcagtttaaaaaataaataattat
tatttctaaatttttaaaatataaaacaaatttggtaaaatggcttaaaacaaactatttcaatgttttctattttaaacaagacagcttctc
aacaaacaaataattacatttcttaaaataaaaaacattgaaataaaaaatatttaaaaaacaaataaaaaaaatcagcaaaccttgaacatg
catagaataaagaagtaagatttagttggcttaccgattttggagtgcaaccagaaataagcagcaacttgggttaacgatctcgcaactct
ttaaactctcaccgctctgctcctgctgcttgggtgagagctcctcgaaggttcccttttgggtgctcactcagaaggtccacatcat
agaagcagctcagatgtgttggcagccttctgtcagcgaaggattttgcaagcctcactcaaacaccactgggaagagatagttgggtg
tgagagaaactataacggcaaacatagactctcaattgctttcaggagttcatcggaaataaacatcgctttggcaagattgtgtcatcaata
tacgttggatctctcctggttgagagcagcatagagtagcactgagcttagctcctcctcaaacctcaagaacacagctgatt
aggttagtatgattcttggtaggttaagggaagaggtctcaacctcaaaccaactgcccattgaaacaacaataagctgtgtttaggtttgga
tgtcgcagctctgacaatggggtaaaatttagtcgcaatttttagagataaaataaataaataaataaataaataaataaataaataaataa
taataa

aataatgagagtgatcgccgctgtttgtggattaaaggaatggcgtgaaactgctgggatagctccgttaat acgacaaaaggagaagggtat
aattataat tcat tgact tgact ttgttattgcaaaactcgttcgtagtgggtaaacattatttttagtgggttggttacogt ttgcaaacgg
gagtttgtataaaaatgat ttgcaaaccttgat ttgggt taaaat taggt ttgtgt taaaat gat ttatgt ttggtaatt ttgtat caaaatg
gattatgtaaaaataaatgttttggttataccatt taaaactcacttttagaaaaaaat taaataaaatggcaactcaacttagat tttttttt
atattatccctatttttaatttagatatttaaatagatttttagtttaattttataataaaaat taatattacttactaaaaataaaaataac
atataaaaaatagacaaaatatttttttaataaaaaataaaaat tatataatttttagatataataataataataataataataataaaagaatatt
taatacaaatatgttacattttttaagtataaaatgtcatttttagtattttttacatcgtactcttattctagtactctcaataatcttattct
taataataatttgaatccaatgtttatgat tttatttatatatgat taat tttttatttatttctttttaataaaatcataat ttat
atataaaaaat tacaactaaaat tagtataggagaat tacaattataaaaaat tagtactgaaaaataaataaaaaaataagtgtatgaaacaa
at ttaaaagtctt taat gat tttttatataataat ttttaata ttttttagtactcttttttagtattgtataattttttactattattatt
at tttacagtttaattttttattttactttaccctaaatagaattaaataaatcatattataaaaaagataataacaaacataat acataaaaaatc
acgat tataaaaagtataatgtcaaacaaaatagttaaaaaaataaaaataaataaaaatagaaaat tagagctcatataaaatagaaaataacaa
gaattttataaaataatacaataacacatacaaaaggataaagttggtaaaaagtaaaaataaaagggttagtatctatggcctaaaagcccggttaagaa
aacgcagaagctaaaaatagttgcttcttgtaaacgctggttttggaacacagatacactctctgcgttcgatgagaaaaaaaat ttgccaaccaa
aagttgaaactttcaaaaaat taaacgctgtttctctctccaacaggtttccaaacacacccttagagttgggcagcagatagttcagat
tcaagttcgtctaagttatgggttataaaaaactagttgttttaatttttttttttttgcataaaatcctcaattttcaagttagaat ttttttat
gatcagttcaatgagttgtatatttaaaaatacaaaatggcaaaaaaaaat taaatttttagtataaaaaaaat ttgcaaaaaaat ttttttg
tgtagttacaatttatatagttataaattttttctagacataaattttcttaaatagcatttttagttagctttcttatataaaaaatct
tttttaacttttagttttcaatttttttacaattttttgaccctgcaaaatagttttctggcacttaaaaagtattttcccaaacgctc
aaaaataaatttataaaaaaaat tattaataaaaat tttctacaatacactaacaccacacactacacatacagatttaactcaaatcagcacaaa
tttagtttaagattttctacaaaacatagttctataaatttttatagttttaatttactctgttttagtttttatagtttttcaaaat ttttaatt
ggctcctaactttttttctcaatttgactttatcacoggttttaatttttataaataaaaat tttctataaaaaaaat tagagttact
gtatattttctataaatttgaaggtctttgttaataaaaaacttaataaacttttaaacgcagttgtat tttggtaaaaaat ttagtttaatt ttaa
tatttttaatatgaaaaagattgaatttataaaaat taaagtagtgacagatttgaccaat taaagaaaaaaaat taaaaaat ttaataaa
aattttggtaaaat tataagaattaaacaaagtaataaacat tttcatataat tttcaagtttaataaagactatataaaaaat ttaacaaaat
actctcgaagtcactgtggt taaactctgcctgtggtgtataggagacact taaataagattttacaaagatttatcaagatagttataactgcagtag
actttacaatctcctaagattttacagatatttacaacacagacactatgtttctaaacaaataaaggtccaccaattcaagtgacttagact
atcaacgttcttctgaaaaaggaataaagaacaataggataaataagttttgttaataatagttgtaaacaaagcagttatactcaacacagagaca
ttgtacaaatattgacacacttttagctgcctaaaaataaataat tagtataaataaaaatata tcaaacacatagaatataat ttaata
ataaatatgtatagaaatatttccaacacatagatgattagtgataaaat taaataat ttaaaatttttaatttttataaaaaaaataa
acaaaaagatttatctatctagaaaatcattccacacaaatagacataaaaccaccagcaccaatagacataaaatcacacttagatgatttttct
ttgtgtgaatttgaagaacatgttccactaattttataaataaataacagaaaataaccattttatcttatatgggtctgaaaataaataat
at tttgattaaagctaaagttacacaaat tttagctcactactcctgtagttttctgtcactacacaaaaaatgttgagttactatcagatttaccg
ttagattcaacgtcgttagagtaactgcagatttaccggcgaatttactgatgagttggacatacaaatcgtcaagttaaatttttcaacgga
tttaacaattctgatggttaaaaatgcccgttaataat tgaaggttaaaataaaaaaaat tagcgcattctttatctttagggtcagattttcgggtg
aaaaaactaataagataaacccaatttaagaaacgctcgttttggcactccgcaatcagtcggcgttgcaatttttcgaggttaaacccagatgg
taactgtgccaataaatttgaagcgcaattctctctccctcattcgaaattttctctctctctctctctctctataactctctcggctcccactctct
ctctaccctctcatttcccgcgacaacccctctctcgcagagcactcctcctcgttctttttgtgtcacctgttagttgcccgaattctaa
cgattgtcaccgctactgctgcccctacaccgctatcgcctatgcccgtct
aacttaggtctactgaaatttcgattttggtatagttttataagttttatttggctgttttgagactctgattttggttatttagttttagttc
taaacatagaatactgaaatttttaaaacttagggttttttaaaattctaaatttagtggttctgaacttattagttttaatttaataaa
catattctgattttggttctcttttaatttagattcttagggtatactgaactgttttaaacctaaagattttctaaaattctgaaatttaggt
actgaactgtttgatttaaaatttttaataaagcattctcttattct
ttaggtatttttaaaattctgaacttagggtactgaacttattagtttcaatttttaataagaacatgttccgattctaaataaaaatagttt
at tttggtgattttggtctgaaatttagttattataaatttataaaaat taaattttgtgtgctgatttttttatattggtgattattgtgattttg
tctctttggtttggttttgattaaagacttcgatagtaaaacacattgattttcaatggttttaagagtaaaaatgatagtaaaataaaaaaaat
caaatagtagatttttaaaattacttatgaattgacttttttttagttagcggcaaaaaataacatgtagtaataattgctaaatattgt
tgaagttagaataatgcaacttaoctcttttggttttatttgatttcaacggataatataattctttaactaaagtttaacgtagtttat tttgacta
agcatttgaatgatgttggtatgttttagcactttttaaattggtatgctctttgcagacatgacaataaggttagaggttatgcccgtcagcc
tagtggctcatagtaggggaggtttctgctggtaacccctgtaattctagactcctcctcctctagtcogactaccagtagctcacaggttg
ggagtgcatcggaaccagccattcatcatggtccctaaatccaactacgtccctacgtcctacggtgacactgcccgcctccatctgcttagtagc
tcaactgtctcggcagcagccactccaacacggacacggcgctggaattctctcatggaagccagtaactgatgccccaccacccctcgca
tcatacatagcagatttgactgatggcggaacgacgaaagtacttttaagcttttatactgctgaaagtgttgataaagctggtgattag
gattcttgatattgacttagtttttagtggatttaggttttaggttttaaaattgataaaaagtgtttgataataactgatattgattcttgatt
attgattctagtttagtggatttagtggattttaggtggttctgatttatgactgtgttcaatgagttgtgacaattgggtgctgtttag
agtttaattgttgatggaacagaatttggggcgactccagttatagatgaaattttgtcgaaattgtgaagaactctataatttagtttatt
gcttctgaaattttaaatttatattgctcatattgggttat tttgtaattgttgatagttttacaccaaacacacacagcagctgactcagatgacca
acatcatcaacttgatatacagacctccatggccgaaccacaaaaaagatcccctttgagatcagagaggtgggttttgaagttggacggta
aagcttaataatagttcaaaccttttagctatgatatgatatatttgattaaactaattttaaagttttggtgtgtagctaaactttataggac
aggagcagcatgcccctcat taagaaaaat acgacctgaaatgggttaagcagttacaatagatgctggaggatgtacgtgaaccacactcact
ctttggcttcgcccagaataatcaagaaggcttagtaactgttactggcgcgatgagaggttcagggcatcgggtctcacaagaagaaacacaaa
tattagccaggtgctccaagtatacggcggtcaacaactaaagacaggtggtatgtaactctctcaat tttggttat taaacttagttatatt
ctttgacatacat tagtaaacctgttgaaaactgatacactcttaacctctgtaactatactgctatgattataaattacattacaataaaat
aaattagttaaaaatacatcccttttaaaatggcattatgaaaaatggaagcaagagactataaagctaccaaataccaatcagatttact
cctgtttttttttttccggtcaaacgtgttaaaacataaaaat ttaattttcagttgcttctgactagaaactgaaactggctgggtgaacttagct
agaaaaaggtggacagtaatttcaactcaagtaggttcaaattggttataaatttataaaacaaagtgaagtttggttacattttactgatgtaaaact
caataattctcaagtaactgactttcttttagtataagcagttatgcagttttccatagataaaaaaattttagtaagataactcgtgaaataaag
gttat tacttttataat tttat tagagttgatttcaattgaaatggtttcaagctcgataaaaaaccataaccactgaagatagaataaataccaa
aagagtagaattgattactcactctgaaacacaaacaaatctactttagttacaagttgaaacatgaaaattcacctaatgtacaacttaaca
agataacttgcgttcaaaatgagttgcaaaaattcattatgcat tttctgagactgcatctatgcatctcctccatgtttcatctggaaatggtt
attctacttcaaatcaaacacatcagcttccaatt
tctttctgttgggctctcagtttccctttgttcaaatgctcaactcactcagttatgcaccacacacagcagagctcaatattttggg
ggatattcaacattgatttttcaaacgcagctcaaaatgaattggctttccaggtttcttgatcacttaaacacatctggtcatcaaaaaa
tacacaggtgataaccttgatagcaaaataggaatgctgctgttagtactatcaccatggcacaacggcgaacaggtgagagggaaaa
ccagcgtttggaagccgacttaggaaacataaagcaatcctgacctttcacttggatggcaatgatgtgggaaaggggactaataagcagtt
ctttctgtctatgacaaaaacatgatggaatattgtcctggaatcaccatttgaagacccaccagggcctaggttagtgcaaaaa

ccgcgcttatcccatatgtaatgtagttcttttatattcaaaaaccatataagcagttatatagtattacagacccaacaggaattactaaaa
ctttaaattcaaatataatgtcaaatataatgtatttagctagtagtataagactaaataatcttttggctgaacatttgctaaaa
aaattttatcgggtccccaccatccaactaatgtactaaagggcctcaacacatcaataatttaccattttcagagttgcatagaccaatt
acttggattgtgtaactaatagaaaatttttaattgtgtgaagaaacctctgtctttcaaaacaaataatagaagaatgaaatatttt
gaaaataatagcagaaaactcgccattctctcaaaaagaaaagactgacactcggcctctaccatttgctttgagacaattgaaacctccat
taagaatttgtaagagtaatatcaccaataataatagtagtagtagagttagaagtataacagtagagaatgagaaaggcataaagtagtca
aaggggagtgagtagtacttgacccccaaaattttgaaactatctttatatagaatgtgctcaaaagataaaaaaaaaacttatgtgga
tagtagtttatatgtaattat

>AD25F09-PL-B03.FM

gaatcaaacacttggcacttctgtcacaattggagcccaacaatacaaaccaagcccttgaagatccatcatgggttaaagccatgtaagag
agcttggccaaattgacaagaatgaggttggaccttagtacctcatcggatggttaagaaagtaccgggtaactaagtggtttaaataa
actttagaggtggaaacttggacttgcgaagagtagtataggcctcaaggtacgatcaagaagagggatggttggattgagttcttt
gctccggtagctagaatggaagcaattaggtgctctctgacctgacgccccataaaggttcaaaaatttcaaattggatgtaaaatgtgct
tccctaatggccttatgatagagaagtgtagtggctcaacccccgggtttgaacataaagaaatttccaaatcattgttttaatttaactaa
ggctctttatggcttagacaagctccaagagcttggtatgaaaggcttagtgcctctctgtggaaaatcattttcaaggggaacccaac
actacttggctcaataagtttcaatgatgacatcctctgttcaagttatgtggatgatattggttggatcagcccaatgagctctgt
gtgaagagttagaaaaactcatgactagtagtggatggaaatgaggttaatggagagactaaacttctctctggcctccaaataaacaactc
ctagtggcactttatcccaaggcaagtatgcaaagaatataaaagaaattggcttagaaaatccaaaccaaattggaacccaatgca
tcaatacaactaaacttgaagaaggatgataatggccaagatgggtgaacacaggtatagaggatgactgggttcaactatgcttctct
ctagaccggatatttcaaagtggtgtagtctcactttcaatctcaccacaaaggaatccatctcagcggctgtaaacgcatcatta
gatatattaaggaacaaagtagtatggcttatggatcctaaactaatgaccttttggtagatgggttttggtagcagatgtagcgggtg
tagagtggatagcagaagtagtctccggcagtagtctccctggagctcactcaacaatgggtcaagcaagaacaaagcagcaggtggctct
tcccagctgaagctgaaatataatccgcatctgctgtgttccaactaatttgggttaaaaacacaattggaagtcaaaatcaaaatca
atagtatacctctgttttgatataatgagtgccatataatatacagaagatcattgtttgactcaagaactatgacatgaaatcaata
tcaattataagagaactgtgcaaaaggtagctattgataattcaatttggaactgaagaacaacttgctgatattttcaaaaacccct
tgtgaagacagatcctattataagaaaagcttgggaatttggtatttaagctctgttgaaaaattttgaaaatttggatctcaatt
atctcatagggttggacagataaaatgcatgcatgtaggaggagctatgataaagggaggcaacgcagaaaataaattttatgggccccac
ctaacctttctgaaacccatcattaccgttaagttgattctctctcttgaaaaaaataacacaaactctcttacttctctcaaatct
gttggaaaatacattgtcaaatcaaatctctctcaactaatccttgaatcaaggaactcttcttggaaacctcttggtaataaac
cgtgctataatggctcatttaactctccattatctctctcaactcaacatttaagtccttgaacctccctccactaccatgtctctggccat
tcttcatttcccactccatcactctctcaccatgaaaaaagaaaccgcccccagaaaagtgaacgcattctctctctcaaaaatctct
actccacagactcactcattccatctcaactcattctctctcaccaccgctcaaaaacaaacaccaccatgagaagaagaa
taattgcaacaaaaagtctcagggccgaggaagaaacaagcaactgttagtgaagaagaagagggaatcaactctcctctcttggaa
ctgaaatggaccccatgggtcattatcagaactgggttaaccgggctgggtcgggaattctatgcaaacatgctgtatgtagggctcaa
tccactcttatgtcaagcagctcatgactctcactcctcaaccataagcgtccttggatagaagatgaagaccggaaggtacat
gactgacaaatggagtagcaagttagagctctcacaacaaaacggttttgcaacaaactctgtgcaaacctcttggatagatgagcttaa
accatcgagcacttggctccgagaactcagctctacaccgcatattacaccacattctcactccacaaagtgggtctcattatagagt
ttctgactcatttatttatttggctcttattaccctcttcccactcatttggcttatattatgatagctcattatgtggagggagat
cacaacaaaggctaaactcttcttaaggatgttcttaccatgattccttggatatttaaggttgaatttgctctaaatgaaatggtt
gtatcaatgttaagggaggaggagagctctgagcaaaaagggaaagcaccacatcctctgacccatccatcagcgtctgagagctgg
ccctgaacccatccatgtgactgcatcaataggggaagtctcaatgaaattctgcacatgacacaaatgtagctcaaatccacaaggct
ctggaatttggcgtatgaaaatgagaaggttggaaaaatgtcatgaaaggggtggctcattgatccaaagctttgaggaagaat
aggaaaatgaaagaagtctagttctgaattcaactcttctgtgtagcttttgtagactgtttcttactgttatgtgattatgagc
tcaatcttttggtagaagcactgtagtactcactgctttctgatactttgatacactcttactactttatattgataattttggttag
gtaagtttttggtagaggttccccttgatgacaaaaggggggagaaataaaaaatgaaacagaaacaggggaggagaaattcaaacagg
ggatgaaatttcaaatgaaatcattgatcaccctgttactgctgctggaaatgcatgttgaatttgcatgtgtttatgtctgtaag
cttggcatttgatttattatgatttatttgaatgataaaatttggatatttggatgcatattctgaaactcttggtaaaaaattg
cttatttcaacttgaataaactgaaatttttggcagtttaataactgtgattaaactgcccacatttgattatgtgtagtcaaaaat
ttcttacttgaataatggctctgatatttggaaaaatttgaacaacatttgattttgtagcagttgtttgtagcagtttgtagtga
aaaatcaattatgataacaaatgagttttgattaccacaaactgctcctcaataatcagcatcacaactttcaaaaatatttgaata
cattgttacttgaagcttgatataatgttataggttagtgcttcccctgataaaaaatggcattatggagggaggctatgtgacaat
agggggagaaatcaaatcaaaaggaggtatcttctactcaactcttaaaattctctccattcaattttaataatggttgcatcaag
aggtgagtaggttggaaaactcacaataatatttgtagactaaacattatataatgataaattgcaaaaataatttggatttaaggt
gattaaatcaatttataatcagggtttatggggcgcaaaataaaaataatttgtagcaagcccaaaataaaaacacaaactcagct
gtgtttcttattatgtagctgaaattgttgcttaaaagggccaagctcaatgttggtagacttagtccatattgattttgatgaa
tatgcttgatacaaatccatcaggaagcaaatgtacaagttggggcaagaataaacattatgcaaccaagccccacataatgaaag
catgcttaccgaatccatgaaagcaagcccaactcactcactcactccttggaaaggtattgaaatttaatttaactgactgact
ggggaacatgagagagaactgacttggatttgatgggaatcatattgatattgagtgctgcacgccacccagggaagttaaagcag
atataattgttcaactcttgaaattggctttcttccactcactcttctcctctctctctctctctctcctctctctctctctct
atggctcctcaatgaaacaaaatggagctaccgaagcaaggggaagcaactcaactcacaatgatggcaagaacataaaaatca
aaatgagctgtgactgagatttcaaccaagattggtaagatttggtaggaacattcgagctctcattgctcaaaaaggaagaagaat
ccagcaaggtgagattcttggaggtgactgtcttggatctctgctcaacccacaggaagtagctagagtgaggcaaggtgagga
gagattgaagcagatgaagcactcatcagatgaagcacaagggccgaaatccatctggagagcaagcgaaggtgagagactcgg
tgaagtagtagatcaagaaaaggattagaggttaattgcatgttgggtttgcaaggttatcctctctcctctctcctctctct
ttcttggagaagaattgagctgggttgatttttggtttcaactatggaggctcactcctctataaaaggggtgcaacagccatgact
gcaagaggttagaagttagcaggtgagagtgcaagccagaaagtctcagagctacctaaagcttacagcttttcttctctctta
ttagtttctctgttcaattttgtcttggagctctgaggtctgagaaaaagggcaacagtagggtttgtagaaaaagccatagag
gcagagagtgcaaaaataaaaagaaaagccatagatgctcttagagttccttggttcactcattattggtttctgagttctgt
tgtaagttgggttagcattacagttttgtaattctggttgattatagtagaattcattcattggttgtagggagactggatgtag
cacttagcagctgaaccaggatatactgggtgtaattttctctctcactcactcaatctggtttcttactgcaacagggacaa
gtctcgtgcaagtgacgagcaaaaatgaaaaggtctcgtgactaggacagctaaaaacagaaaagtctcctcaatccagcaac
caacaaaaaaagggggctaaatcaacccccctctcttagcactgaaacctcaattggtatcagagcttggtctcaagagat
cttctgagcttggagaaaatccagatggtagagaacaggtgcaaaaactgctcctcaaatcttactggaagcagctcaagca
cctctctcaatgggaaaaactatacctttagaaggaataatgagacttttagcaagcagtagactacagactgaggaacattctg
aagccctcaatttcaaaactgacacaaggtgtagaggttaactctctgaaaccagaaagtaggactgagaagatgagaagaag

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)