



Universidade de Brasília

Instituto de Ciências Biológicas

Departamento de Pós-Graduação em Fitopatologia

Prospecção e validação de genes-candidatos envolvidos na resistência de amendoim silvestre (*Arachis stenosperma*) ao *Passalora personata*, o agente causal da mancha preta do amendoim

Andressa da Cunha Quintana Martins

Orientador: Prof. Dr. Robert Neil Gerard Miller

Brasília – DF

2011



Universidade de Brasília

Instituto de Ciências Biológicas

Departamento de Pós-Graduação em Fitopatologia

Prospecção e validação de genes-candidatos envolvidos na resistência de amendoim silvestre (*Arachis stenosperma*) ao *Passalora personata*, o agente causal da mancha preta do amendoim

Andressa da Cunha Quintana Martins

Orientador: Prof. Dr. Robert Neil Gerard Miller

Dissertação apresentada ao Departamento de Fitopatologia do Instituto de Biologia, da Universidade de Brasília, como requisito parcial para obtenção do grau de mestre em Fitopatologia, Área de concentração: interação planta-patógeno.

Brasília – DF

2011

Dissertação de autoria de Andressa da Cunha Quintana Martins, intitulada “Prospecção e validação de genes-candidatos envolvidos na resistência de amendoim silvestre (*Arachis stenosperma*) ao *Passalora personata*, o agente causal da mancha preta do amendoim”, realizada junto ao Departamento de Fitopatologia, do Instituto de Ciências Biológicas da Universidade de Brasília e à Embrapa Recursos Genéticos e Biotecnologia, sob orientação do Professor Robert Neil Gerard Miller, com apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e da Fundação de Apoio à Pesquisa do Distrito Federal (FAP/DF).

Aprovado por:

Prof. Dr. Robert Neil Gerard Miller (Orientador)

Departamento de Fitopatologia e Departamento de Biologia Celular

Universidade de Brasília

Profa. Dra. Marisa Alvares da Silva Velloso Ferreira (Examinadora)

Departamento de Fitopatologia

Universidade de Brasília

Dra. Natália Florêncio Martins (Examinadora)

Laboratório de Bioinformática

Embrapa Recursos Genéticos e Biotecnologia

Prof. Dr. Cléber Furlanetto (Examinador - Suplente)

Departamento de Fitopatologia

Universidade de Brasília

Dedicatória

Á Deus, aos meus pais, Rose e Eivaldo, às minhas irmãs, Anna Paula e Larissa, à Silvana, ao meu namorado, Pedro, aos meus amigos, “orientadores” e professores, dedico.

AGRADECIMENTO

A Fundação de Apoio à Pesquisa do Distrito Federal (FAP/DF) pelo financiamento do projeto, possibilitando a realização do mesmo.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudos.

A Embrapa Recursos Genéticos e Biotecnologia pela oportunidade de execução do projeto proposto, bem como pela disponibilização da infra-estrutura e recursos operacionais necessários.

Ao meu orientador, professor Dr. Robert Neil Gerard Miller, pelos seus ensinamentos durante toda minha pós-graduação, seja dentro ou fora de sala de aula, ora como aluna, ora como educadora.

A Dra. Patrícia Messenberg Guimarães, pela confiança e disponibilização do projeto, pelos conselhos e ajudas, incentivo, carinho e compreensão durante todo o meu processo de amadurecimento científico. Obrigada por tudo! Afinal de contas, hoje sou fitopatologista por incentivo seu.

A Dra. Ana Cristina Miranda Brasileiro, por sempre estar disposta a me ajudar e, principalmente, a me ensinar. Obrigada por todo o aprendizado acumulado durante esses anos, pelos recém ensinamentos acerca de RT-qPCR e imensa ajuda na análise dos resultados.

A Dra. Soraya Cristina de Macedo Leal Bertoli, pelos ensinamentos fitopatológicos, extremamente úteis para que eu pudesse aprender a isolar, repicar e montar experimentos, com o fungo *Passalora personata*, essenciais para execução do meu trabalho.

Ao professor Dr. David John Bertoli, pelas sugestões e ajuda nas análises geradas pela “caliandra”, bem como outros ensinamentos que me auxiliaram no desenvolvimento deste trabalho.

A Dra. Ana Cláudia Guerra de Araújo, pelas risadas e ajuda nos meus incansáveis bioensaios.

A Dra. Carolina Vianna Morgante pelos ensinamentos repassados durante nosso trabalho juntas e pelas várias sugestões concedidas.

Aos professores, Dra. Marisa A. S. Velloso Ferreira e Dr. Cléber Furlanetto, pelos ensinamentos, incentivos e sugestões para a melhoria do meu projeto de trabalho, bem como pela disponibilidade em participar da minha banca.

A Dra. Natália Florêncio Martins, pela prontidão no aceite do convite para participar da minha banca.

A equipe do laboratório de bioinformática, em especial ao Dr. Orzenil Bonfim da Silva Junior, pela imensa ajuda nas análises de bioinformática.

Ao Leandro Mesquita, nosso maravilhoso guardião dos amendoins. Sem ele, nada do que há nas casas de vegetação vingariam. Seu cuidado minucioso e carinhoso com os *Arachis* sp. faz com que nossos trabalhos sejam viabilizados. Nunca vi ninguém com tanto trato e zelo. Adoro trabalhar com você!

A todos os professores do departamento da Fitopatologia e colaboradores (José Dianese, Connie McManus, Robert Miller, Carlos Uesugi, Marisa Ferreira, Adalberto Café, Cléber Furlanetto, Renato Resende, Alice Nagata, Rita de Cássia Carvalho, Cláudio Costa, Helson Martins, Juvenil Ceres, Luiz Blum e Carlos Inácio), os quais sem eles, o desenvolvimento do conhecimento fitopatológico seria impossível.

Aos meus queridos amigos, do laboratório de interação molecular planta-praga III (LPP III) da Embrapa Recursos Genéticos e Biotecnologia (Amanda Kristina, Thaís Nicolini, Alinne Silva, Rosana Blawdi, Lilian Travassos, Stephan Nielen, Bruna Vidigal, Larissa Muniz, Uiara Cavalcante, Iugo Galhardo, Lígia Almeida, Thatianny Alves, Rafaela Fontenele, Vânia Cristelly) que me ajudaram, apoiaram, incentivaram, riram e dividiram momentos muito agradáveis. Aos meus amigos da Fito (Ana Paula Pagani, Érica Santos, Maíra Andrade, Mariana Severo, Justino Neto, Daniel Lage, Liamar Anjos, Claudênia Ferreira, Cecília Rodrigues, Esdras Henrique, Jessica Monteiro, Kamila Araújo, Mikail Leastro, Mônica Alves, Natália Lucinda, Leonardo Braúna, Flávio Romero, Pablo Oliveira, Ribamar Frazão, Cezar Castro e os mais novos, os quais não me atreverei em citar, para não correr o risco de esquecer de alguém), pela companhia, amizade, ajuda e compartilhar do mesmo desespero durante as disciplinas cursadas no Programa de Pós-Graduação do departamento de Fitopatologia. Aos meus outros, mas não menos especiais, amigos, que me ajudaram de forma indireta, mas sempre com pensamentos positivos, e que compreenderam minha ausência e torceram pelo meu sucesso.

Ao meu querido namorado, Pedro Schwerz Junior, o qual é um companheiro inigualável, que me ajudou a descobrir que eu era muito mais capaz do que imaginava ser, que me consolou nas horas de desespero e dúvidas, que em hipótese nenhuma me

deixou desistir e que entendeu meus momentos “ausentes”. Amor, muito obrigada pelo seu imenso carinho, compreensão e incentivo. Amo você.

A minha família, que amo incondicionalmente, por todo amor, carinho e incentivo despendidos, mesmo nos momentos em que me ausentei ou me abduquei em participar por estar envolvida até o topo, com este trabalho. Sei que foram muitos os momentos de ausência. Mas acabou! Agora sou uma pessoa com vida social!

E, principalmente, a Deus, pois sem Ele a insanidade teria tomado conta de mim, meus pensamentos não seriam sensatos, minha paciência teria esgotado, minha persistência teria se perdido e o desespero se alastrado em mim. Obrigada meu Senhor, pelas bênçãos a mim concedidas.

É por isso, e por muito mais coisas, que sou imensamente grata a todos vocês!

SUMÁRIO

Resumo	x
Abstract	xi
Lista de tabelas.....	xii
Lista de figuras	xiii
1. Introdução	1
2. Referencial teórico.....	3
2.1. O amendoim (<i>Arachis hypogaea</i>)	3
2.2. O gênero <i>Arachis</i>	6
2.3. Uso de espécies silvestres no melhoramento genético.....	9
2.4. <i>Arachis stenosperma</i>	10
2.5. Principais fatores responsáveis pela redução da produtividade	11
2.6. Mancha preta.....	11
Nomenclatura	12
Etiologia	12
Sintomatologia.....	14
Epidemiologia.....	17
Controle.....	18
2.7. Interação planta patógeno	19
2.8. Análise da expressão gênica	22
2.9. Pirosequenciamento	24
2.10. PCR quantitativo (RT-qPCR)	26
3. Objetivos	31
3.1. Geral.....	31
3.2. Específicos	31
4. Material e Métodos	32
4.1. Bioensaio de <i>A.stenosperma</i> e <i>P. personata</i>	32
4.1.1. Plantio das sementes de <i>A. stenosperma</i> e <i>A. hypogaea</i> cv. IAC-Tatu.....	32
4.1.2. Isolamento e preparo do inóculo com <i>P. personata</i>	33
4.1.3. Inoculação das plantas com <i>P. personata</i> e coleta do material vegetal	34
4.2. Extração de RNA total	35
4.3. Análise qualitativa e quantitativa do RNA total e preparação das amostras para sequenciamento	36
4.4. Pirosequenciamento massal	38

4.5.	Análise bioinformática dos dados	42
4.6.	Avaliação dos unigenes para seleção de genes-candidatos	44
4.7.	Validação dos genes-candidatos por RT-qPCR.....	44
4.7.1.	Desenhos dos <i>primers</i> para os genes-candidatos selecionados.....	44
4.7.2.	Preparação das amostras para validação dos genes-alvo em RT-qPCR.....	45
	Tratamento das amostras com DNase e síntese de cDNA	45
4.7.3.	PCR quantitativa (RT-qPCR)	46
	Análise da expressão dos genes-candidatos selecionados	47
5.	Resultados	49
5.1.	Desenho experimental, inoculação e coleta das plantas.....	49
5.2.	Extração e purificação de RNA total	50
5.3.	Análise bioinformática dos dados gerados pelo pirosequenciamento massal (sequenciamento 454)	51
	Genes diferencialmente expressos in silico	56
5.4.	Escolha dos genes-candidatos	57
5.5.	Análise da qualidade do cDNA	60
5.6.	Validação dos genes-candidatos por RT-qPCR.....	60
6.	Discussão.....	68
7.	Conclusão	85
8.	Referências Bibliográficas	87
9.	Apêndices	111

RESUMO

O amendoim cultivado (*Arachis hypogaea*) é uma das oleaginosas mais importantes no mundo, pelo seu grande valor energético e nutricional. Contudo, o amendoim possui baixa variabilidade genética quando consideradas características de interesse agrônomo, o que difere de suas espécies silvestres, as quais são fontes de resistência a doenças e de adaptação a diversos ambientes. *Arachis stenosperma* é uma espécie silvestre resistente a algumas pragas e doenças, dentre elas, a mancha preta do amendoim, causada pelo fungo *Passalora personata*. Visando a identificação de genes-candidatos associados à resposta de defesa da planta, o transcriptoma de plantas de *A. stenosperma* acesso V10309, desafiadas com o fungo *P. personata*, e seu controle, foram analisados através de pirosequenciamento massal (Roche), resultando em 362.631 genes transcritos, o que gerou 17.912 unigenes, dos quais 7.723 eram constituídos por *contigs*. Dentre as anotações obtidas por BLASTx, 50 fatores de transcrição, 48 RGAs e vários genes-candidatos envolvidos em processos de defesa da planta foram identificados. Dentre eles, 16 genes-candidatos, cinco deles RGAs, foram avaliados através de RT-qPCR, quanto ao nível de expressão mediante a presença do patógeno, sendo cinco dos 11 primeiros, negativamente regulados, enquanto que todos os RGAs foram positivamente regulados. Esta é a primeira análise massal do transcriptoma de espécie silvestre de *Arachis* inoculado com *P. personata*. Os dados gerados neste trabalho promoverão o enriquecimento do banco de dados de genes transcritos de *A. stenosperma*, os quais poderão ser utilizados para descoberta de genes-candidatos à resistência, a outras características de interesse agrônomo e novos marcadores moleculares (SNPs e SSRs), que uma vez mapeados, poderão auxiliar na identificação de QTLs no melhoramento da cultura do amendoim, constituindo um importante recurso para a comunidade científica interessada na descoberta de genes relacionados à resposta de defesa ou ao estresse biótico em amendoim e em outras espécies.

Palavras-chave: Amendoim, *Arachis stenosperma*, *Passalora personata*, mancha preta, pirosequenciamento 454, RT-qPCR, genes-candidatos.

ABSTRACT

Peanut (*Arachis hypogaea*) is one of the most important oilseeds in the world due to its high energy value and nutrition. However, peanut has a low genetic variability, especially for agronomical traits, which differs from its wild species, which are sources of disease resistance and adaptation to various environments. *Arachis stenosperma* is a wild species resistant to some pests and diseases, among them, the Late Leaf Spot, caused by the fungus *Passalora personata*. Aiming to identify candidate genes associated with the plant defense response, the transcriptome of *A. stenosperma* V10309, challenged with the fungus *P. personata* and from a non-infected control, was analyzed by large scale pyrosequencing (Roche), resulting in 362,631 expressed genes, which generated 17,912 unigenes, which consisted of 7,723 contigs. Among the annotations obtained by BLASTx, 50 transcription factors, 48 RGAs and several candidate genes involved in plant defense processes were identified. Among them, 16 candidate genes were analyzed through RT-qPCR for expression in the presence of the pathogen. Five RGAs were up-regulated, whilst the other 11 were down-regulated. This is the first large scale report on the transcriptome analysis of wild peanut inoculated with *P. personata*. The data generated in this study promote the enrichment of the database of expressed genes from *A. stenosperma*, which can be used for discovery of candidate genes resistance, other agronomic traits of interest and new molecular markers (SNPs and SSRs). Once mapped, these can assist in the identification of QTLs in peanut crop improvement. This data is an important resource to the scientific community in the discovery of genes related to defense response or biotic stress in peanuts and other species.

Key-words: Peanut, *Arachis stenosperma*, *Passalora personata*, Late Leaf Spot, pyrosequencing 454, RT-qPCR, candidate genes.

LISTA DE TABELAS

- Tabela 1** – Resumo dos dados (em números) de leituras de sequências, unigenes e *contigs* obtidos pelo sequenciamento 454 GS FLX Titanium XLR70 e estimativa de cobertura genômica..... 53
- Tabela 2** – *Primers* desenhados para os genes candidatos selecionados em *A. stenosperma*. Valores de *e-value* obtidos pelo BLASTx..... 59
- Tabela 3** – Dados obtidos pela curva-padrão de diluição seriada dos pares de *primers* selecionados neste estudo. ID = Identidade do par de *primers*; Indicativo dos valores do *Slope*, R^2 (coeficiente de correlação) e média dos C_q das triplicatas na diluição 10^{-1} 62
- Tabela 4** – Cem primeiros unigenes (formados por dois ou mais *contigs*), diferencialmente expressos indicados pela subtração *in silico*, entre as bibliotecas de folhas de *A. stenosperma* controle (SFC) e inoculadas com *P. personata* (SFI). ID = identidade do *contig*; SFI/SFC = razão entre amostras inoculadas e controle, em vezes. Regulados negativamente (em azul); Regulados positivamente (em vermelho); Exclusivos nas amostras controle (SFC – em verde) e exclusivo nas amostras inoculadas (SFI – em roxo)..... 111

LISTA DE FIGURAS

- Figura 1** – Rendimento de produção (em milhões de toneladas/ano) das principais oleaginosas produzidas no mundo, nas safras de 2008 a 2011 (USDA-FAS, 2011a). 3
- Figura 2** – Países produtores de amendoim no mundo. Divisão (cores) por área colhida (ha). Imagem adaptada aos dados de USDA-FAS (2011b). 4
- Figura 3** – Produção de amendoim no Brasil (em milhões de toneladas). Imagem adaptada aos dados contidos no relatório de Acompanhamento da Safra Brasileira - Grãos - Safra 2010/2011 (Conab, 2011b). 5
- Figura 4** – Área de ocorrência das espécies silvestres da seção *Arachis* (Silva, 2008)..... 7
- Figura 5** – Esquema ilustrativo do cruzamento eventual de duas espécies diplóides (genoma AA + genoma BB) de *Arachis*, gerando um híbrido estéril (AB). 8
- Figura 6** – Imagens de microscopia óptica da fase anamórfica do agente causal da mancha preta, *Passalora personata*. (A) Estroma globular denso. (B) Conidióforos geniculados em fascículos densos. (C) Detalhe da estrutura fasciculada dos conidióforos e em seu aspecto liso. (D) Cicatrizes nos conidióforos (setas). (E) Crescimento simpodial das células conidiogênicas, formando conídios solitários (seta). (F) Mostra o padrão de formato encontrado de conídios, cilíndricos, podendo ser retos ou ligeiramente curvados, parede levemente áspera, podendo ser uni ou multisseptados, com ponta arredondada e base cônica com hilo conspícuo. Fotos de Liberato e Shivas (2006). 14
- Figura 7** – Ciclo de vida do *P. personata*, agente causal da mancha preta no amendoim. Imagem adaptada de McDonald e colaboradores (1985). 15
- Figura 8** – Folhas de *A. hypogaea* infectadas por *P. personata*. (A-G) *A. hypogaea* cv. IAC-Tatu, obtidos e cultivados em casa de vegetação da Embrapa/Cenargen, Brasília, DF. (A, C e E) Superfície adaxial; (B, D e F) Superfície abaxial. (A-D) Folhas destacadas, mantidas umedecidas em placa de Petri, mostrando sintomas iniciais da mancha preta (pequenas manchas cloróticas). Inoculadas artificialmente; (E-F) Cultivar crescida e mantida em copo plástico, em casa de vegetação. Inoculadas artificialmente; (C-D) Aumento do tamanho das lesões e a uniformização das bordas arredondadas. (G) Amostra coletada de casa de vegetação, contendo sintomas mais avançados da mancha preta e de ferrugem (*Puccinia arachidis*), ambas mostrando seus corpos de frutificação na parte abaxial. Infecção natural; H) Imagem de um campo de plantio de amendoim, no sul dos Estados Unidos, infectado por *P. personata* (DWSPL & Cattlin, s/d). Fotos (A-G) por Andressa da C. Q. Martins. DAI: Dias após a inoculação. 16
- Figura 9** – Países que possuem programas de melhoramento de amendoim no mundo. Adaptada de <http://www.guiageo-mapas.com/mapa-mundi.htm>. América do Norte: Estados Unidos (Texas, Flórida, Geórgia, Alabama, Carolina do Norte, Novo México, Oklahoma e Virgínia); América do Sul: Brasil (Distrito Federal¹, São Paulo², Paraíba³), Bolívia e Argentina; Europa:

- França; África: Senegal, Mali, Nigéria, Uganda, Tanzânia, Zâmbia, Malawi, Zimbábue, Moçambique e África do Sul; Ásia: Índia, China, Tailândia, Japão e Indonésia; Oceania: Austrália. ¹IAC – Instituto Agrônômico de Campinas – SP; ²Embrapa Recursos Genéticos e Biotecnologia – DF; ³Embrapa Algodão – Campina Grande – PB)..... 19
- Figura 10** – Esquema resumido do processo de pirosequenciamento. 24
- Figura 11** – Plotagem mostrando o número de ciclos da RT-qPCR e o número de cópias do material amplificado, em cada ciclo durante a RT-qPCR, além dos três estágios de amplificação: Exponencial (fase onde os reagentes da PCR estão em abundância, dobrando a cada ciclo os *amplicons* gerados); Linear (Reagentes começam a ficar escassos, diminuindo a quantidade de *amplicons* produzidos e desacelerando a reação de PCR) e Platô/Plateau (os reagentes já esgotaram e cessando a reação). Imagem adaptada de Life Technologies, sem data. 28
- Figura 12** – Curva de amplificação do RT-qPCR, número de ciclos da PCR x ΔRn (magnitude da fluorescência emitida), mostrando alguns componentes e valores que devem ser analisados. Seta e linha pontilhada verde: linha de base (*baseline*) há pouca mudança no sinal de fluorescência. Seta pontilhada azul: C_q (Ciclo de quantificação), número do ciclo no qual a fluorescência ultrapassa o *threshold*. Seta e reta vermelha: controle sem DNA (NTC – *No Template Control*) verifica a qualidade da amplificação e possíveis contaminações. Seta e reta lilás: limiar ou o *threshold*, nível de sinal do fluoróforo (corante *repórter* – corante de referência interna), corrige flutuações causadas por alterações na concentração ou volume. 28
- Figura 13** – Gráfico com os dados recolhidos durante o estágio de curva de dissociação (*melting curve*). Picos na curva de dissociação podem indicar a temperatura de dissociação (T_m – *melting temperature*) do alvo ou pode identificar amplificação inespecífica. O pico mais evidente localizado no centro do gráfico (seta azul) indica o ponto de dissociação do *amplicon* (amplificação de produto específico/esperado), com T_m a 77,36°C. Contaminantes, gerando produtos inespecíficos ou dímeros de *primers* aparecem como picos adicionais, vistos à esquerda do gráfico (seta vermelha). $-Rn$ corresponde a primeira derivada negativa da fluorescência normalizada pelo *reporter* (SYBR Green) durante a amplificação da PCR. 29
- Figura 14** – Sementes de *A. stenosperma* (V10309) e *A. hypogaea* cv. IAC-Tatu, germinadas em copos plásticos de 500 mL. Foto por Andressa da C. Q. Martins. 32
- Figura 15** – Divisão das plantas em dois grupos (controle e inoculado). Foto por Andressa da C. Q. Martins. 33
- Figura 16** – Esquema de inoculação, coleta e armazenamento das plantas de *A. stenosperma* e *A. hypogaea* cv. IAC-Tatu, durante os bioensaios. (A) Inoculação das plantas. (B) Coleta. (C) Identificação. (D) Armazenamento durante a coleta (N_2 líquido). (E) Acondicionamento final após a coleta (refrigerador -80°C). 35
- Figura 17** – Esquema ilustrativo da formação dos dois *pools* (A e B), por meio da junção equitativa do RNA total de diferentes plantas extraídas individualmente, do mesmo ponto de coleta (24, 48 ou 72h) e mesmo grupo (controle ou inoculado). 37

- Figura 18** – Esquema ilustrativo da formação de um *pool* controle e outro inoculado, contendo, de forma equitativa, as amostras dos *pools* A e B, de todos os pontos de coleta (24, 48 ou 72h), adicionados de etanol absoluto, para envio a empresa de sequenciamento massal (pirossequenciamento). SFC – *A. stenosperma* Folha Controle. 38
- Figura 19** – Esboço esquemático da síntese de cDNA SMART. Imagem adaptada (Evrogen, s/d). 39
- Figura 20** – Esquema ilustrativo das diferentes etapas de sequenciamento 454 (Ellegren, 2008):
 A) DNA fragmento utilizado como molde, com os adaptadores A e B ligados as suas extremidades 3' e 5', respectivamente. B) Fragmentos de DNA imobilizados por uma sequência alvo de biotina (adaptador B), que se liga às microesferas cobertas com estreptavidina. C) As microesferas são emulsionadas em água e óleo, contendo os reagentes necessários para reação de PCR, funcionando, como microrreator para amplificação clonal. Após a deposição dessas microesferas em placa de picotitulação, dá-se o início do pirossequenciamento, realizado por meio do fluxo seqüencial de reagentes sobre a placa. D) A intensidade do sinal captado por meio da incorporação de um novo nucleotídeo a nova fita, é gerado um sinal luminoso, representado em um pirograma. A altura de cada sinal é proporcional ao número de nucleotídeos idênticos incorporados consecutivamente. 41
- Figura 21** – Fluxograma utilizado para análise *in silico* dos dados gerados pelo sequenciamento 454..... 43
- Figura 22** – Esquema ilustrativo da formação de um único *pool* contendo amostras dos *pools* A e B de todos os pontos (24, 48 e 72h) e dos dois grupos trabalhados, controle e inoculado, para análise da curva de diluição pelo RT-qPCR. SFC – *A. stenosperma* Folha Controle; SFI – *A. stenosperma* Folha Inoculada. 47
- Figura 23** – Controles positivos (*A. hypogaea* cv. IAC-Tatu inoculado com *P. personata*). (A e B) Controle realizado no 6º bioensaio, com plantas mantidas durante todo o experimento até a aparição dos primeiros sintomas em casa de vegetação. (A) Sintomas registrados 31 dias após a inoculação. (B) Sintomas registrados 35 dias após inoculação. (C e D) Controle realizado no 8º bioensaio, com a técnica de folhas destacadas (Moraes & Salgado, 1982) acondicionadas em placas de Petri, mantidas em câmara de crescimento. (C) Sintomas registrados 24 dias após a inoculação dos folíolos em placa de Petri. (D) Sintomas registrados 27 dias após a inoculação. 49
- Figura 24** – Gel não desnaturante de agarose a 1,5%, corado com brometo de etídio, contendo RNA total de folha de *A. stenosperma* (1 µL de RNA total/poço). Amostras no gel: (M) Marcador molecular (1Kb Plus – Invitrogen); (1-2) SFC 24h; (3-4) SFI 24h; (5-6) SFC 48h; (7-8) SFI 48h e (9-10) SFI 72h. Onde, SFC = *A. stenosperma*/Folha/Controle, SFI = *A. stenosperma*/Folha/Inoculada e 24, 48 e 72h refere-se aos pontos de coleta..... 50
- Figura 25** – Gel não desnaturante de agarose a 1,5% corado com brometo de etídio, contendo amostras combinadas de RNA total de folha de *A. stenosperma* e passadas nas colunas do Kit Invisorb® Spin Plant Mini (Invitex). Amostras no gel: (M) Marcador de peso molecular (DNA do fago Lambda digerido com Hind III); (1) SFC 24h *Pool* A; (2) SFC 24h *Pool* B; (3)

- SFC 48h *Pool A*; (4) SFC 48h *Pool B*; (5) SFC 72h *Pool A*; (6) SFC 72h *Pool B*; (7) SFI 24h *Pool A*; (8) SFI 24h *Pool B*; (9) SFI 48h *Pool A*; (10) SFI 48h *Pool B*; (11) SFI 72h *Pool A*; (12) SFI 72h *Pool B*. Onde SFC = *A. stenosperma*/Folha/Controle, SFI = *A. stenosperma*/Folha/Inoculada e 24, 48 e 72h refere-se aos pontos de coleta..... 51
- Figura 26** – Frequência dos comprimentos de leituras (pb) obtidos por meio do sequenciamento 454. À esquerda, em vermelho, dados referentes à biblioteca de folhas de *A. stenosperma* inoculadas com *P. personata*. À direita, em verde, dados referentes à biblioteca de folhas *A. stenosperma*, não submetidas à inoculação (controle). 52
- Figura 27** – Distribuição dos *contigs* por número de genes transcritos. 53
- Figura 28** – Distribuição dos *contigs* por comprimento. 54
- Figura 29** – Análises de InterProScan e de anotação dos termos de ontologia gênica (GO) dos *contigs* de *A. stenosperma*. (A) Número de termos de processo biológico; (B) Número de termos de componentes celulares; (C) Números de termos de função molecular. 55
- Figura 30** – Distribuição de *contigs* por famílias de fatores de transcrição em *A. stenosperma*. ... 56
- Figura 31** – Agrupamentos gerados no programa BioLayout 3D, baseado na similaridade dos unigenes das bibliotecas SFI e SFC com RGAs de *A. thaliana*, além de similaridades diferenciais de alguns grupos..... 58
- Figura 32** – Gel não desnaturante de agarose a 1,5% corado com brometo de etídio. (A) Produto da RT-PCR onde foram avaliadas as primeiras amostras de cDNA sintetizadas; (1-4) Produto que teve como molde cDNA (SFI 24h *Pool A*, SFI 24h *Pool B*; SFI 48h *Pool A* e SFI 72h *Pool A*; (5) Produto que teve como molde DNA genômico (*A. hypogaea* cv. IAC-Tatu – controle positivo) e (6) Controle negativo da reação, sem cDNA molde. (B) Produto da RT-PCR dos cDNAs sintetizados de todas as amostras. (1) SFC 24h *Pool A*; (2) SFC 24h *Pool B*; (3) SFI 24h *Pool A*; (4) SFI 24h *Pool B*; (5) SFC 48h *Pool A*; (6) SFC 48h *Pool B*; (7) SFI 48h *Pool A*; (8) SFI 48h *Pool B*; (9) SFC 72h *Pool A*; (10) SFC 72h *Pool B*; (11) SFI 72h *Pool A* e (12) SFI 72h *Pool B*; (13) Controle negativo (sem cDNA molde) e (14) Controle positivo (cDNA de *A. stenosperma*). Marcador de peso molecular (M) utilizado foi 1Kb Plus (Invitrogen) SFC = *A. stenosperma*/Folha/Controle, SFI = *A. stenosperma*/Folha/Inoculada, onde e 24, 48 e 72h referem-se aos pontos de coleta..... 60
- Figura 33** – Exemplo de uma curva padrão de diluição seriada. Dados obtidos pela amplificação dos *primers* RAPS, utilizando-se como molde o cDNA dos *pools* de folhas controle e inoculadas de *A. stenosperma*, diluído a 10^{-1} , 10^{-2} e 10^{-3} 61
- Figura 34** – Curva de dissociação obtida para cada par de *primers* selecionados neste estudo, mostrando a formação de um único pico, referente à amplificação de um único fragmento específico. (A) *Actina*; (B) *60S*; (C) *PDI*; (D) *P450*; (E) *RAPS*; (F) *HD-Zip*; (G) *Proteassoma*; (H) *LRR-TM-PK*; (I) *Bomba de Ca²⁺*; (J) *ARF*; (K) *NPL*; (L) *DNA-RNA pol*; (M) *BURP*; (N) *EDR1*; (O) *GSK*; (P) *CIPKs*; (Q) *STPKs* e (R) *CRKs*. T_m = Temperatura de dissociação (*melting temperature*). OBS.: O pico à esquerda observado em todas as imagens refere-se a um problema gerado pela calibração do equipamento F7300 Real-Time PCR System, da Applied Biosystem. 63

- Figura 35** – Análise da quantificação relativa (RQ) por RT-qPCR para os 11 primeiros genes-candidatos, a partir de dois *pools* de *A. stenosperma* (SFC e SFI) do 6º bioensaio. Os níveis relativos de mRNA (razão entre SFI/SFC) estão representados por barras e o desvio-padrão das triplicatas pelas linhas verticais. (A) Dados normalizados com o gene de referência *Actina*. (B) Dados normalizados com o gene de referência *60S*. 65
- Figura 36** – Análise da RQ por RT-qPCR para os genes-alvo *P450*; *HD-Zip*; *LRR-TM-PK*; *NPL* e *BURP*. Os níveis relativos de mRNA (razão SFI/SFC) estão representados por barras e o desvio-padrão das médias entre os mesmos pontos de coleta (24, 48 e 72h) do 4º, 5º e 8º bioensaios pelas linhas verticais. 1
- Figura 38** – Análise da quantificação relativa (RQ) para os cinco RGAs-candidatos, a partir de dois *pools* de *A. stenosperma* (SFC e SFI) do 6º bioensaio. Os níveis relativos de mRNA (razão entre SFI/SFC) estão representados por barras e o desvio-padrão das triplicatas pelas linha verticais. Dados normalizados com o gene referência *60S*. 67

1. INTRODUÇÃO

O amendoim (*Arachis hypogaea* L.) é uma cultura alimentar importante, cultivada em regiões tropicais, subtropicais e temperadas, sendo a quarta oleaginosa mais produzida no mundo. Apesar da grande variabilidade morfológica apresentada entre as diferentes cultivares de *A. hypogaea* (alotetraploide AABB), esta espécie apresenta estreita base genética e alta suscetibilidade ao ataque de fitopatógenos. Por outro lado, seus parentes silvestres diploides apresentam maior variabilidade genética, sendo uma potencial fonte de alelos de resistência a doenças e à adaptação a diversos ambientes, como é o caso da espécie silvestre *Arachis stenosperma* (acesso V10309). Muitas são as doenças que acometem e prejudicam a produtividade do amendoim. As doenças foliares são as que causam as maiores perdas. Dentre elas, duas são consideradas as doenças foliares mais graves em todo o mundo: a mancha preta (*Passalora personata*) e a ferrugem (*Puccinia arachidis*).

Apesar das plantas estarem sob constante exposição a vários fitopatógenos, elas são naturalmente resistentes a muitos deles. No entanto, planta e patógeno, estão em processo contínuo de co-evolução, resultando em complexos mecanismos de defesa da planta, para reconhecer e se proteger da invasão do mesmo. Diante disso, a identificação de genes envolvidos em respostas de defesa da planta é essencial para o estabelecimento de estratégias em programas de melhoramento de plantas. Neste intuito, o estudo da expressão gênica tem demonstrado ser uma importante ferramenta para o entendimento dos processos biológicos envolvidos na defesa de plantas, na caracterização de resistências e na identificação de genes que possam estar envolvidos nestes mecanismos de defesa. Para tal, a construção de banco de dados de ESTs (*Expressed Sequence Tags*) ou genes transcritos (genes sequenciados sem a necessidade de clonagem) é de grande valia no estudo de expressão gênica, principalmente para espécies em que a sequência genômica completa não está disponível, como é o caso do amendoim.

O uso de tecnologias de sequenciamento de segunda geração, como a plataforma 454 GS-FLX, tem-se mostrado muito útil no sequenciamento em larga escala de genes transcritos, representando um método rápido e eficiente de grande potencial para analisar regiões transcritas do genoma. Essa plataforma permite a montagem *de novo* de um genoma, pois produz sequências de 250-300 pb (atualmente, são de 500-700 pb) facilitando a montagem das sequências consenso. Esta tecnologia também possibilita a

análise de expressão diferencial entre amostras submetidas a tratamentos contrastantes, possibilitando a identificação de sequências expressas envolvidas em resposta de defesa ou ao estresse biótico.

Após a identificação de genes-candidatos, uma importante etapa é a sua validação por meio da análise mais detalhada de seu perfil de expressão. Esse processo foi facilitado pela técnica de RT-qPCR (Reação em cadeia da polimerase com transcrição reversa quantitativa), que combina a amplificação de um alvo específico com métodos de quantificação por meio de medidas da fluorescência associada à síntese de *amplicon* ao longo dos ciclos da PCR.

O presente trabalho teve como objetivo, descobrir genes associados à resposta de defesa da planta durante interação planta-patógeno, por meio da análise massal do transcrito da espécie silvestre de *Arachis* (*A. stenosperma*) resistente ao fungo causador da mancha preta do amendoim, *P. personata*, bem como o desenvolvimento de um banco de dados dos genes transcritos durante esta interação.

2. REFERENCIAL TEÓRICO

2.1. O amendoim (*Arachis hypogaea*)

O amendoim (*Arachis hypogaea* L.) é uma das culturas mais populares do mundo (Kumar, 2007), sendo cultivado extensivamente em regiões tropicais, subtropicais e temperadas (Chirinos, 2011), em mais de 100 países, nos seis continentes (Kumar, 2007). Depois da soja (*Glycine max* L. Merr.; 257,47 milhões de toneladas/ano – “Mt/ano”), canola (*Brassica napus* L.; 58,52 Mt/ano) e algodão (*Gossypium hirsutum* L.; 46,2 Mt/ano), o amendoim é a quarta oleaginosa mais cultivada no mundo, com uma superfície de área plantada superior a 21,2 milhões de hectares e produção global acima de 35,2 milhões de toneladas/ano (Figura 1) (USDA-FAS, 2011a). Apesar de ser originalmente brasileira, esta cultura assim como tantas outras, tornou-se mais importante em regiões do mundo longe do seu país de origem (Bertioli, et al., 2011). A grande maioria do amendoim é produzida na Ásia e na África: 70% (24,19 Mt/ano) e 21% (7,29 Mt/ano) da produção global, respectivamente (USDA-FAS, 2011b). Os outros 9% (3,3 Mt/ano) vem da América do Norte, Caribe, América do Sul, Europa e Oceania (Dwivedi, et al., 2003; USDA-FAS, 2011b).

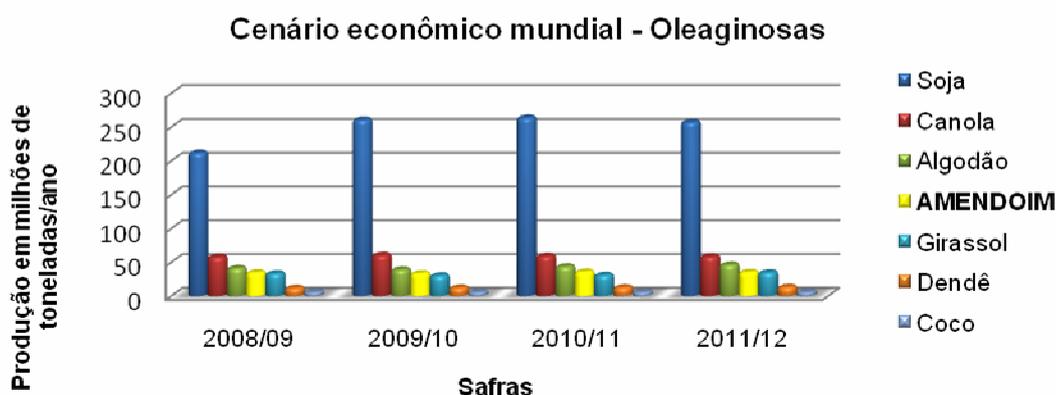


Figura 1 – Rendimento de produção (em milhões de toneladas/ano) das principais oleaginosas produzidas no mundo, nas safras de 2008 a 2011 (USDA-FAS, 2011a).

Os países em desenvolvimento são responsáveis por mais de 85% da produção mundial de amendoim (Figura 2). Dentre os cinco maiores produtores, a China ocupa a primeira posição (15,2 milhões de toneladas – Mt), seguida pela Índia (6 Mt), Estados Unidos (1,63 Mt), Nigéria (1,55 Mt) e Indonésia (1,25 Mt) (USDA-FAS, 2011b). O Brasil já esteve entre os sete primeiros produtores, sendo que o principal produto comercializado

na época, década de 70, era o óleo (Santos, 1999). Contudo vários fatores tecnológicos e mercadológicos, tais como: custos crescentes de produção da cultura do amendoim acompanhados por baixo rendimento por área, suscetibilidade às variações climáticas com influências negativas na qualidade do produto e intensas variações de preço durante a comercialização (Nogueira Júnior, 1976; Freitas e Margarido, 2003) e expansão das culturas de soja, cana de açúcar e laranja (Caser, et al., 1994), influenciaram a redução do cultivo nacional que hoje ocupa a 18ª posição (USDA-FAS, 2011b), com uma estimativa para as safras de 2010/2011 de 84,4 mil hectares (ha) de área plantada e 224,1 mil toneladas de produção de grãos. Com esta produção, de janeiro a junho deste ano, o Brasil exportou 21.974 toneladas de amendoim em grãos, além de 13.373 toneladas de óleo bruto e 218 toneladas de óleo refinado, perfazendo uma receita de quase US\$50 milhões (Conab, 2011a). O estado de São Paulo é o principal produtor brasileiro de amendoim, respondendo por 80% da produção nacional (Figura 3), e cerca de 70% do total da área plantada no Brasil (Conab, 2011b).

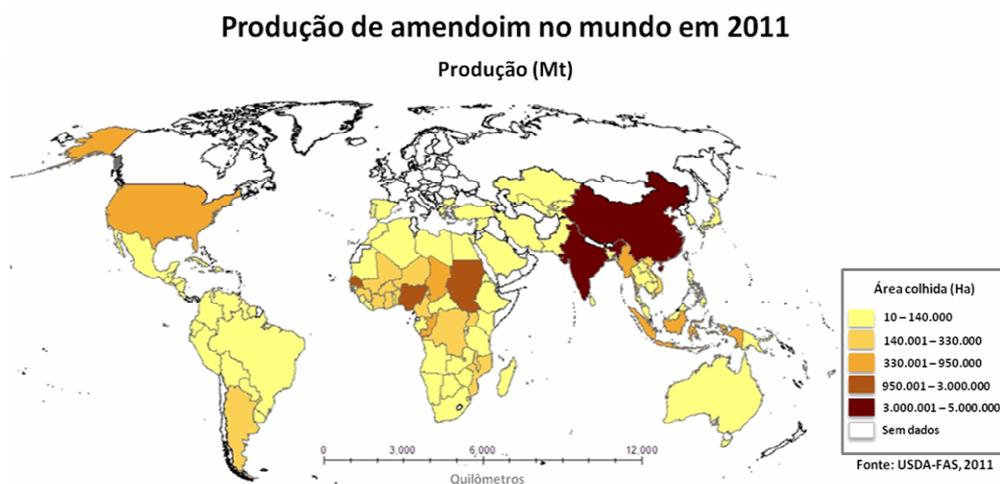


Figura 2 – Países produtores de amendoim no mundo. Divisão (cores) por área colhida (ha). Imagem adaptada aos dados de USDA-FAS (2011b).

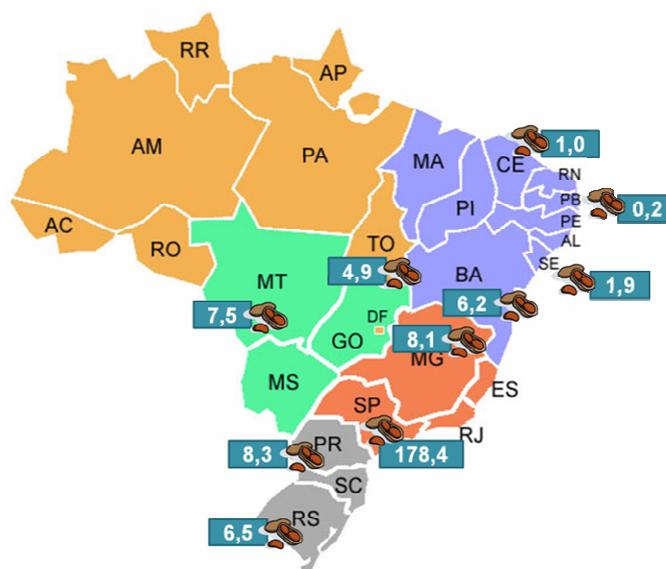


Figura 3 – Produção de amendoim no Brasil (em milhões de toneladas). Imagem adaptada aos dados contidos no relatório de Acompanhamento da Safra Brasileira - Grãos - Safra 2010/2011 (Conab, 2011b).

Os grãos de amendoim são uma excelente fonte de proteína vegetal, pois contém 45-50% de óleo, 27-33% de proteína, assim como vitaminas e minerais essenciais (Pretorius, 2006). O óleo contido nesses grãos é composto por uma mistura de acilgliceróis, contendo uma alta proporção de ácidos graxos insaturados, em particular, o oléico (50-65%) e linoléico (18-30%) (Pattee, 2005). Desta forma, o grão inteiro do amendoim é comestível e muito apreciado como alimento humano e nos produtos processados (Chirinos, 2011). Um exemplo são os salgadinhos e doces de amendoim. Além disso, derivados de amendoim, como farinha, farelo e isolados de proteína, são úteis como suplementos de proteínas nas produções de alimentos enriquecidos que podem ajudar na alimentação, especialmente, em países em desenvolvimento, onde a proteína animal é muito cara para a maioria da população (Singh e Singh, 1991). De acordo com Howard Valentine (2011), a cada ano, no mundo, 5,1 milhões de crianças menores de cinco anos morrem de desnutrição, e para reverter tal estatística, um novo produto, a base de amendoim, vem sendo usado. Fora os destinos alimentícios acima citados e a reserva de óleo, o resíduo vegetativo (*haulms*) pode ser usado para alimentar o gado (Chirinos, 2011), principalmente na época seca onde há certa escassez de forragem disponível para alimentação animal.

2.2. O gênero *Arachis*

O gênero *Arachis* pertence à família das leguminosas (Fabaceae ou Leguminosae) e à subfamília Papilionideae (Bertioli, et al., 2011), sendo todas as espécies deste gênero geocárpicas (Krapovickas & Gregory, 1994).

A maioria das espécies pertencentes a este gênero é autofecundada (autógamas), mas a fecundação cruzada pode ocorrer por polinização de abelhas (Luo, et al., 2005a).

Exclusivo da América do Sul, o gênero *Arachis* está distribuído dentro de uma grande região deste continente. Apresenta distribuição natural restrita ao Brasil, Bolívia, Argentina, Paraguai e Uruguai (Silva, 2008). E das 64 espécies silvestres encontradas no Brasil, 47 são exclusivas de sua flora, incluindo espécies de interesse econômico (Valls e Simpson, 2005; Silva, 2008).

Com base na morfologia, compatibilidade de cruzamento, viabilidade de híbridos, distribuição geográfica e citogenética, as 80 espécies já descritas do gênero *Arachis* estão dispostas em nove seções taxonômicas: *Trirectoides*, *Erectoides*, *Procumbentes*, *Rhizomatosae*, *Heteranthae*, *Caulorrhizae*, *Extrabervosae*, *Triseminatae* e *Arachis* (Krapovickas e Gregory, 1994; Valls e Simpson, 2005; Bertioli, et al., 2011).

A seção *Arachis* abriga 32 das 80 espécies descritas, inclusive *A. hypogaea*, e é considerada a mais diversificada e derivada. Abriga tanto espécies anuais quanto perenes, e espécies com números diferentes de cromossomos, diferentes níveis de ploidia e estrutura de cariótipos. Além disso, possui vasta distribuição geográfica (Figura 4) com espécies exclusivas do Brasil e de interesse para o melhoramento genético de *A. hypogaea* (Silva, 2008; Bertioli, et al., 2011).



Figura 4 – Área de ocorrência das espécies silvestres da seção *Arachis* (Silva, 2008).

Assim como no gênero *Arachis*, a maioria das espécies da seção *Arachis* é diploide ($2n=2x=20$) e somente duas espécies são tetraploides ($2n=4x=40$ cromossomos), *A. monticola* e *A. hypogaea* (Krapovickas e Gregory, 1994). Espécies diploides com $2n=20$ são atribuídas a três genomas diferentes, A, B e D. As espécies, com o genoma A, são caracterizadas por um pequeno par simétrico de cromossomos, enquanto que nas espécies com o genoma B os pares de cromossomos são maiores que os encontrados no A. Já o genoma D possui cariótipo assimétrico (Smartt, 1964, citado por Smartt, et al., 1978; Singh & Moss, 1982; Stalker, 1991). As espécies tetraploides têm fórmula genômica AABB, isto é, quatro conjuntos de cromossomos, dois do genoma A e dois do genoma B (Krapovickas & Gregory, 1994) e surgiram, provavelmente, de um cruzamento eventual entre duas espécies diploides, *A. duranensis* (genoma AA) e *A. ipaënsis* (genoma BB). Este teria resultado em um híbrido estéril (AB), cujos cromossomos foram duplicados, através de um evento natural, levando a restauração da fertilidade (Figura 5) (Kochert, et al., 1996; Seijo G. J., et al., 2004).

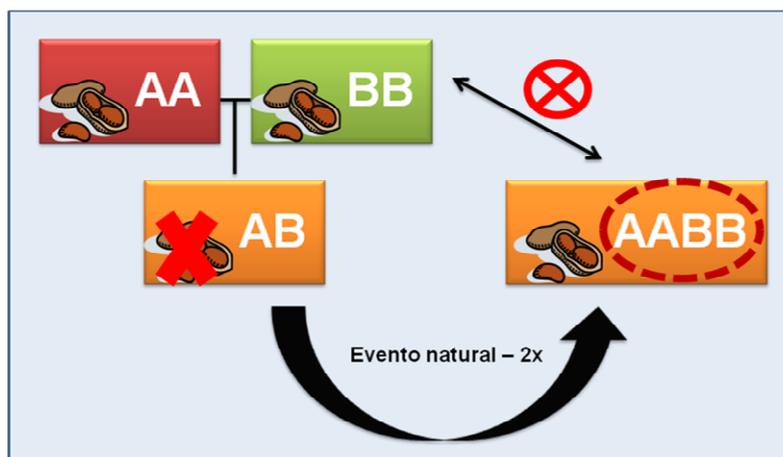


Figura 5 – Esquema ilustrativo do cruzamento eventual de duas espécies diploides (genoma AA + genoma BB) de *Arachis*, gerando um híbrido estéril (AB).

A provável origem de *A. hypogaea*, por meio deste evento de poliploidização seguido por sucessivas seleções durante sua domesticação, resultou em um genoma tetraploide conservado (Young, et al., 1996).

A poliploidia de *A. hypogaea* levou a um isolamento reprodutivo, restringindo a partilha de genes com seus parentes silvestres (Figura 5), criando, dessa forma, um grande gargalo genético (Kochert, et al., 1996; Bertoli, et al., 2011).

Apesar de uma notável quantidade de variabilidade morfológica acumulada durante a seleção artificial ao longo de milhares de anos de domesticação, o isolamento reprodutivo pode ser, possivelmente, a principal causa de sua base genética estreita, que para outras características como resistência a doenças e pragas apresenta claras limitações ao melhoramento da cultura (Kochert, et al., 1991; Bertoli, et al., 2011).

Uma forma de introgressão de genes de parentes silvestres na espécie cultivada seria por meio de duplicação cromossômica artificial, onde os híbridos estéreis gerados pelo cruzamento entre espécies silvestres com genomas AA e BB são submetidos ao tratamento com colchicina para duplicar o número de cromossomos, obtendo-se assim anfiploides sintéticos (AABB), que a princípio são sexualmente compatíveis com o amendoim (Simpson, 1991).

2.3. Uso de espécies silvestres no melhoramento genético

O uso de espécies silvestres no melhoramento de plantas torna-se importante, quando a resistência a vários estresses bióticos e abióticos em germoplasma cultivado é baixa, ou quando a variabilidade genética é estreita e pressões seletivas resultam em novos biótipos de pragas e de fitopatógenos virulentos. Neste caso, a descoberta e a incorporação de genes adicionais para a resistência de espécies silvestres tornam-se fundamentais para sustentar a produtividade das culturas (Rao, et al., 2003). Isto porque ao contrário da espécie cultivada, durante a evolução as espécies silvestres têm se adaptado a diversos nichos ecológicos, ou seja, a uma variedade de ambientes e estresses bióticos, constituindo uma rica fonte de diversidade de alelos com potencial para uso no melhoramento da cultura do amendoim (Guimarães, et al., 2008; Bertoli, et al., 2011).

Não surpreendentemente, a maior contribuição das espécies silvestres tem sido no melhoramento para resistência a doenças (Rao, et al., 2003). Diversos trabalhos têm mostrado que as espécies silvestres de *Arachis* são muito mais resistentes a doenças causadas por fungos, vírus e nematoides do que a espécie cultivada (Stalker & Moss, 1987; Holbrook, et al., 2000; Pande & Rao, 2001; Rao, et al., 2003).

O Brasil é o centro de origem de grande parte das espécies silvestres de *Arachis*. Muitas destas espécies silvestres estão conservadas no banco de germoplasma da Embrapa Recursos Genéticos e Biotecnologia (Embrapa/Cenargen), o maior de *Arachis* silvestre do mundo, podendo ser explorado visando o melhoramento genético do amendoim (Simpson, et al., 2001).

Assim, é necessária uma maior e mais detalhada investigação desse material coletado e conservado, sabendo que o grande interesse pela prospecção, resgate e caracterização de germoplasma dessas espécies silvestres reside, justamente, na utilização dessas informações genéticas para o melhoramento do amendoim cultivado (Stalker & Moss, 1987; Stalker, 1992; Simpson, et al., 2001). Com isso, o gênero *Arachis* vem sendo estudado com crescente intensidade (Fávero & Suassuna, 2007), visando à identificação e o isolamento de novos alelos de seus parentes silvestres com potencial interesse agrônomo.

Além do melhoramento clássico, existe a possibilidade de por meio da engenharia genética se transferir os genes de interesse evitando assim o arraste de ligação deletéria, que muitas vezes se torna um problema durante a transferência de genes de espécies silvestres através de técnicas de melhoramento convencional (Kumar & Kirti, 2011).

2.4. *Arachis stenosperma*

Arachis stenosperma Krapov. e W. C. Gregory é uma espécie silvestre diploide do gênero *Arachis* (seção *Arachis*), reconhecida como uma fonte de genes importantes para o melhoramento genético do amendoim (*A. hypogaea* L.), por causa da estreita relação entre as duas espécies (Pacheco, et al., 2008) e por apresentar resistência a inúmeras doenças, como por exemplo, mancha preta, mancha marrom, nematoide das galhas e nematoide dos bulbos (Proite, et al., 2007).

Sabe-se que *A. stenosperma* é de ciclo perene e possui 20 cromossomos, incluindo um par denominado par "A" (Fernández & Krapovickas, 1994). É sexualmente compatível com o mais provável doador do genoma AA (*A. duranensis*) do amendoim cultivado, e, portanto, é um excelente candidato a doador do genoma para introgressão de genes (Proite, et al., 2007), além de ser cultivada por índios guaranis no litoral Atlântico (Silva, 2008), mostrando ser objeto de seleção para características de cultivo.

Esta espécie é geograficamente a mais amplamente difundida entre as espécies de *Arachis* e, em consequência, foi selecionada em diversos ambientes que vão desde a savana às dunas costeiras (Proite, et al., 2007). Devido a isso, diversos acessos de *A. stenosperma* do Banco Ativo de Germoplasma (BAG) da Embrapa Recursos Genéticos e Biotecnologia já foram testados. Mais especificamente, o acesso V10309 desta espécie, coletado em Rondonópolis – MT, é resistente ao nematoide das galhas (*Meloidogyne arenaria* raças 1 e 2, *M. javanica* raça 4 e *M. hapla*), ao nematoide dos bulbos (*Ditylenchus destructor*) e aos fungos foliares [*Cercospora arachidicola*, *Passalora personata* (Sin. *Cercosporidium personatum*) e *Puccinia arachidis*] (Veiga, et al., 2001; Leal-Bertioli, et al., 2005; Proite, et al., 2007; Fávero, et al., 2009; Leal-Bertioli, et al., 2010).

Segundo Leal-Bertioli e colaboradores (2010), 35 dias após a inoculação, as amostras de *A. hypogaea* cv. IAC-Tatu exibiam sintomas típicos visíveis causados por *P. personata*, contudo, a espécie silvestre testada, *A. stenosperma* acesso V10309, não apresentou nenhum sintoma macroscópico, o que também foi confirmado por Fávero e colaboradores (2009) após 42 dias da primeira inoculação.

2.5. Principais fatores responsáveis pela redução da produtividade

O amendoim é amplamente cultivado nos trópicos semiáridos por agricultores detentores de poucos recursos, onde muitos fatores abióticos e bióticos podem limitar sua produtividade e a qualidade das sementes (Dwivedi, et al., 2003).

A maioria das perdas econômicas desta cultura, em todo o mundo, é imposta por doenças foliares causadas por diversos agentes patogênicos (Nobile, et al., 2008), dentre eles: *Phoma arachidicola* Marasas, Pauer & Boerema (agente causal da mancha barrenta), *Cercospora arachidicola* Hori (mancha castanha), *Sphaceloma arachidis* Bitanc. & Jenkis (verrugose), *Passalora personata* (Berk. & M.A. Curtis) S.A Khan & M. Khamal (mancha preta) e *Puccinia arachidis* Speg. (ferrugem) (Godoy, et al., 1999; Izge, et al., 2007; Leal-Bertioli, et al., 2010). As duas últimas, mancha preta e ferrugem, são consideradas as doenças foliares mais graves no amendoim no Brasil e no mundo (Shokes & Culbreath, 1997; Pande & Rao, 2001; Leal-Bertioli, et al., 2005; Nobile, et al., 2008). Na maioria das áreas ambas as doenças ocorrem juntas, mas a incidência e gravidade de cada doença variam com o ambiente, localização e cultivar (Pande & Rao, 2001; Cantonwine, et al., 2008). Ambas as doenças causam perdas de rendimento de 50% ou mais na safra se cultivares suscetíveis não forem protegidas com produtos químicos (Subrahmanyam, et al., 1984; Aquino, et al., 1995; Pande & Rao, 2001), e além de prejudicar o rendimento e qualidade da vagem, essas doenças também afetam o rendimento e qualidade da área foliar (*haulm*), que poderia ser usada para forragem (Khedikar, et al., 2010).

Portanto, cultivares bem adaptadas com múltipla resistência ao estresse biótico e tolerância aos estresses abióticos (seca, baixa disponibilidade de ferro, especialmente sob solos em condições ácidas ou a não disponibilidade dele em solos calcários, por exemplo) e com alto rendimento, proporcionaria uma maior e mais sustentável produção de amendoim, principalmente aos agricultores de subsistência nas regiões dos trópicos semiáridos, reduzindo os custos de produção e os efeitos nocivos de produtos químicos no ecossistema (Dwivedi, et al., 2003; Izge, et al., 2007).

2.6. Mancha preta

A mancha preta, também conhecida como mancha tardia, pinta preta, *viruela del maní* ou *late leaf spot*, é causada pelo fungo *Passalora personata*. Esta espécie é exclusivamente patogênica ao gênero *Arachis* e causa a mais severa entre as doenças

foliares do amendoim no estado de São Paulo, o maior produtor nacional, além de ser mundialmente difundido e aparecer em destaque também em outros países produtores (Moraes, et al., 1994; Butler, et al., 1994; Nobile, et al., 2008). Dificilmente encontra-se uma cultura em fim de ciclo sem sintomas dessa doença, pois a mancha preta geralmente ocorre mais tarde na estação e é, muitas vezes, vista como um complexo com outras manchas foliares (Pretorius, 2006).

Nomenclatura

Há ainda certa controvérsia no que se refere à nomenclatura correta da fase anamórfica do agente causal da mancha preta.

De acordo com o Index Fungorum (2011), a nomenclatura mais recente válida e disponibilizada para este patógeno é *Phaeoisariopsis personata* (Berk. & M.A. Curtis) Arx (1983). Outras sinonímias são encontradas, tais como: *Passalora personata* (Berk. & M.A. Curtis) S.A Khan & M. Khamal (1961) e *Cercosporidium personatum* (Berk. & M.A. Curtis) Deighton (1967).

O gênero *Cercosporidium* tornou-se sinonímia de *Passalora*, que é essencialmente distinta de *Phaeoisariopsis* por não ter conidiomas sinemáticos e por apresentar, visivelmente, células conidiogênicas geniculadas (Stewart, et al., 1999).

Segundo alguns pesquisadores (Braun, 1995; Stewart, et al., 1999; Crous, Kang, & Braun, 2001) e com base em seus estudos morfológicos e moleculares, optou-se neste trabalho, utilizar o nome científico *Passalora personata* (Berk. & M.A. Curtis) S.A Khan & M. Khamal (1961), por se tratar de ser o mais apropriado.

Etiologia

Mycosphaerella berkeleyi W.A. Jenkins (1938), é a forma teleomórfica ou sexuada de *Passalora personata*. É um ascomiceto, pertencente à família Mycosphaerellaceae, apresenta ascomas periteciais espalhados, principalmente ao longo das margens da lesão, anfigeno, parcialmente dentro do tecido do hospedeiro, erupente, amplamente ovado a globoso, em um tamanho de 84-140 x 70-112µm, na cor preta, ostíolo levemente papilado, ascos cilíndricos, em forma de trevos, estipes pequenas, fasciculadas, 30-40 x 4-6µm, aparafisados, bitunicados, com oito (8) esporos, esporos unisseriados a imperfeitamente bisseriados no asco, bicelular, a célula superior um pouco maior,

ligeiramente constricto no septo, hialino e ascósporos com 10,9-19,6 x 2,9-3,8 μ m (média de 14,9 x 3,4 μ m) de tamanho (McDonald, et al., 1985).

O agente causal da mancha preta, apesar da existência conhecida da sua fase teleomórfica, é encontrado em sua fase anamórfica (assexual), conhecida como *P. personata* (Berk. & M.A. Curtis) S.A Khan & M. Khamal (1961), um fungo cercosporóide, que possui estroma globular denso, pseudoparenquimatoso, com até 130 μ m de diâmetro (Figura 6A), além de inúmeros conidióforos de cor pálida a marrom oliváceo, lisos, 1-3 genículos, com tamanhos de 1-10 x 3,0-6,5 μ m, cicatrizes conidiais conspícuas, proeminentes, um tanto escuras, com 2-3 μ m de largura (Figura 6B-D), os conídios são solitários e pigmentados (castanho claro), cilíndricos, obclavados, geralmente retos ou ligeiramente curvados, parede geralmente finamente áspera, arredondadas no ápice, base pouco cônica com hilo conspícuo, 1-9 septos não constrictos, principalmente com 3-4 septos e variando de 20-70 x 4-9 μ m de tamanho (Figura 6E-F) (Jenkis, 1938; McDonald, et al., 1985; Pretorius, 2006; Freire & Braun, 2009; Leal-Bertioli, et al., 2010).

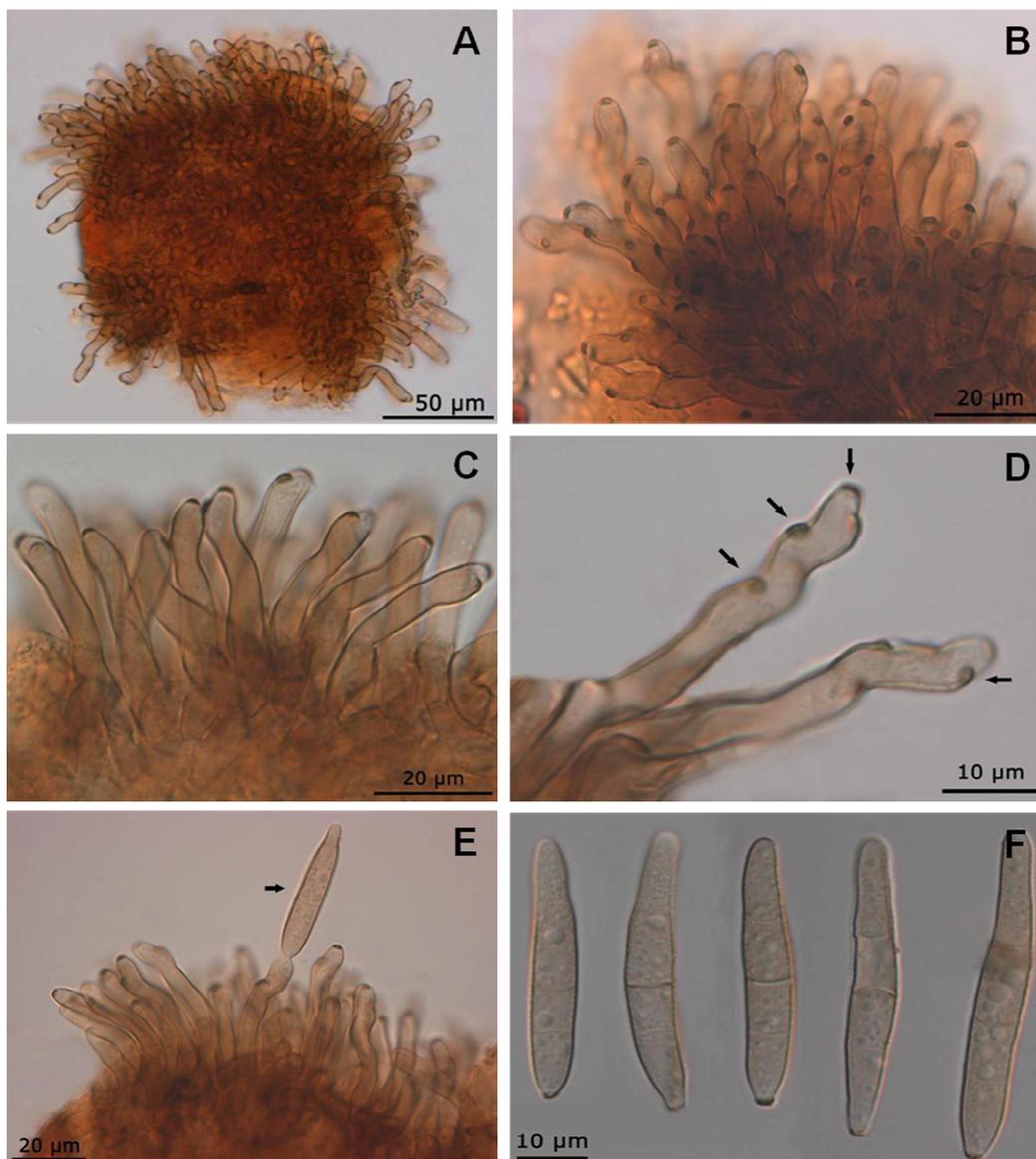


Figura 6 – Imagens de microscopia óptica da fase anamórfica do agente causal da mancha preta, *Passalora personata*. (A) Estroma globular denso. (B) Conidióforos geniculados em fascículos densos. (C) Detalhe da estrutura fasciculada dos conidióforos e em seu aspecto liso. (D) Cicatrizes nos conidióforos (setas). (E) Crescimento simpodial das células conidiogênicas, formando conídios solitários (seta). (F) Mostra o padrão de formato encontrado de conídios, cilíndricos, podendo ser retos ou ligeiramente curvados, parede levemente áspera, podendo ser uni ou multisseptados, com ponta arredondada e base cônica com hilo conspicuo. Fotos de Liberato e Shivas (2006).

Sintomatologia

A patogênese de *P. personata* começa com a disseminação (por respingos de água, vento, insetos ou utensílios agrícolas), deposição e adesão dos conídios na superfície da planta e o crescimento, durante os períodos de alta umidade, do tubo germinativo sobre essa superfície. Em seguida, as pontas das hifas se modificam em apressórios (estrutura de fixação do fungo na superfície do tecido hospedeiro), onde com

o aumento da pressão de turgor, impulsiona a penetração na parede da célula hospedeira através da epiderme, ou mais frequentemente, pelos estômatos, originando ramificações hifais especializadas, os haustórios (responsáveis pela captação de nutrientes do hospedeiro), permitindo assim o crescimento intracelular do micélio e a sua invaginação na membrana plasmática da célula hospedeira (McDonald, et al., 1985; Pretorius, 2006; Nobile, et al., 2008; Leal-Bertioli, et al., 2010). O ciclo de infecção é concluído em cerca de 20-25 dias e os primeiros sintomas em cultivares suscetíveis são visíveis depois de 10-14 dias (Shokes & Culbreath, 1997), podendo todo o processo ser recomeçado pela dispersão de novos conídios ou ascósporos (Figura 7).

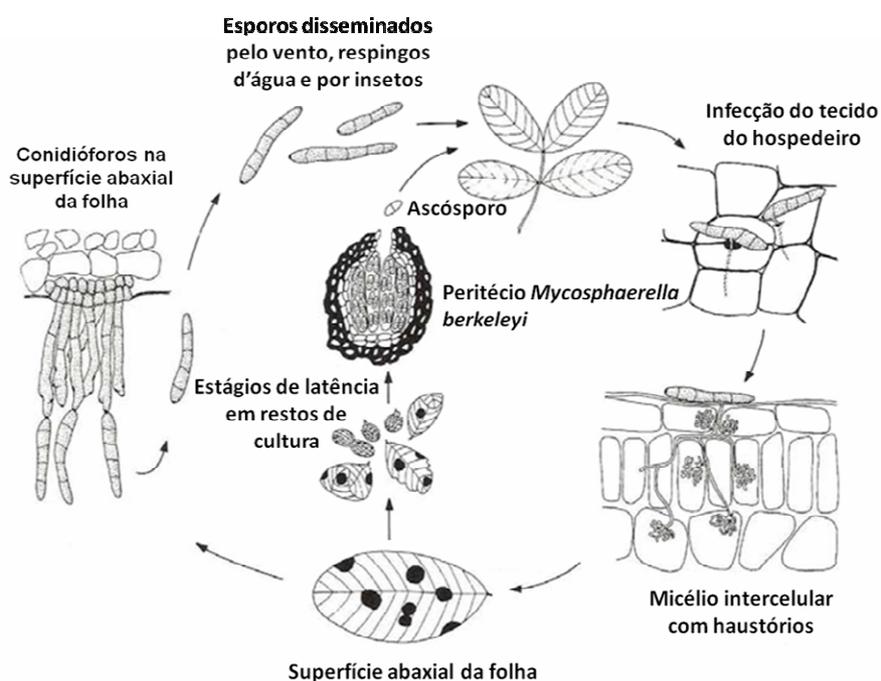


Figura 7 – Ciclo de vida do *P. personata*, agente causal da mancha preta no amendoim. Imagem adaptada de McDonald e colaboradores (1985).

Os sintomas da mancha preta começam nas folhas mais velhas, avançando progressivamente para as folhas mais novas (Almeida, 2001). Aparecem como pequenas manchas marrons com halos cloróticos (Figuras 8A e B) (Hagan, 1998), que mais tarde se alargam tornando-se castanho-claras e castanho-escuras enegrecidas (Jyosthna, et al., 2004) medindo até 7mm de diâmetro com bordas mais uniformes nos folíolos e halo amarelo indistinto ou ausente (Figura 8C-H). Porém, em manchas maduras, ocasionalmente, desenvolve-se um halo amarelo (Barreto, 1997; Leal-Bertioli, et al., 2010). Além disso, manchas podem aparecer nos pecíolos, hastes central e lateral, pedúnculo (*peg*) e vagens, e são mais alongadas, do que as vistas nos folíolos, mas, só

são encontradas em amendoins severamente afetados pela mancha preta (Hagan, 1998; Almeida, 2001).

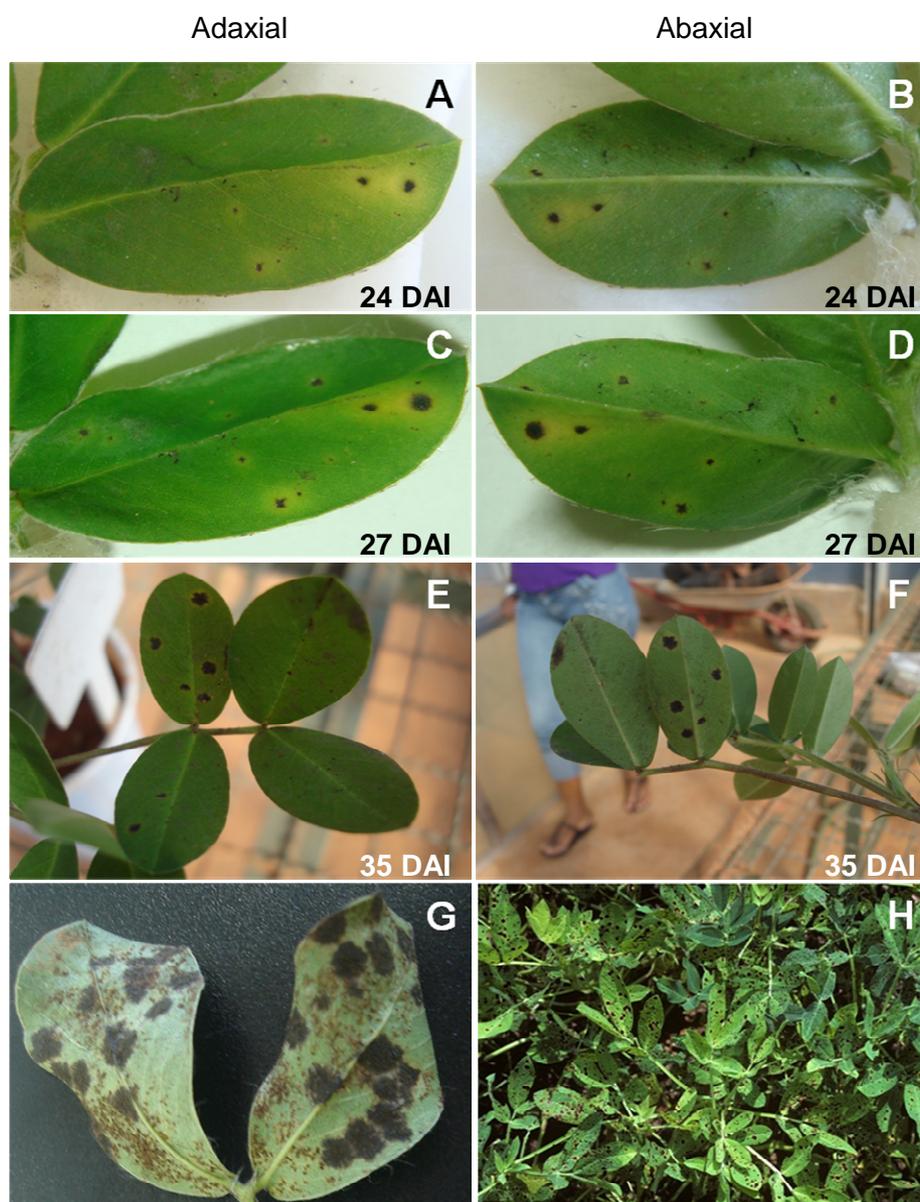


Figura 8 – Folhas de *A. hypogaea* infectadas por *P. personata*. (A-G) *A. hypogaea* cv. IAC-Tatu, obtidos e cultivados em casa de vegetação da Embrapa/Cenargen, Brasília, DF. (A, C e E) Superfície adaxial; (B, D e F) Superfície abaxial. (A-D) Folhas destacadas, mantidas umedecidas em placa de Petri, mostrando sintomas iniciais da mancha preta (pequenas manchas cloróticas). Inoculadas artificialmente; (E-F) Cultivar crescida e mantida em copo plástico, em casa de vegetação. Inoculadas artificialmente; (C-D) Aumento do tamanho das lesões e a uniformização das bordas arredondadas. (G) Amostra coletada de casa de vegetação, contendo sintomas mais avançados da mancha preta e de ferrugem (*Puccinia arachidis*), ambas mostrando seus corpos de frutificação na parte abaxial. Infecção natural; H) Imagem de um campo de plantio de amendoim, no sul dos Estados Unidos, infectado por *P. personata* (DWSPL & Cattlin, s/d). Fotos (A-G) por Andressa da C. Q. Martins. DAI: Dias após a inoculação.

A esporulação conidial ocorre em quase todas as lesões, principalmente na superfície abaxial da folha, onde as lesões são pretas, a partir de um visível estroma escuro com aparência pouco áspera, elevada ou acolchoada (Figura 8G). Isso é devido ao fato do *P. personata* produzir muitos esporos, principalmente, na parte abaxial da folha (Hagan, 1998; Almeida, 2001; Jyosthna, et al., 2004; Ramos, 2007). Em ataques severos, essa esporulação também pode ser observada na superfície adaxial (Barreto, 1997), onde as lesões são quase pretas.

Dependendo das condições ambientais e do histórico de cultivo, os sintomas foliares geralmente aparecem entre 30-50 dias após o plantio (Damicone & Melouk, sem data).

A doença reduz a área foliar e provoca queda prematura das folhas, além de enfraquecer as hastes e os *pegs*, podendo reduzir a produtividade em até 70% (Coutinho & Suassuna, 2008). Folhas severamente afetadas tornam-se verdes pálidas a amarelas, murcham e caem. A desfolha normalmente começa nas folhas mais baixas e continua para cima até que todas as folhas mais jovens do ápice também sejam perdidas (Damicone & Melouk, sem data; Hagan, 1998).

Epidemiologia

Duas variáveis meteorológicas que afetam fortemente o processo de infecção e favorecem as epidemias de mancha preta são a temperatura (20-26°C) e o molhamento foliar (≥ 10 horas) (Butler, et al., 1994; Barreto, 1997). A temperatura mais favorável para produção de esporos durante períodos de umidade relativa $\geq 95\%$ é 20°C (Alderman & Nutter Jr, 1994), sendo que a porcentagem de germinação de conídios diminui com o aumento da temperatura, acima de 20°C, enquanto que a 20°C, leva cerca de 12 horas para que 50% dos conídios germinem (Sommartya & Beute, 1986). Além disso, outro requisito importante para esporulação é a luz (Pretorius, 2006).

Embora o teleomorfo desse fungo seja conhecido, os ascósporos não são geralmente considerados como fontes importantes de inóculo primário (McDonald, et al., 1985). Conídios produzidos por conidióforos em restos culturais no solo e em plantas de amendoim fora de época servem como a principal fonte primária de inóculo inicial (Pretorius, 2006), e de onde os conídios são disseminados a curta distância pelo vento ou por gotas de água, dando início às infecções primárias (Sanine, 2006). Os esporos

produzidos nessas infecções são disseminados principalmente pela água, conduzindo a ciclos secundários de infecção (Damicone & Melouk, sem data).

Mesmo sendo a principal fonte de inóculo (conídios), os ascósporos, clamidósporos e fragmentos de micélio também podem ser infecciosos (Leal-Bertioli, et al., 2010), pois o fungo sobrevive de uma estação a outra como conídios ou massa micelial, em restos de culturas ou plantas voluntárias. Estes podem permanecer por até 10 meses e é por este motivo que a doença se mantém quando se faz uma ou duas safras/ano (Almeida, 2001).

Controle

O controle da mancha preta pode ser feito através da integração de diversas medidas que visa reduzir o inóculo inicial e a taxa de infecção, retardando o início da epidemia (Almeida, 2001). É recomendável práticas agrícolas com rotação de culturas com espécies não pertencentes ao gênero *Arachis*, por dois a três anos, medida esta já adotada na renovação de canaviais, impedindo o acúmulo do patógeno no solo. Outras medidas são: incorporação de restos culturais através de aração profunda, retardando o início das epidemias; destruição de plantas voluntárias, que servem de inóculo secundário para novas epidemias; limpeza dos equipamentos; aplicações múltiplas de produtos químicos em intervalos regulares. Dependendo da persistência do patógeno na planta, estes podem ser aplicados 3-4 vezes durante o ciclo, a partir dos 35-40^o dias de emergência ou por meio de modelos de previsão de ocorrência da doença, visando à racionalização do seu uso (Damicone & Melouk, sem data; Almeida, 2001; Kishore, et al., 2001; Pensuk, et al., 2003).

Aplicações de fungicidas são eficazes no controle da mancha preta, no entanto, a maioria dos agricultores nos países em desenvolvimento não tem recursos para comprar o fungicida (Pande & Rao, 2001; Pensuk, et al., 2003). Embora a aplicação de fungicida seja eficaz no controle das doenças foliares do amendoim, os problemas associados à sua aplicação como alto custo, poluição ambiental e as chances de desenvolvimento de resistência do patógeno não fazem deste uma boa escolha para o seu uso continuado (Almeida, 2001; Luo, et al., 2005b; Kumar & Kirti, 2011), necessitando o desenvolvimento de componentes alternativos de manejo da doença mais econômica e ecologicamente amigáveis (Kishore, et al., 2001). O desenvolvimento de cultivares resistente é considerado um dos meios mais eficazes para reduzir as perdas de rendimento causadas por patógenos na agricultura (Luo, et al., 2005b; Nobile, et al., 2008).

Assim, um dos principais objetivos dos programas de melhoramento de amendoim em todo o mundo (Figura 9) é o desenvolvimento de cultivares tolerantes ou resistentes a doenças, que possam ser usadas no manejo ou no controle de doenças do amendoim (Wynne, et al., 1991). Com isso, o caminho alternativo é o de se identificar no germoplasma de amendoim resistente a mancha preta e incorporar a resistência em cultivares adaptadas (Pensuk, et al., 2003).

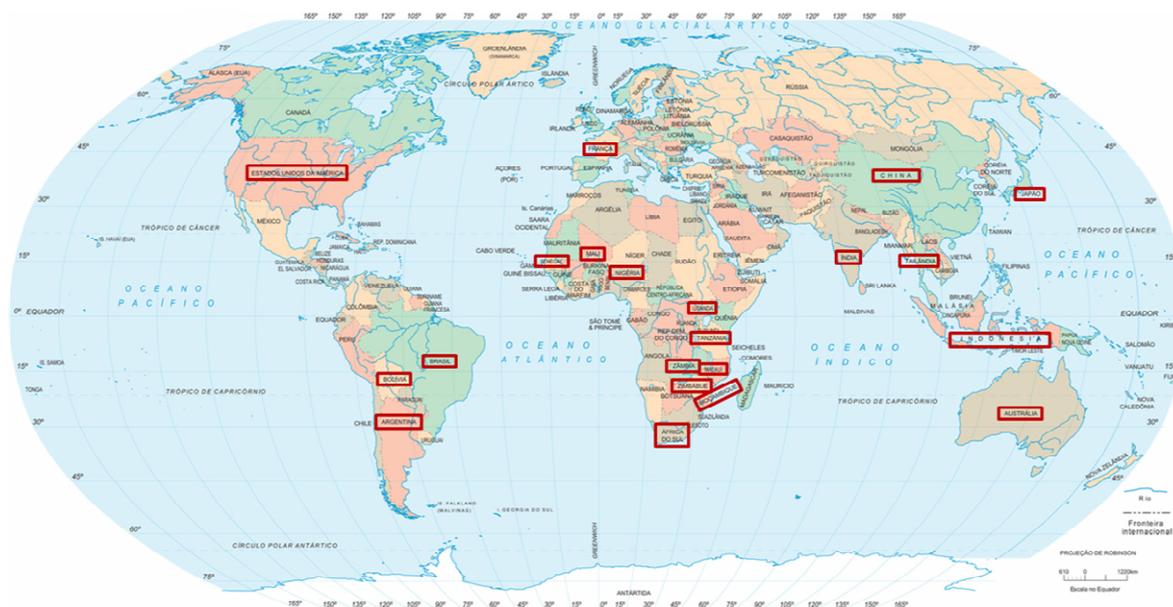


Figura 9 – Países que possuem programas de melhoramento de amendoim no mundo. Adaptada de <http://www.quiageo-mapas.com/mapa-mundi.htm>. América do Norte: Estados Unidos (Texas, Flórida, Geórgia, Alabama, Carolina do Norte, Novo México, Oklahoma e Virgínia); América do Sul: Brasil (Distrito Federal¹, São Paulo², Paraíba³), Bolívia e Argentina; Europa: França; África: Senegal, Mali, Nigéria, Uganda, Tanzânia, Zâmbia, Malawi, Zimbábue, Moçambique e África do Sul; Ásia: Índia, China, Tailândia, Japão e Indonésia; Oceania: Austrália. ¹IAC – Instituto Agrônomo de Campinas – SP; ²Embrapa Recursos Genéticos e Biotecnologia – DF; ³Embrapa Algodão – Campina Grande – PB.

2.7. Interação planta patógeno

Apesar das plantas estarem sob exposição contínua a vários micro-organismos na natureza, elas são geralmente resistentes à maioria dos patógenos, pois a habilidade de um micro-organismo causar doença na planta é tratada como uma exceção e não como a regra (Staskawicz, 2001). A resistência a doenças de plantas é resultado da co-evolução entre planta e patógeno (Graham & Graham, 1991), e se dá por um complexo sistema multicomponente de mecanismos de defesa estruturais e bioquímicos da planta, podendo ser de origem constitutiva (pré-formado) ou induzida (pós-formado), que atuam de maneira dinâmica e coordenada para reconhecer e impedir a invasão do patógeno (Pascholati & Leite, 1995).

As barreiras pré-formadas são o primeiro obstáculo que o patógeno encontra ao tentar invadir e colonizar a planta, deparando-se logo em seguida, com as respostas induzidas (barreiras pós-formadas), onde pode ocorrer a resposta de hipersensibilidade (HR) com a indução de fitoalexinas e várias proteínas, incluindo as proteínas relacionadas à patogênese (PR). Além dessa resposta de defesa induzida, existem ainda, diversos compostos que desencadeiam rotas de sinalização específicas e coordenam resposta de defesa contra o patógeno, tais como a do ácido jasmônico (JA), ácido salicílico (SA) e do hormônio etileno (ET), que precisam ser vencidos, para então o patógeno conseguir completar seu ciclo e causar a doença na planta (Mysore & Ryu, 2004).

Mecanismos de defesa podem ser ativados por PAMPs (padrões moleculares associados a patógenos – *Pathogen-associated Molecular Patterns*), denominada PTI (imunidade desencadeada por PAMPs – *PAMP-triggered immunity*), ou desencadeada por efetores e denominada ETI (*Effector-triggered immunity*) (Göhre & Robatzek, 2008). A PTI está geralmente associada à imunidade não-hospedeira e à imunidade basal (respostas imediatas ao contato com o patógeno). Enquanto que a ETI está associada à HR e à resistência sistêmica adquirida (SAR).

A PTI e a ETI fazem parte do modelo atualmente conhecido como Dogma Central da Fitopatologia, que descreve um processo de co-evolução denominado quebra de braço evolucionária (*evolutionary arms-race*).

A resistência mostrada por toda espécie de planta a maioria dos patógenos potenciais, é conhecida como resistência de não-hospedeira, sendo a forma mais comum e durável de resistência. De acordo com Mysore e Ryo (2004) a resistência não-hospedeira pode ser dividida em dois tipos: tipo I – não produz nenhum sintoma visível (necrose), e tipo II – está sempre associado à rápida formação de necrose localizada (HR). Em contraste a resistência de não-hospedeira, a resistência de hospedeira, é expressa por alguns genótipos de plantas dentro de uma espécie hospedeira suscetível, sendo geralmente patógeno-específica (Heath, 2000), podendo ser mediada por genes *R*.

Dentre os mecanismos de defesa das plantas contra a infecção por patógenos, aqueles mediados pelos genes *R* são os melhores caracterizados (Elvira, et al., 2008). No modelo clássico de interação gene a gene, proposto por Flor (1971), onde barreiras químicas e estruturais impedem o sucesso do patógeno, essa resistência ocorre quando planta possui o gene *R* (de resistência dominante) e o patógeno o gene *Avr* (de avirulência dominante), gerando uma reação incompatível, onde o produto do gene *R* reconhece o produto do gene *Avr*, que atua como elicitador, desencadeando sinais para

uma ativação rápida dos mecanismos de defesa e resistência da planta. A ausência do gene *R* na planta, ou do gene *Avr* no patógeno ou de ambos, resulta na interação compatível, isto é, não ocorre o reconhecimento do patógeno pela planta, desencadeando uma resposta de defesa mais lenta e a instalação da doença (Hutcheson, 1998; Cordeiro & Sá, 1999; Staskawicz, 2001). Os genes *Avr* são assim denominados, porque na presença dos genes *R* causam avirulência (Bent & Mackey, 2007).

Essa teoria de Flor pode ser interpretada, por meio da nova terminologia empregada pelo modelo do Dogma Central da Fitopatologia, da seguinte forma: as plantas detectam componentes genéricos conservados dos micro-organismos (PAMPs) desencadeando a PTI. No entanto, alguns micro-organismos desenvolvem efetores (fatores de virulência) que podem suprimir a resposta de defesa das plantas. Diante disso, as plantas desenvolveram genes de resistência (genes *R*) específicos, onde seu produto detecta os efetores do patógeno, desencadeando a ETI (Bent & Mackey, 2007).

Logo após o reconhecimento genético entre planta-patógeno e a transdução de sinais ocorre a ativação de genes, onde dois grupos podem estar envolvidos nesta fase, os genes de ativação precoce e os de ativação tardia, sendo que, os genes de ativação precoce podem, hipoteticamente, estar envolvidos na regulação e atividade dos genes de ativação tardia (Cordeiro & Sá, 1999).

Uma série de alterações fisiológicas e bioquímicas, juntamente com a resposta molecular, ocorre durante o período entre o aparecimento do estresse na planta, a partir da invasão do patógeno e subsequente interação planta-patógeno (You-Xiong, et al., 2011). Portanto, as interações planta-patógeno são caracterizadas por mudanças na expressão gênica de ambos os organismos (Benito, et al., 1996).

A complexidade dos mecanismos de defesa vegetal implica em uma grande mudança na atividade metabólica, resultante da expressão regulada de genes responsivos ao estresse.

Dessa forma, torna-se essencial a identificação de genes envolvidos nestas respostas tanto para o entendimento desses mecanismos de defesa da planta contra fitopatógenos (Somssich & Hahlbrock, 1998; Lee & Choi, 2002; Corrent, et al., 2009; Govind, et al., 2009) quanto para estabelecer estratégias em programas de melhoramento de plantas, especialmente para cultura de amendoim (Nobile, et al., 2008).

2.8. Análise da expressão gênica

Progressos no campo da genômica funcional têm permitido o estudo de respostas de plantas em todos os níveis do transcrito, revelando a natureza complexa das respostas multigênicas em plantas (Payton, et al., 2009).

O transcrito é um conjunto completo de transcritos (RNAs) em uma célula e sua quantidade, para um estágio de desenvolvimento específico ou condição fisiológica.

A compreensão transcritômica ajudará a interpretar os elementos funcionais do genoma, revelando os componentes moleculares de células e tecidos, visando: catalogar todas as espécies de transcritos, incluindo mRNA (RNA mensageiro), RNAs não codificadores e pequenos RNAs; determinar a estrutura de transcrição dos genes, em termos de locais de início, extremidades 5' e 3', os padrões de *splicing* e outras modificações pós-transcricionais; quantificar a mudança nos níveis de expressão de cada transcrito durante o desenvolvimento e sob diferentes condições, tornando seu estudo importante para entendimento dos processos de transcrição, uma vez que a alteração nos níveis de expressão está diretamente relacionada à modificações na fisiologia, metabolismo e conseqüentemente, ao processo de adaptação celular (Passos, et al., 2000; van Vliet, 2009; Wang, et al., 2009).

Com isso, o estudo da expressão gênica tem demonstrado ser uma importante ferramenta no entendimento dos processos biológicos em nível molecular, podendo este, contribuir para caracterização da resistência de plantas.

Por meio da comparação da expressão gênica diferencial entre organismos em condições normais e submetidos a algum tipo de estresse, é possível, identificar genes que podem estar envolvidos nos mecanismos de defesa, além da possibilidade de tentar identificar em quais vias atuam ou são reguladas por esses genes (Leal-Bertioli, et al., 2003).

O sequenciamento do genoma revela todos os genes presentes dentro de um organismo, mas não revela quais dos genes são ativos em diferentes vias metabólicas, tecidos ou estágios de desenvolvimento (PGP, 2011). Por isso que, é importante, mesmo para espécies para as quais a sequência completa do genoma já está disponível, incluindo algumas leguminosas de grande expressão produtiva como: soja, canola e, recentemente, feijão-guandu, a construção de banco de genes transcritos (ou de ESTs – *Expressed Sequence Tags*, na terminologia, anteriormente utilizado para o sequenciamento de clones de bibliotecas de cDNA). No caso de plantas cujo genoma

ainda não está disponível, como o caso do amendoim, essa ferramenta se torna ainda mais essencial para estudo de genômica e prospecção de genes.

ESTs são sequências curtas (parciais) de DNA, geradas pelo sequenciamento em larga escala de clones selecionados, aleatoriamente, de bibliotecas de cDNA (DNA complementar) construídas a partir de mRNA (RNA mensageiro), isolado em um estágio de desenvolvimento e/ou tecido em particular (Luo, et al., 2005a; Luo, et al., 2005b). Trata-se de uma abordagem genômica eficaz para a rápida identificação de genes expressos ou genes transcritos (Guo, et al., 2008), capaz de fornecer uma fonte para descobertas de novos genes, especialmente em plantas não modelo para as quais as sequências do genoma completo não estão ainda disponíveis, e para análises comparativas entre os organismos (Morozova, et al., 2009; Li, et al., 2010; Franssen, et al., 2011; Mondego, et al., 2011). Além disso, a partir de ESTs ou de genes transcritos é possível, ainda, analisar expressão em grande escala, fazer anotações genômicas, mapeamento genético, analisar traços quantitativos (QTLs – *Quantitative Traits Loci*), desenvolver marcadores moleculares, incluindo os microssatélites (SSR – *Simple Sequence Repeats*) e os de variações em um único nucleotídeo (SNPs – *Single Nucleotide polymorphism*) e produzir catálogos de genes (Mekhedov, et al., 2000; Ogihara, et al., 2003; Luo, et al., 2005a; Alves, et al., 2008; Guo, et al., 2009; Liang, et al., 2009; Payton, et al., 2009; Stalker, et al., 2009; Song, et al., 2010; Kumar & Kirti, 2011).

Um grande número de ESTs já está depositado no banco de dados do GenBank (dbEST) para as três maiores culturas de oleaginosas, concorrentes do amendoim em escala mundial de produção: soja (1.461.624), canola (643.937) e algodão (274.247 – dados das espécies cultivadas). Diferente disso, havia apenas 226.009 ESTs, para todo o gênero *Arachis*, depositados no GenBank, entre os quais 151.667 foram provenientes da espécie cultivada (*A. hypogaea*) e o restante oriundos de *A. duranensis* (35.291), *A. ipaënsis* (32.787) e *A. stenosperma* (6.264). Até o momento, a maioria dos ESTs gerados, para o gênero *Arachis*, tem tido como objetivo o de identificar genes potencialmente envolvidos em resposta ao estresse abiótico e biótico e o desenvolvimento de marcadores moleculares (Alves, et al., 2008). Enquanto que os ESTs depositados no dbEST para *A. stenosperma* são referentes a folhas e raízes saudáveis e inoculadas com *Bradyrhizobium* sp. e *M. arenaria*.

Recentemente foi desenvolvida a tecnologia de sequenciamento 454 GS-FLX para o sequenciamento de genes transcritos que representa um método rápido e eficiente de grande potencial para analisar as regiões transcritas do genoma (Torres, et al., 2008; Li, et al., 2010), pois remove muitas das etapas envolvidas no sequenciamento de genes

transcritos pelo método Sanger, que consumiam muito tempo, facilitando o sequenciamento do transcrito (Parchman, et al., 2010), além de dispensar a necessidade de clonagem, tirando vantagem de um eficiente método de amplificação *in vitro* de DNA, conhecido como PCR em emulsão (Margulies, et al., 2005; Roche A. S., 2006).

2.9. Pirosequenciamento

Desenvolvida pela *454 Life Sciences Corporation*, agora também conhecida por *Roche Applied Science* (<http://www.my454.com>), a tecnologia de sequenciamento 454 foi a primeira plataforma de sequenciamento de nova geração (NGS – *Next Generation Sequencing*) a ser comercializada. Esta tecnologia é baseada em uma plataforma de pirosequenciamento que utiliza volumes em picolitros de DNA em um ambiente altamente multiplexados de fluxos capazes de produzir de 20 a 40 milhões de bases por corrida (Cheung, et al., 2006; Rothberg & Leamon, 2008; Carvalho & Silva, 2010).

Em resumo, trata-se de um sequenciamento baseado em síntese – o pirosequenciamento, onde a leitura da sequência é realizada pela combinação de reações enzimáticas que se inicia com a adição do dNTP, ocorrendo logo em seguida a liberação de um pirofosfato inorgânico (PPi), o qual é convertido em ATP, pela enzima ATP sulfurilase. A ATP é utilizada pela enzima luciferase para oxidar a luciferina, produzindo, desta forma um sinal de luz. As emissões de luzes são então captadas por uma câmera (CCD – *Charge-coupled device*) acoplada ao sistema 454 (Figura 10) (Rothberg & Leamon, 2008; Carvalho & Silva, 2010).

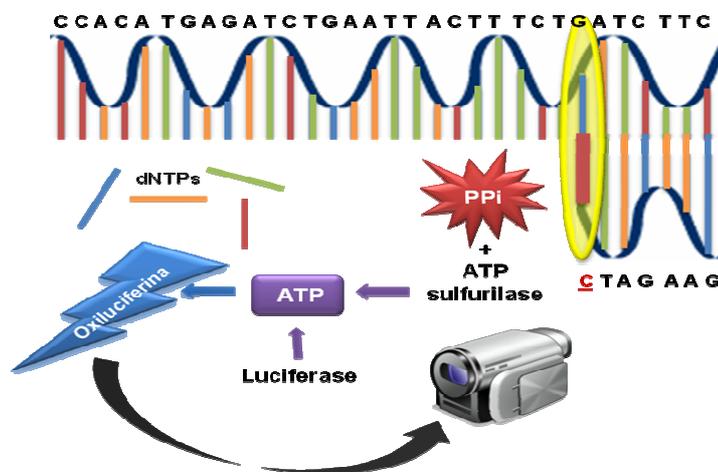


Figura 10 – Esquema resumido do processo de pirosequenciamento.

A plataforma *GS FLX Titanium* tem capacidade de produzir em média um milhão de leituras, com tamanho variando entre 300-450 bases, podendo ser até de 500 bases em condições de bibliotecas *shotgun* (o DNA ou cDNA é dividido aleatoriamente, em fragmentos, e cada um é seqüenciado), ou de 500-1.000pb (com a atual tecnologia *GS FLX Titanium XL+*) em 10 horas (Andreote, 2011; Gilles, et al., 2011; Natarajan & Parani, 2011).

O primeiro trabalho publicado mostrou que a plataforma 454 foi capaz de gerar a mesma quantidade de informação obtida com o sequenciamento Sanger com alta qualidade, a partir de clones BACs (*Bacterial Artificial Chromosome* – Cromossomo Artificial de Bactérias) de cevada (*Hordeum vulgare* L.) (Wicker, et al., 2006). Depois vieram outros trabalhos. Com *Medicago* (*Medicago truncatula* Gaertn.), confirmou-se que o pirosequenciamento era capaz de proporcionar descoberta de novos genes, além de confirmar o valor de leituras curtas, originadas de bibliotecas de cDNA normalizadas, para as anotações gênicas (Cheung, et al., 2006). Em milho (*Zea mays* L.), foi evidenciado que o uso conjugado da técnica de microdissecação com captura a laser (MCL) de transcritos aumenta a chance de transcritos raros, principalmente em órgãos altamente especializados como o meristema apical (Emrich, et al., 2006). Além disso, com *Arabidopsis* (*Arabidopsis thaliana* L. *sensu* Heynh) testou-se a capacidade dessa tecnologia em gerar genes transcritos e fornecer uma ampla cobertura do transcritoma, além da indicação de uma boa representatividade de todas as regiões transcritas e da adição de mais de 16.000 novos genes transcritos ao banco de dados já existente (Weber, et al., 2007). Novaes e colaboradores (2008) demonstraram, ainda, o potencial da plataforma 454 para descoberta de polimorfismos de única base (SNPs – *Single Nucleotide Polymorphism*), em Eucalipto (*Eucalyptus grandis* Hill ex Maiden). No caso, do Feijão-guandu (*Cajanus cajan* L.), a tecnologia 454 foi capaz de gerar dados mais abrangentes de transcritos, permitindo a análise para o entendimento da arquitetura do transcritoma e organização genômica com relação à dinâmica evolucionária, além da identificação de genes-candidatos envolvidos com resistência a doenças e desenvolvimento de marcadores SNPs e SSR (*Simple Sequence Repeat*). Essa produção de marcadores SNPs em grande escala para construção de um mapa de ligação de alta densidade, foi demonstrada no trabalho de Russel e colaboradores (2011), com groselha (*Ribes nigrum* L.). Enquanto que a identificação e produção de marcadores moleculares SSR, foi evidenciada, no estudo com grão de bico (*Cicer arietinum* L.) (Garg, et al., 2011). Do mesmo modo, para lentilha (*Lens culinaris* Medik.), produziu-se novos unigenes com subsequente anotação genômica e identificou-se

marcadores SSRs (Kaur, et al., 2011). Não obstante, com o uso do pirosequenciamento foi possível determinar não somente a sequência genômica completa do cloroplasto de feijão-da-china (*Vigna radiata* L.), como também a relação filogenética muito estreita entre *Vigna* e *Phaseolus* (Tangphatsornruang, et al., 2010) e desenvolvimento de marcadores SNPs e SSRs (Moe, et al., 2011).

Além dessas espécies, acima mencionadas, várias outras não modelos se beneficiaram do sequenciamento de segunda geração para criar banco de dados de transcrito, tais como: mangue-vermelho (*Rhizophora mangle* L.) e *Heritiera littoralis* Aiton (Dassanayake, et al., 2009), azeitona (*Olea europaea* L.) (Alagna, et al., 2009), Artemísia (*Artemisia annua* L.) (Wang, et al., 2009), castanha Americana (*Castanea dentata* Marsh. Borkh) e castanha chinesa (*Castanea mollissima* Blume) (Barakat, et al., 2009), *Tragopogon miscellus* G. B. Ownbey (Buggs, et al., 2010), ervilha (*Pisum sativum* L.) (Franssen, et al., 2011), pinhão manso (*Jatropha curcas* L.) (Natarajan & Parani, 2011), são alguns exemplos.

Como visto, este tipo de sequenciamento permite a montagem *de novo* de um organismo, pois possibilita uma maior profundidade de cobertura (número de leituras por nucleotídeo na fita molde – *template*). Dado pela fórmula $C = \frac{N \times L}{G}$, onde C = cobertura, N = número de leituras, L = tamanho médio das leituras e G = tamanho do genoma original) do genoma (Parchman, et al., 2010), com tamanho razoável de sequência (400-500pb), e tem sido o método amplamente utilizado no sequenciamento de espécies cujo genoma é pouco estudado, como o caso do amendoim.

2.10. PCR quantitativo (RT-qPCR)

As tecnologias e métodos utilizados na análise da expressão gênica para o estabelecimento geral dos padrões de expressão, bem como na seleção de expressão diferencial, requerem uma etapa adicional de refinamento e validação dos resultados que pode ser realizada por meio de *Northern Blotting*, a qual requer grande quantidade de RNA, além de ser laboriosa. Com isso, esta técnica vem sendo substituída pelo método de reação em cadeia da polimerase com transcrição reversa quantitativa (RT-qPCR), devido a sua alta sensibilidade, boa reprodutibilidade, precisão e resultados rápidos (Rajeevan, et al., 2001; Pfaffl, et al., 2002; Morgante, et al., 2009; Derveaux, et al., 2010; Gan, et al., 2011).

Lu e colaboradores (2005) avaliaram o nível de expressão de alguns genes relacionados à defesa, a partir de coroas de trigo (*Triticum aestivum* L.) inoculados com *Tilletia tritici* (Bjerk.), utilizando o RT-qPCR. Uzarowska e colaboradores (2009) utilizaram o RT-qPCR para validarem seis das 65 sequências diferencialmente expressas encontradas no experimento de microarranjo, em resposta a infecção do vírus do mosaico da cana-de-açúcar (SCMV – *Sugarcane mosaic virus*) em milho (*Zea mays* L.). Os seis genes-candidatos tinham associação à resposta de defesa. Meyer e colaboradores (2009) utilizaram o RT-qPCR para determinar os níveis de expressão relativa dos genes *Rpp4C1* ao *Rpp4C3* da linha suscetível de soja [*Glycine max* (L.) Merr.] (Wm82) e *Rpp4C1* ao *Rpp4C5* da linha resistente de soja (P1459025B), tanto no grupo controle quanto no grupo desafiado com *Phakospora pachyrhizi*, agente causal da ferrugem asiática da soja. Wagner (2008), se utilizou dessa tecnologia para validar genes-candidatos diferencialmente expressos, determinados pela técnica de hibridização Affymetrix GeneChip® com amostras de dois genótipos de cevada (*Hordeum vulgare* L.), Igri (suscetível) e Triton (resistente), desafiados com *Rhynchosporium secalis*, agente causal da escaldadura na cevada, que foram sequenciados a procura de SNPs e INDELS, visando a detecção de QTL; Shi e colaboradores (2011) confirmaram o padrão de expressão de dez transcritos derivados de polimorfismos de comprimento de fragmentos amplificados (AFLP – *Amplified Fragment Length Polymorphism*) a partir de folhas de feijão (*Phaseolus vulgaris* L.) inoculados com *Xanthomonas axonopodis* pv. *phaseoli*. Esses são apenas alguns exemplos do uso do RT-qPCR na validação da expressão gênica de genes-candidatos envolvidos na resposta de defesa.

O desenvolvimento de RT-qPCR facilitou a detecção de transcritos, aumentou o rendimento experimental e reduziu a quantidade de RNA de entrada necessária (Morozova, et al., 2009), em virtude de sua capacidade de detectar em cada ciclo da PCR, a quantidade de fragmentos amplificados (*amplicons*) acumulados (fase exponencial da reação), utilizando o monitoramento contínuo da fluorescência emitida durante a reação de amplificação (Ponchel, et al., 2003). O primeiro aumento significativo na quantidade de produto amplificado (*amplicon*) correlaciona-se com a quantidade inicial da amostra-alvo (*template*).

A RT-qPCR combina a amplificação de um alvo específico com métodos de quantificação, por meio de medidas da fluorescência associada à síntese de *amplicon* ao longo dos ciclos da PCR. Esta técnica funciona essencialmente da mesma maneira que a PCR convencional, com ciclos subsequentes em que ocorre a desnaturação do DNA, a hibridização de um par de *primers* (iniciadores) a sequências específicas e a extensão, mediada por uma DNA polimerase termoestável, de uma fita de DNA complementar,

resultando em um aumento exponencial do número de *amplicons* ao longo dos ciclos da reação. No entanto, na RT-qPCR esse aumento é registrado em tempo real ao longo da fase exponencial e não somente ao final da reação (Applied Biosystem, sem data).

Sendo assim, nesta fase exponencial (fase que fornece os dados mais precisos e exatos para a quantificação) dois valores são calculados: o limiar (*threshold line*) e o C_T (*Cycle threshold*) ou C_q (Ciclo de quantificação – *Quantification cycle value*) (Bustin, et al., 2009). O aumento do sinal fluorescente é captado a cada ciclo até atingir um limiar (*threshold*). E C_q é o ponto que detecta o ciclo na qual a reação atinge o limiar da fase exponencial (Figuras 11 e 12) (Life Technologies, s/d; Almeida & Saddi, 2007; Peirson & Butler, 2007).

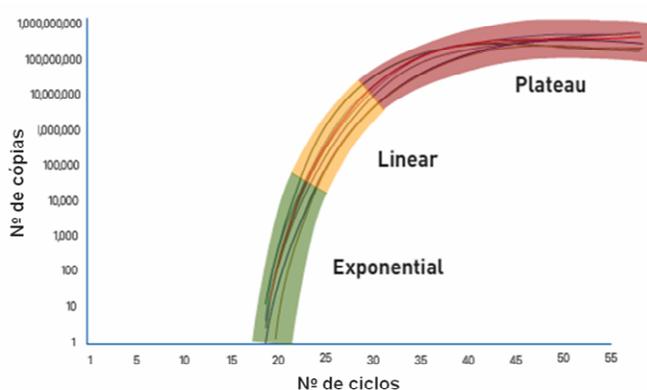


Figura 11 – Plotagem mostrando o número de ciclos da RT-qPCR e o número de cópias do material amplificado, em cada ciclo durante a RT-qPCR, além dos três estágios de amplificação: Exponencial (fase onde os reagentes da PCR estão em abundância, dobrando a cada ciclo os *amplicons* gerados); Linear (Reagentes começam a ficar escassos, diminuindo a quantidade de *amplicons* produzidos e desacelerando a reação de PCR) e Platô/Plateau (os reagentes já esgotaram e cessando a reação). Imagem adaptada de Life Technologies, sem data.

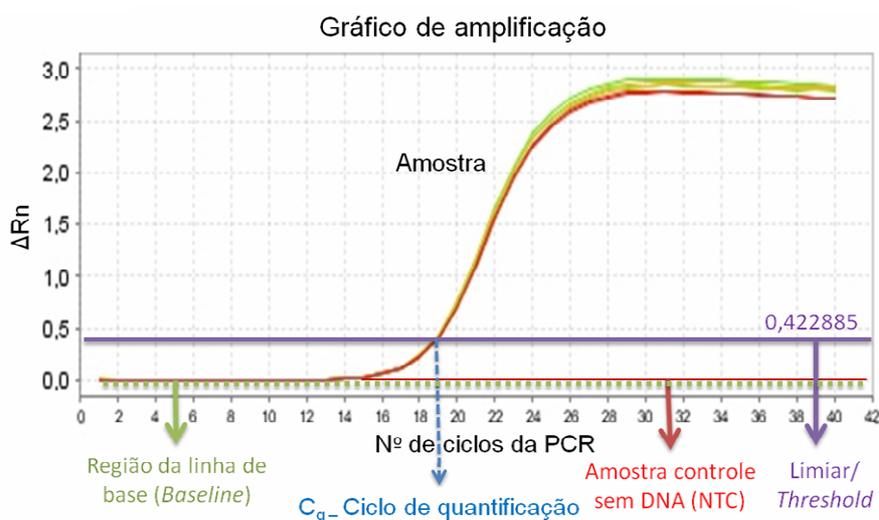


Figura 12 – Curva de amplificação do RT-qPCR, número de ciclos da PCR x ΔRn (magnitude da fluorescência emitida), mostrando alguns componentes e valores que devem ser analisados. Seta e linha pontilhada verde: linha de base (*baseline*) há pouca mudança no sinal de fluorescência. Seta

pontilhada azul: C_q (Ciclo de quantificação), número do ciclo no qual a fluorescência ultrapassa o *threshold*. Seta e reta vermelha: controle sem DNA (NTC – *No Template Control*) verifica a qualidade da amplificação e possíveis contaminações. Seta e reta lilás: limiar ou o *threshold*, nível de sinal do fluoróforo (corante *repórter* – corante de referência interna), corrige flutuações causadas por alterações na concentração ou volume.

A fluorescência mais utilizada é SYBR® Green I, um corante fluorescente que se incorpora ao DNA dupla fita, para detectar o produto da PCR conforme ele se acumula durante os ciclos da reação, permitindo o monitoramento da amplificação de qualquer sequência de DNA dupla fita, não sendo necessárias sondas, por meio da emissão de luz verde, captada pelo instrumento de RT-qPCR.

Na detecção utilizando o sistema SYBR® Green, é importante executar uma análise da curva de dissociação (*melting curve*) após a RT-qPCR para garantir que o *amplicon* desejado foi detectado, isto porque, este corante intercala em todo e qualquer DNA dupla fita presente na amostra, podendo gerar falso-positivos, ou seja, ele também pode se ligar à sequências não específicas (produtos inespecíficos) e, dímeros de *primers* (Figura 13) (Life Technologies, s/d; Applied Biosystem L. T., s/d; Novais, et al., 2004).

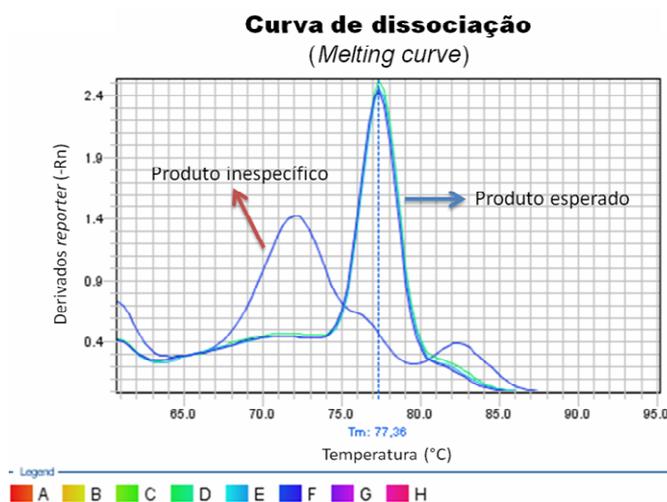


Figura 13 – Gráfico com os dados recolhidos durante o estágio de curva de dissociação (*melting curve*). Picos na curva de dissociação podem indicar a temperatura de dissociação (T_m – *melting temperature*) do alvo ou pode identificar amplificação inespecífica. O pico mais evidente localizado no centro do gráfico (seta azul) indica o ponto de dissociação do *amplicon* (amplificação de produto específico/esperado), com T_m a 77,36°C. Contaminantes, gerando produtos inespecíficos ou dímeros de *primers* aparecem como picos adicionais, vistos à esquerda do gráfico (seta vermelha). $-R_n$ corresponde a primeira derivada negativa da fluorescência normalizada pelo *repórter* (SYBR Green) durante a amplificação da PCR.

Todos os formatos de PCR quantitativa detectam fluorescência em tempo real, e usa o valor de C_q para quantificar montantes iniciais de DNA/cDNA ou para estabelecer uma curva padrão para estudos de expressão gênica ou outras análises comparativas (Life Technologies, s/d).

Devido à diversidade quase que infinita de possíveis condições experimentais e da incapacidade de controlar todos os parâmetros experimentais, um gene de referência deve ser considerado como sendo válido, desde que tenha sido validado nas mesmas condições que os usados para avaliar a expressão do gene-alvo, permitindo a avaliação da estabilidade de expressão de genes-candidatos sob uma variedade de condições (Gutierrez, et al., 2008).

Genes de referência são genes presentes em todos os tipos de células nucleadas, envolvidos em processos celulares básicos e que são expressos constitutivamente na célula, uma vez que são necessários para a sobrevivência da célula. Porém, estudos indicam que a expressão desses genes constitutivos pode variar, e não há um gene de referência universal para todos os experimentos (Pfaffl, 2001; Gutierrez, et al., 2008; Morgante, et al., 2009; Lee, et al., 2010). Com isso, um estudo recente foi realizado por Morgante e colaboradores (2011) onde foi avaliada a estabilidade de dez genes de referência candidatos para normalização em RT-qPCR de folhas e raízes de três espécies silvestres (*A. magna*, *A. duranensis* e *A. stenosperma*) e a espécie cultivada (*A. hypogaea*) do gênero *Arachis*, submetidos aos estresses bióticos e abióticos, permitindo o uso de combinações específicas de genes de referência de acordo com a espécie, órgão e estresse avaliados por meio do RT-qPCR. De acordo com Morgante e colaboradores (2011), *Actina* (*ACT1*), proteína ribossomal 60S (*60S*) e duas *ubiquitinas* (*UBI1* e *UBI2*) foram os quatro principais genes de referência por eles avaliados, e que se mostraram serem adequados como genes de referência universal inter-espécies de *Arachis*.

3. OBJETIVOS

3.1. Geral

Descobrir genes associados à resposta de defesa da planta durante interação planta-patógeno, por meio da análise massal do transcrito de espécie silvestre de *Arachis* (*A. stenosperma*) resistente ao fungo causador da mancha preta do amendoim, *P. personata*, bem como o desenvolvimento de um banco de dados dos genes transcritos durante esta interação.

3.2. Específicos

- a) Realizar bioensaios com *A. stenosperma* e *P. personata*, visando à obtenção de RNA total da interação e controle;
- b) Isolar, quantificar e produzir cDNA, a partir do RNA total de folhas de *A. stenosperma* inoculada e não inoculada com *P. personata*;
- c) Realizar anotação genômica de um banco de dados de genes transcritos em *A. stenosperma* obtidos através da tecnologia 454;
- d) Identificar genes diferencialmente expressos durante a interação *A. stenosperma* x *P. personata* e seu controle, visando à identificação daqueles associados à resposta de resistência da planta;
- e) Validar a expressão diferencial de genes-candidatos selecionados por meio da tecnologia de reação em cadeia da polimerase com transcrição reversa quantitativa (RT-qPCR).

4. MATERIAL E MÉTODOS

4.1. Bioensaio de *A. stenosperma* e *P. personata*

4.1.1. Plantio das sementes de *A. stenosperma* e *A. hypogaea* cv. IAC-Tatu

Os experimentos foram conduzidos em casa de vegetação e no laboratório de Interação Molecular Planta Praga III, no prédio de Biotecnologia da Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF, em períodos intercalados entre dezembro de 2009 a maio de 2011.

Sementes de *A. stenosperma*, $2n=2x=20$, acesso VSv 10309, foram obtidas do Banco Ativo de Germoplasma (BAG) da Embrapa Recursos Genéticos e Biotecnologia, Brasília – DF, e germinadas, sem adição de Ethrel ou fungicida, em copos plásticos de 500 mL, perfurados na parte inferior, contendo uma mistura de solo estéril, sulfato de amônia, calcário e Mastermins, e mantidas em casa de vegetação (Figura 14). A mesma metodologia de germinação foi utilizada com sementes de *A. hypogaea* cv. IAC-Tatu ($2n=4x=40$) também obtidas a partir do BAG da Embrapa/Cenargen.



Figura 14 – Sementes de *A. stenosperma* (V10309) e *A. hypogaea* cv. IAC-Tatu, germinadas em copos plásticos de 500 mL. Foto por Andressa da C. Q. Martins.

Após 30 dias da germinação, as plantas foram divididas em dois grupos (controle/testemunha e inoculado/tratado – Figura 15), com o cuidado para que cada grupo incluísse plantas com o mesmo estágio fisiológico.



Figura 15 – Divisão das plantas em dois grupos (controle e inoculado). Foto por Andressa da C. Q. Martins.

4.1.2. Isolamento e preparo do inóculo com *P. personata*

O isolado de *P. personata*, utilizado nos experimentos, foi coletado de lesões foliares de *A. hypogaea* cv. IAC-Tatu, em campo naturalmente infectado ao redor da Embrapa Cerrados (Planaltina – DF). Os esporos do fungo, encontrados nestas lesões, foram raspados utilizando-se uma espátula e, posteriormente, transferidos para placa de Petri contendo, aproximadamente 20 mL de meio BDA (Batata-Dextrose-Ágar) com o antibiótico levofloxacina (100mg/mL), e incubadas a 23°C em câmara de crescimento, sob 24 horas de luz, por 20 dias, período mínimo necessário para o início da esporulação do fungo. Após esporulação, o fungo foi observado quanto à morfologia dos seus conídios, conidióforos, bem como o formato de seus corpos de frutificação, por meio de microscópio ótico, para sua correta identificação. A partir disso, o fungo era periodicamente repicado e armazenado em câmara de crescimento.

A partir do 20º dia, procedeu-se a repicagem do fungo, transferindo-se, após breve agitação, 1 mL de suspensão de raspagem do meio, contendo micélio e conídios, em água, para novo meio BDA (100 mL) em Erlenmeyer, em mesmas condições de cultivo acima citadas, proporcionando superfície e quantidade de nutrientes maiores para

o crescimento do fungo, do que o oferecido em placa de Petri. Após esporulação, os Erlenmeyers foram transferidos para câmara fria a 4°C.

Para o preparo de meio BDA, utilizou-se 13g de BDA (Acumedia) completando com quantidade suficiente de água destilada para um volume final de 330 mL. Após homogeneização do meio, distribuiu-se 100 mL deste para cada Erlenmeyer de 250 mL, fechados com um chumaço de algodão, gazes e papel pardo, para serem autoclavados por 40 minutos, a 121°C e 1atm de pressão.

Uma suspensão fúngica foi preparada em Tween 20 (USB®) a 0,5%, a partir de raspagens de colônias de *P. personata* repicadas em meio BDA. Após breve agitação da suspensão para liberação dos conídios, estimou-se, por média de contagem em câmara de Neubauer, uma suspensão de 5×10^4 esporos/ml. Outro tubo Falcon de 50 mL foi preparado contendo apenas a solução de Tween 20 a 0,5%.

4.1.3. Inoculação das plantas com *P. personata* e coleta do material vegetal

Sete plantas, de cada espécie utilizada no experimento, tiveram sua parte aérea (caule, ramos e folhas) coletada como ponto zero ou ponto inicial do experimento (para análise e comparação da expressão gênica antes do início do experimento com os demais pontos coletados), e rapidamente envolvidas em papel alumínio, devidamente identificados e submersos em nitrogênio líquido até posterior armazenamento em refrigerador a -80°C e isolamento de RNA total. Após a coleta do ponto zero, deu-se início a inoculação das plantas do grupo inoculado, as quais tiveram suas folhas pinceladas com a suspensão fúngica, tanto na parte adaxial quanto na parte abaxial de todos os seus folíolos. O mesmo procedimento foi realizado com as plantas do grupo controle, sendo pinceladas também, ambas as partes dos folíolos, contudo, apenas com uma solução de Tween 20 a 0,5% (Figura 16).



Figura 16 – Esquema de inoculação, coleta e armazenamento das plantas de *A. stenosperma* e *A. hypogaea* cv. IAC-Tatu, durante os bioensaios. (A) Inoculação das plantas. (B) Coleta. (C) Identificação. (D) Armazenamento durante a coleta (N₂ líquido). (E) Acondicionamento final após a coleta (refrigerador -80°C).

A parte aérea de plantas do grupo controle e do grupo inoculado, de ambas as espécies estudadas, foram coletadas às 24, 48 e 72 horas após a inoculação (HAI), seguindo o mesmo procedimento de coleta realizado no ponto zero.

Este ensaio foi realizado oito vezes, em épocas diferentes do ano, porém, somente quatro deles (4º bioensaio – Dezembro/2009; 5º bioensaio – Março/2010; 6º bioensaio – Julho/2010 e o 8º bioensaio – Maio/2011), foram utilizados neste estudo, tendo em vista a obtenção de repetições biológicas para estudos de expressão gênica.

4.2. Extração de RNA total

Para o isolamento de RNA total utilizou-se o protocolo modificado de LiCl por Morgante e colaboradores (2011). Usando-se as seguintes soluções: tampão NTES (NaCl 0,1M, Tris-HCl 0,01M, EDTA 1mM, SDS 1% e quantidade suficiente de H₂O DEPC 0,1% para 100 mL de tampão), Clorofane (25:24:1 – Fenol:Clorofórmio:Álcool Isoamílico), AcNa – Acetato de Sódio 3M pH 4,5, LiCl – Cloreto de Lítio 4M, H₂O DEPC (Dietilpirocarbonato) 0,1%, Etanol absoluto (100%) a -20°C e Etanol 70% a -20°C. Com exceção do SDS, clorofane e etanol, todas as outras soluções foram autoclavadas a 120°C por 20 minutos.

No primeiro dia, o tecido vegetal coletado e armazenado no refrigerador -80°C foi macerado em almofariz, previamente refrigerado com nitrogênio líquido, até obtenção de um pó fino. O material foi mantido congelado durante todo o processo de maceração. Cerca de 650mg do pó obtido com a maceração foi transferido, com o auxílio de espátula estéril refrigerada, para tubo Falcon de 15 mL, também refrigerado, o equivalente a aproximadamente 2,5 mL, voltando a ser mantido em nitrogênio líquido até a etapa inicial de extração.

Foram adicionados tampão NTES e clorofane, em quantidades equivalentes ao volume inicial de material vegetal, aos tubos, agitando-os vigorosamente, a temperatura ambiente, a fim de obter uma homogeneização da amostra, sendo logo em seguida, centrifugados a 10.000 rpm por 30 minutos à 4°C. Transferiu-se a fase superior aquosa para novo tubo Falcon de 15 mL. Centrifugou-se a 10.000 rpm por 30 minutos à 4°C, para retirada de qualquer resquícios de contaminantes. Transferiu-se, cuidadosamente, a fase aquosa para novo tubo Falcon de 15 mL. A partir desse ponto, todas as amostras foram constantemente mantidas em gelo. Ao volume obtido desta fase aquosa foi adicionado 1/10 de acetato de sódio e 2,5 vezes de etanol absoluto gelado. As amostras foram agitadas manualmente, tentando homogeneizá-las, e incubadas a -20°C por 2 a 4 horas. Em seguida, as amostras foram centrifugadas a 10.000 rpm por 20 minutos à 4°C. O sobrenadante foi descartado por inversão do tubo. O *pellet* (precipitado) obtido foi dissolvido em 1 mL de H₂O DEPC 0,1%, transferido para microtubo de polietileno de 1,5 mL e centrifugado a 5.000 rpm por 3 minutos à 4°C. Transferiu-se o sobrenadante, por inversão, para microtubo de polietileno de 2,0 mL e adicionou-se o mesmo volume de LiCl, sendo incubadas durante toda a noite (*overnight*) a 4°C. No segundo dia, as amostras precipitadas foram centrifugadas a 10.000 rpm por 20 minutos à 4°C. Descartou-se o sobrenadante e o *pellet* foi ressuspendido em 500 µL de H₂O DEPC.

Adicionou-se 1/10 de acetato de sódio e 2,5 vezes de etanol absoluto gelado ao volume ressuspendido em H₂O DEPC. Os tubos foram agitados manualmente e incubados a -20°C por 2 horas. Em seguida, centrifugou-se a 10.000 rpm por 20 minutos à 4°C. Descartou-se o sobrenadante, e lavou-se o *pellet* com 500 µL de etanol 70% gelado. Centrifugou-se as amostras a 10.000 rpm por 5 minutos à 4°C, descartando o sobrenadante. Secou-se o *pellet* à temperatura ambiente, e o RNA foi ressuspendido em 30 µL de H₂O DEPC 0,1%.

4.3. Análise qualitativa e quantitativa do RNA total e preparação das amostras para sequenciamento

A avaliação qualitativa de integridade do RNA total extraído foi conduzida por meio da análise de eletroforese em gel de agarose 1,5%, não desnaturante, corado com brometo de etídio, com tampão de corrida TAE (Tris-Acetato-EDTA) 1% e voltagem de 80V, por aproximadamente 30 minutos.

A avaliação quantitativa deu-se pela quantificação do RNA total a 260nm, utilizando o espectrofotômetro Nanodrop ND-1000 (Thermo Scientific). Sua pureza foi

confirmada com a relação 260/280nm acima de 1,8 U.A. (Unidade de Absorbância). Após a quantificação do RNA total de cada planta extraída individualmente, formou-se dois *pools*, de cada ponto e para cada grupo, denominados A e B (Figura 17). Para isso, juntou-se o material extraído de diferentes plantas do mesmo grupo (controle ou inoculado) e ponto (24, 48 e 72hs), com perfis de rendimento próximos. Tomou-se o cuidado para que cada *pool* contivesse no máximo 150µg de RNA total.

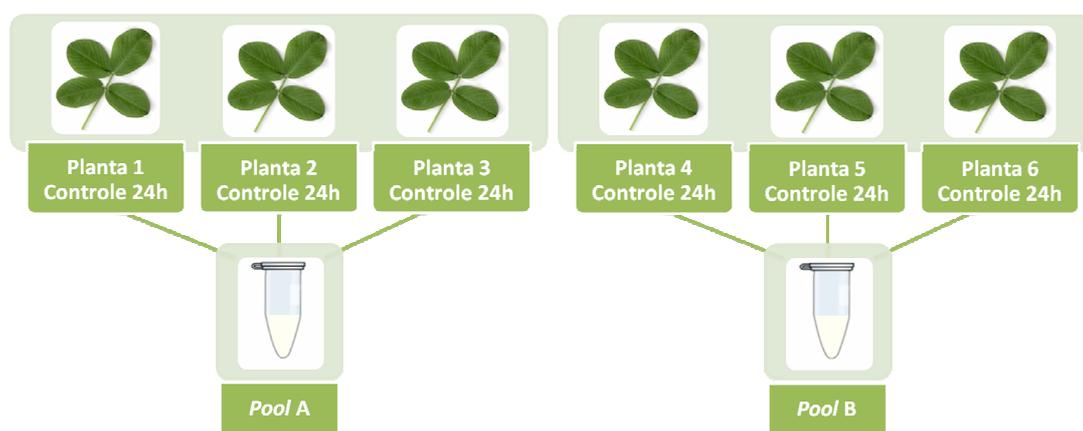


Figura 17 – Esquema ilustrativo da formação dos dois *pools* (A e B), por meio da junção equitativa do RNA total de diferentes plantas extraídas individualmente, do mesmo ponto de coleta (24, 48 ou 72h) e mesmo grupo (controle ou inoculado).

Estes materiais foram então purificados em colunas do Kit Invisorb® Spin Plant RNA Mini (Invitek), adicionando-se 900 µL da solução de lise DCT, agitando-os manualmente. As amostras foram centrifugadas na velocidade máxima (13.200 rpm) por 1 minuto, a temperatura ambiente, formando um *pellet* gelatinoso e transparente. O sobrenadante transparente foi transferido, por inversão do microtubo, para o pré-filtro contido em um tubo receptor de 2,0 mL (tubo e filtro verdes), os quais foram centrifugados a 10.000 rpm por 1 minuto. Descartou-se o pré-filtro. Ao filtrado, aproximadamente 800 µL contendo o RNA, foi adicionado, misturando completamente por pipetagem, 500 µL de etanol absoluto gelado. Transferiu-se 750 µL desse filtrado para o filtro RTA *spin* (filtro amarelo), incubando-o por 1 minuto a temperatura ambiente, sendo logo em seguida, centrifugado a 10.000 rpm por 1 minuto. Descartou-se o filtrado, e retornou o filtro RTA *spin* ao seu tubo receptor RTA (tubo amarelo). Repetiu-se essa etapa, transferindo o volume residual do filtrado, para o filtro RTA *spin*, centrifugando novamente por 1 minuto e descartando o filtrado. Ao filtro RTA *spin*, adicionou-se 500 µL do tampão de lavagem R1. Centrifugou-se a 10.000 rpm por 30 segundos. Descartou-se o filtrado. Ao filtro RTA *spin* adicionou-se 700 µL do tampão de lavagem R2 e centrifugou-se a 10.000 rpm por 30 segundos, descartando o filtrado. Esta etapa de lavagem com o

tampão R2 foi repetida mais uma vez. Para eliminar qualquer traço de etanol, as amostras foram centrifugadas a 12.000 rpm por 3 minutos. O filtro RTA *spin* foi transferido para o tubo de eluição livre de RNase, onde foram adicionados, diretamente na membrana do filtro RTA *spin*, de 30-60 μ L (dependendo do volume de RNA inicial usado) do tampão de eluição R. As amostras foram incubadas por 2 minutos a temperatura ambiente, sendo logo em seguida, centrifugadas a 10.000 rpm por 1 minuto e adicionadas imediatamente ao gelo, após o descarte dos filtros RTA *spin*.

A qualidade e quantidade das amostras de RNA foram novamente avaliadas conforme descrito acima. Em seguida, um novo *pool* contendo 100 μ g, de forma equitativa, de todas as amostras dos *pools* A e B do grupo controle (24, 48 e 72hs) foi formado (Figura 18). O mesmo foi feito com as amostras dos *pools* A e B do grupo inoculado. A estes dois novos *pools* (SFC – *Stenosperma* Folha Controle e SFI – *Stenosperma* Folha Inoculada) foram adicionados 2,5 vezes (volume inicial) de etanol absoluto. Esse material foi acondicionado em refrigerador -80 e transportado para a empresa de sequenciamento massal, CD-Genomics, Nova Iorque – EUA, em gelo seco. O restante do RNA total purificado foi armazenado em refrigerador -20 $^{\circ}$ C para posterior uso na validação de genes-candidatos identificados pelo sequenciamento.

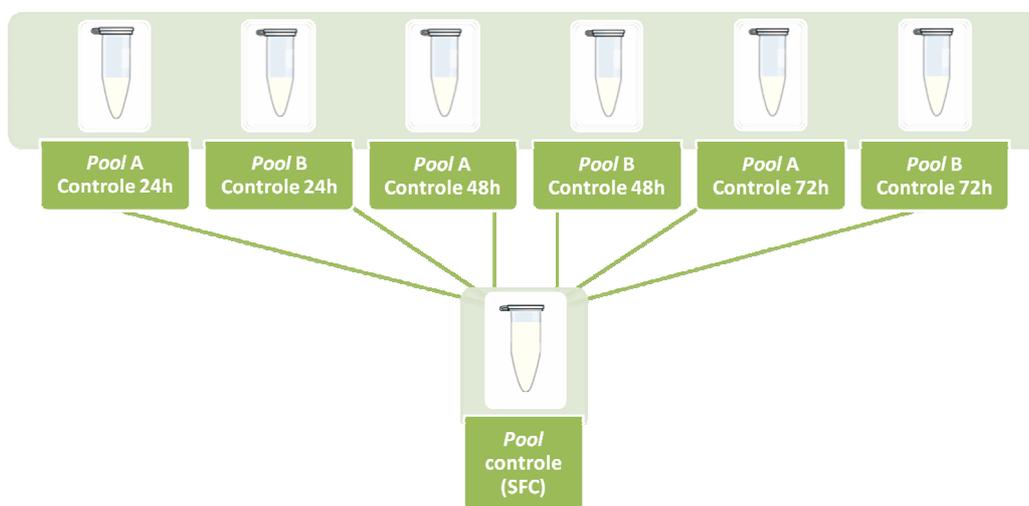


Figura 18 – Esquema ilustrativo da formação de um *pool* controle e outro inoculado, contendo, de forma equitativa, as amostras dos *pools* A e B, de todos os pontos de coleta (24, 48 ou 72h), adicionados de etanol absoluto, para envio a empresa de sequenciamento massal (pirosequenciamento). SFC – *A. stenosperma* Folha Controle.

4.4. Pirosequenciamento massal

Todo o processo de construção de bibliotecas de cDNAs e o próprio sequenciamento 454 foi realizado pela empresa CD Genomics. (<http://www.cd->

genomics.com). O ds-cDNA (DNA complementar de dupla fita) foi sintetizado usando o kit Creator™ SMART™ cDNA Library Construction (Clontech) e a biblioteca de DNA, a PCR em emulsão (emPCR) e o sequenciamento foi realizado utilizando o GS-FLX Titanium Fragment Series Kits, da empresa Roche/454 (<http://www.my454.com>), de acordo com o protocolo da Roche, a partir dos dois *pools* (SFC e SFI) de RNA total enviados à empresa.

Esse protocolo de preparação elimina a necessidade de clonagem e repicagem de colônias, portanto, a primeira etapa, de preparação das bibliotecas de DNA/cDNA, consiste em algumas etapas simples, individualizadas por amostra, onde para isso, utilizou-se o método de síntese de cDNA SMART que utiliza características específicas da transcriptase reversa (RT) do vírus da leucemia murina de Moloney para adicionar dNTPs, principalmente citosina, sem molde, à extremidade 3' de uma fita recém sintetizada de cDNA, ao chegar na extremidade 5' do RNA molde (Schmidt & Mueller, 1999). Quando adicionados a reação de RT, oligonucleotídeos contendo uma sequência de oligo-riboG na extremidade 3' (também chamados TS-oligo - *Template-switch oligonucleotide*) formam pares com o trecho de desoxicitidina produzido pela RT. A RT, em seguida, continua a replicação usando o TS-oligo como um molde, levando a adição 3' da sequência TS complementar, servindo como sítio terminal 5' para pareamento do *primer* para obtenção de amostras amplificadas de sequências completas de cDNA enriquecido (Figura 19) (Matz, et al., 1999). De acordo com Zhu e colaboradores (2001), a tecnologia SMART (*Switching mechanism at the 5' end of the RNA transcript*) de síntese de cDNA é bem adequada para geração de bibliotecas de cDNA para diversas aplicações.

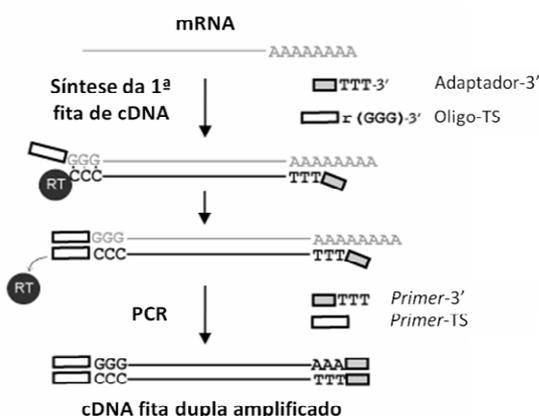


Figura 19 – Esboço esquemático da síntese de cDNA SMART. Imagem adaptada (Evrogen, s/d).

Para amostras de ácidos nucleicos com tamanho entre 70-800 pb, não se faz necessária a etapa de nebulização (fragmentação mecânica). Portanto, todos os produtos de PCR passaram, diretamente, para a etapa de avaliação de qualidade, para serem analisados no sistema Agilent 2100 Bioanalyser (Capiler, USA), em chip de DNA - *DNA 1000 LabChip* (Agilent Technologies), verificando se os fragmentos obtidos estavam entre o intervalo determinado de 70-800pb. Os fragmentos abruptos são submetidos ao procedimento de polimento e desfosforilação de suas extremidades, para ligação dos adaptadores necessários para a PCR em emulsão (emPCR) e o pirosequenciamento (Roche, 2009). Adaptadores com sequências específicas, denominados “A” (contendo a sequência complementar, para pareamento do *primer* contido nas microesferas) e “B” (contendo biotina), são ligados às extremidades 3’ e 5’, respectivamente, dos fragmentos, formando quatro tipos de produtos: a) fragmentos com adaptadores “AB”, nas extremidades 3’ e 5’, respectivamente; b) fragmentos com adaptadores “BA”, nas extremidades 3’ e 5’, respectivamente; c) fragmentos apenas com adaptador “B” (BB) e d) fragmentos apenas com adaptador “A” (AA). Contudo, apenas os fragmentos “AB” são eluídos na reação de purificação, onde a extremidade 5’, contendo o adaptador “B” (hidrofílico), é imobilizada nas microesferas revestidas de estreptavidina, que possui alta afinidade a biotina, enquanto que a extremidade 3’, contendo o adaptador “A” (hidrofóbico), é mantida expostas a superfície das microesferas. Os *gaps* presentes nos fragmentos de DNA são reparados pela DNA polimerase. Os ds-DNAs são, então, desnaturados pela aplicação de tampões específicos. E os fragmentos moldes (*templates*) de DNA fita simples (ssDNA), contendo os adaptadores “A” e “B” em cada uma de suas extremidades 3’ e 5’, respectivamente, são, portanto, os fragmentos selecionados para sequenciamento (Figura 20A) (Margulies, et al., 2005; Roche A. S., 2006; Carvalho & Silva, 2010).

Logo depois, os fragmentos de ssDNA selecionados pelo procedimento de purificação são ligados a microesferas magnéticas, por meio do pareamento de milhares de oligômeros existente, com a sequência complementar e exata ao adaptador “A”, presentes em sua superfície. Apenas um único tipo de fragmento se liga a uma determinada microesfera (Figura 20B). Logo em seguida, adiciona-se às microesferas magnéticas, que já possuem polimerase, uma emulsão contendo uma mistura de água, óleo e reagentes para PCR. Essas microesferas são, então, capturas, individualmente, em gotículas oleosas, onde a emPCR – *in vitro* ocorre, parecendo microrreatores para amplificação clonal, produzindo assim, muitas cópias idênticas de um mesmo fragmento alvo, isoladamente, em um microsuporte (Figura 20C) (Margulies, et al., 2005; Ellegren, 2008; Mardis, 2008a; Mardis, 2008b; Carvalho & Silva, 2010).

Para o sequenciamento em si, as microesferas magnéticas, com os fragmentos amplificados ligados a ela, são liberadas das micelas (gotículas de óleo) e carregadas em uma placa de picotitulação de chip de fibra ótica, com uma mistura de incubação, contendo DNA polimerase e co-fator. Nanoesferas de enzimas (contendo sulfúrilase e luciferase) são, também, depositadas por toda a placa, envolvendo por completo as microesferas, garantindo, portanto, que elas permaneçam posicionadas nos poços durante a reação de sequenciamento, e permitindo que uma única microesfera esteja contida em cada poço de $\approx 44\mu\text{m}$. Esse procedimento faz com que as reações de sequenciamento ocorram em poços distintos, contendo uma única microesfera, com um único tipo de fragmento ligado a ela, evitando a competição por reagentes com os outros fragmentos da biblioteca e propiciando o sequenciamento paralelo dos 1,6 milhões de poços (Ellegren, 2008; Mardis, 2008a; Carvalho & Silva, 2010).

Com isso, ocorre então o sequenciamento baseado na síntese por pirosequenciamento, onde a cada ciclo, após um determinado nucleotídeo ser incorporado à síntese, um sinal de luz é emitido, e capturado pela câmera acoplada ao sistema (CCD). A intensidade do sinal capturado é o reflexo do número de nucleotídeos específicos, incorporados, sucessivamente, na molécula. O sinal capturado é plotado em um pirograma, onde a altura de cada sinal é proporcional à intensidade do sinal (Figura 20D). No entanto, a intensidade do sinal para homopolímeros se estende até oito nucleotídeos consecutivos, após isso, o sinal cai rapidamente.

As leituras geradas são processadas pelo software de análise do 454, sendo em seguida, analisadas por vários filtros de qualidade para removerem as sequências de má qualidade, sequências mistas ou distintas (mais de um fragmento de DNA inicial), advindas da mesma microesfera, bem como as leituras sem as iniciais TCGA (correspondente aos quatro primeiros nucleotídeos do adaptador “A”).

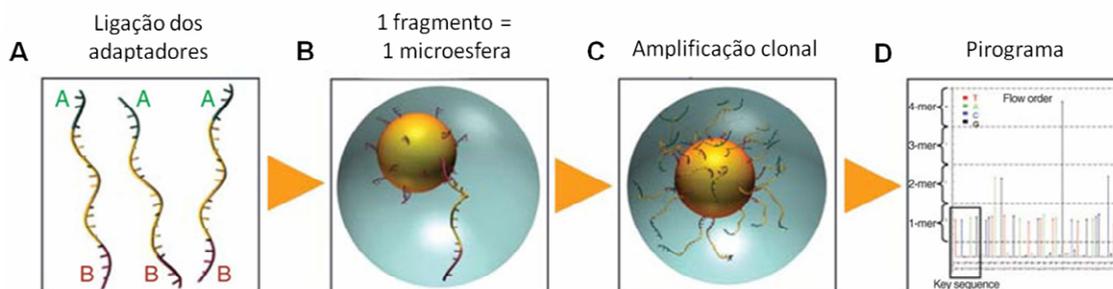


Figura 20 – Esquema ilustrativo das diferentes etapas de sequenciamento 454 (Ellegren, 2008): A) DNA fragmento utilizado como molde, com os adaptadores A e B ligados as suas extremidades 3' e 5', respectivamente. B) Fragmentos de DNA immobilizados por uma sequência alvo de biotina (adaptador B), que se liga às microesferas cobertas com estreptavidina. C) As microesferas são emulsionadas em água e óleo, contendo os reagentes necessários para reação de PCR, funcionando, como microrreator para amplificação clonal. Após a deposição dessas microesferas em placa de

picotitulação, dá-se o início do pirosequenciamento, realizado por meio do fluxo sequencial de reagentes sobre a placa. D) A intensidade do sinal captado por meio da incorporação de um novo nucleotídeo a nova fita, é gerado um sinal luminoso, representado em um pirograma. A altura de cada sinal é proporcional ao número de nucleotídeos idênticos incorporados consecutivamente.

4.5. Análise bioinformática dos dados

Os dados gerados pelo pirosequenciamento massal em 454 foram submetidos a um pré-processamento, que incluiu a leitura dos arquivos *.sff (formato de arquivo binário usado para codificar resultados de pirosequenciamento) fornecidos pela CD Genomics, seguido da extração das informações deste arquivo e conversão em formato FASTA/Phred-compatível, com a identificação de pequenas sequências de adaptadores com o software SSAHA2 (*Sequence Search and Alignment by Hashing Algorithm* – <http://www.sanger.ac.uk/resources/software/ssaha2/>), bem como a identificação e mascaramento de sequências indesejadas com o auxílio do programa Blast2GO (*Basic Local Alignment Search Tool*), além do mascaramento de sequências repetitivas, realizada com o software *RepeatMasker* (<http://www.repeatmasker.org>) e remoção da poliA/T.

Logo em seguida, realizou-se a avaliação de cobertura *k-mer* mais freqüentes, de forma empírica. O conjunto de sequências geradas por esse método foi submetido ao procedimento de montagem com o software Mira (<http://sourceforge.net/projects/mira-assembler/files/>). Uma nova inspeção com o programa Blast2GO (<http://www.blast2go.com/b2ghome>) foi realizada para conciliar as montagens (unigenes) obtidas anteriormente para cada um dos conjuntos de sequências visando eliminar possíveis redundâncias de agrupamentos de sequências (Figura 21). Após a obtenção desses agrupamentos (*clusters*) a anotação foi realizada com o uso dos programas Blast2GO, com o programa BLASTx para obtenção de proteínas/genes conhecidos e o com o InterproScan, para anotação de domínios protéicos com funções celulares específicas.

Uma análise de expressão diferencial foi conduzida mediante o emprego da técnica de DDD (*Digital Differential Display*). Essa técnica faz uso do método estatístico de comparação – o teste de Audic-Claverie (1997), para identificação de diferenças significativas de alteração da expressão gênica a partir da diferença do número de sequências de cada biblioteca comparada presentes em um agrupamento particular. Testes estatísticos foram realizados com o software IDEG6 (*Identification Differentially Expressed Genes*) confirmando os genes diferencialmente expressos indicados pela subtração *in silico*.

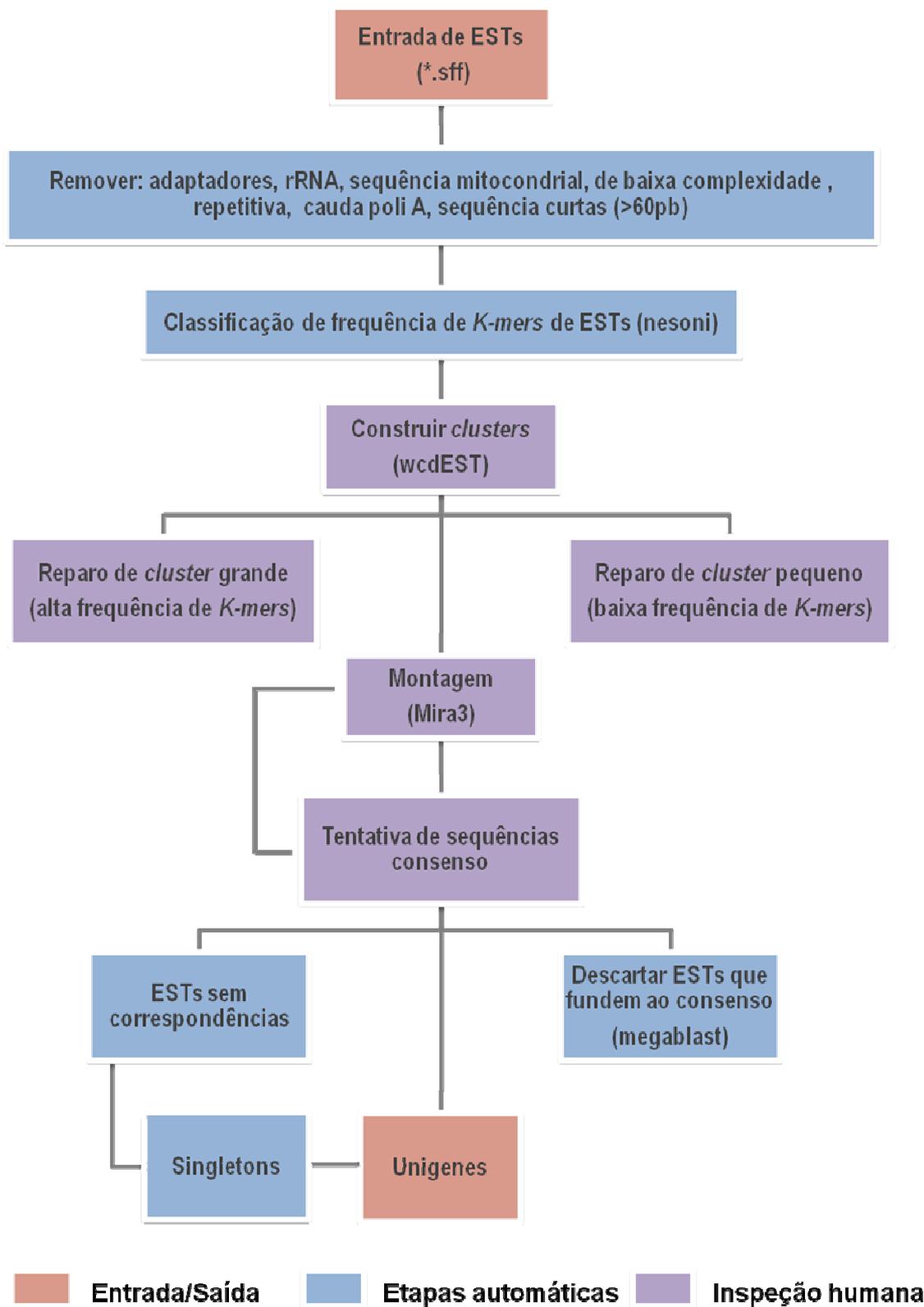


Figura 21 – Fluxograma utilizado para análise *in silico* dos dados gerados pelo sequenciamento 454.

4.6. Avaliação dos unigenes para seleção de genes-candidatos

A partir da análise da expressão diferencial e da anotação em BLASTx (Tabela 4 – Apêndice), foram selecionados 11 genes como candidatos. A escolha foi realizada nesta tabela pela avaliação da função putativa, a eles relacionados, de genes que possivelmente estão associados à resposta de defesa ou ao estresse biótico, já previamente, descritos em literatura.

A seleção dos cinco RGAs candidatos foi realizada por meio da busca por similaridade das sequências das duas bibliotecas (SFI e SFC) obtidas com RGAs NBS-LRR de *Arabidopsis thaliana* (L.) Heynh, utilizando-se o programa BLASTx e o software BioLayout Express^{3D} (<http://www.biolayout.org/>).

4.7. Validação dos genes-candidatos por RT-qPCR

4.7.1. Desenhos dos *primers* para os genes-candidatos selecionados

A partir das sequências dos 11 genes-candidatos selecionados, realizou-se um BLASTn contra o banco de ESTs, o qual foi limitado a sequências do gênero *Arachis* (<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi?PAGE=Nucleotides>), visando a verificação de sequências similares existentes. Já para os cinco RGAs candidatos, por estarem usualmente presentes em regiões conservadas do genoma, tentou-se identificar o(s) domínio(s) conservado(s) presente(s) no *contig* avaliado, e a similaridade entre as proteínas depositadas no banco de sequências de proteínas não-redundantes (nr), por meio do BLASTx. Esses domínios foram identificados pelo banco de dados *Conserved Domain Database* v2.32 (CDD – <http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>), pelo alinhamento das sequências pesquisadas com as depositadas neste banco. As sequências com similaridade ao domínio pesquisado, e que não fazem parte do domínio conservado, foram selecionadas para desenho dos *primers* (iniciadores) gene-específicos.

Com o auxílio do programa *Gene Runner* versão 3.05 (Hastings Software, Inc., disponível em <http://www.generunner.net>), o intervalo de bases similares localizado entre as sequências resultantes dos BLAST acima citados, foi selecionado e, a partir dele, pelo programa *Primer3Plus* (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>), foram desenhados os *primers* a serem utilizados na validação desses genes por RT-

qPCR. Os parâmetros utilizados neste programa para desenho dos *primers* foram: fragmento amplificado (*amplicon*) com tamanho entre 150-200pb, *primers* com tamanho entre 19-22pb, com temperatura de dissociação (T_m) entre 59-61°C e porcentagem de bases GC no *primer* entre 45-55%.

4.7.2. Preparação das amostras para validação dos genes-alvo em RT-qPCR

Os *pools* A e B de RNA total, formados e purificados em colunas do Kit Invisorb® para cada um dos três pontos coletados e para cada grupo (controle e inoculado), como descrito no item 4.3, armazenados a -20°C, foram utilizados como molde para síntese de cDNA utilizado na validação da expressão gênica por qPCR, a seguir.

Tratamento das amostras com DNase e síntese de cDNA

Para eliminação de eventuais contaminações das amostras de RNA total com DNA genômico, as amostras foram tratadas com desoxirribonuclease (DNase). Para isso, 2µg de RNA total de cada *pool* (Figura 17), individualizados por ponto e tratamento, foi tratado com 2U de TURBO™ DNase (Ambion® - Applied Biosystem™), em uma reação com 10 µL de volume final, a qual foi incubada a 37°C por uma hora, seguida da desativação da enzima, com 2,5mM de EDTA por 10 minutos a 65°C.

Após o tratamento com a DNase, cada amostra de RNA total foi diretamente usada para a síntese de cDNA. Primeiramente, para desestabilizar a estrutura secundária do RNA, adicionou-se um *mix*, contendo 1 µL de Anchored Oligo (dT)₂₀ 50mM (Invitrogen™), 1 µL de dNTP 10mM ao volume final de cada amostra tratada com DNase. Incubaram-se as amostras a 65°C por 5 minutos e depois no gelo por mais 5 minutos. Em seguida, adicionou-se para cada amostra o *mix* para transcrição reversa, contendo 4 µL do tampão 5X FS (*First Strand Buffer*) da enzima, 2 µL de DTT 0,1M e 1 µL da enzima transcriptase reversa SuperScript II (Invitrogen™), e incubou-se a 42°C por 60 minutos e por último, a 70°C por 15 minutos, para desativação da enzima.

Para conferir a ausência de contaminação por DNA genômico, nas amostras de cDNA sintetizadas, realizou-se uma PCR convencional, utilizando-se um par de *primers* leg066, leg066_*Forward* (5'AGCTCCACCTCTTTCCGACAGA³) e leg066_*Reverse* (5'AGTTTCTACAGCACGTATCCTTTCC³), que flanqueiam uma região intrônica

conservada em *Arachis* (Hougaard, et al., 2008; Morgante, et al., 2011) e permite a distinção entre os produtos de PCR amplificados a partir de um DNA genômico ou de um cDNA (Morgante, et al., 2011). A reação de PCR foi realizada em um volume final de 13,5 μL , contendo tampão 1x, MgCl_2 1,44mM, dNTP 2,44mM, BSA 0,5mg/mL, 0,15 μM de cada *primer*, 1,5U da enzima Taq DNA Polimerase Recombinante (Invitrogen™®) e 1 μL de cDNA. Esta reação foi realizada no termociclador Master Cycler (Eppendorf), sendo submetida a 94°C por 5 minutos, para uma desnaturação inicial, seguido de 30 ciclos repetitivos de amplificação: 94°C por 1 minuto, 56°C por 1 minuto e 72°C por 1 minuto, e a um ciclo final de extensão a 72°C por 5 minutos. O resultado da PCR foi analisado por eletroforese em gel de agarose 1,5%, não desnaturante, corado com brometo de etídio.

4.7.3. PCR quantitativa (RT-qPCR)

A reação de RT-qPCR (PCR quantitativa precedida de transcrição reversa utilizando o cDNA como molde) foi realizada no F7300 Real Time PCR System (Applied Biosystem) utilizando-se o reagente Platinum® SYBR® Green RT-qPCR SuperMix-UDG w/ROX (Invitrogen™), composto por um *mix* que contém, com exceção dos *primers* e da fita molde, todos os outros componentes necessários para amplificação e detecção do DNA por qPCR (PCR quantitativa). Em um volume final de 10 μL , esta reação foi preparada com 5 μL do SuperMix SYBR Green, 0,2 μM de cada *primer* e 2 μL do cDNA diluído a 10^{-1} . Esta diluição foi determinada a partir da análise de uma curva de diluição, com a qual se pode estabelecer parâmetros como a eficiência dos *primers*, que deve estar entre 90 e 110%, o coeficiente de correlação (R^2), a qualidade do *primer*, fornecida pela curva de dissociação, e a diluição ideal de cDNA a ser utilizada no RT-qPCR. Para o estabelecimento da curva de diluição, um *pool* formado por amostras dos *pools* A e B, contendo quantidades iguais de cDNA dos três pontos coletados (24, 48, e 72hs) dos dois grupos, controle e inoculado (Figura 22). Esse *pool* foi diluído a 10^{-1} , 10^{-2} , 10^{-3} , e a diluição ideal foi determinada para obtenção de C_q com valores entre 18 e 25. As reações de RT-qPCR foram realizadas em triplicatas (três repetições técnicas da mesma amostra) para cada amostra testada, e, como controle negativo dessas reações, utilizou-se para cada *primer* testado, o NTC (*No template Control*) composto pelo mesmo *mix* de reação utilizado nas demais amostras, porém sem a adição da fita molde, isto é, sem o cDNA. Além disso, a fim de determinar a ausência de contaminação com DNA genômico, nas amostras de RNA total usadas para síntese de cDNA, realizou-se também o controle negativo NAC (*No Amplification Control*), onde ao *mix* da reação, foi adicionado o RNA total. Os ciclos da RT-qPCR consistiram em quatro etapas: 50°C por 2 minutos, 95°C por

10 minutos, 40 ciclos de 95°C por 15 segundos e 60° C por 1 minuto, e por fim, uma curva de dissociação de 95°C por 15 segundos, 60°C por 60 segundos e 95°C por 15 segundos. As eficiências de amplificação e os valores de coeficiente de correlação (R^2) foram calculados pelo método da curva-padrão, avaliando os C_q das triplicatas das três diluições utilizadas (10^{-1} , 10^{-2} e 10^{-3}) para cada *primer*.

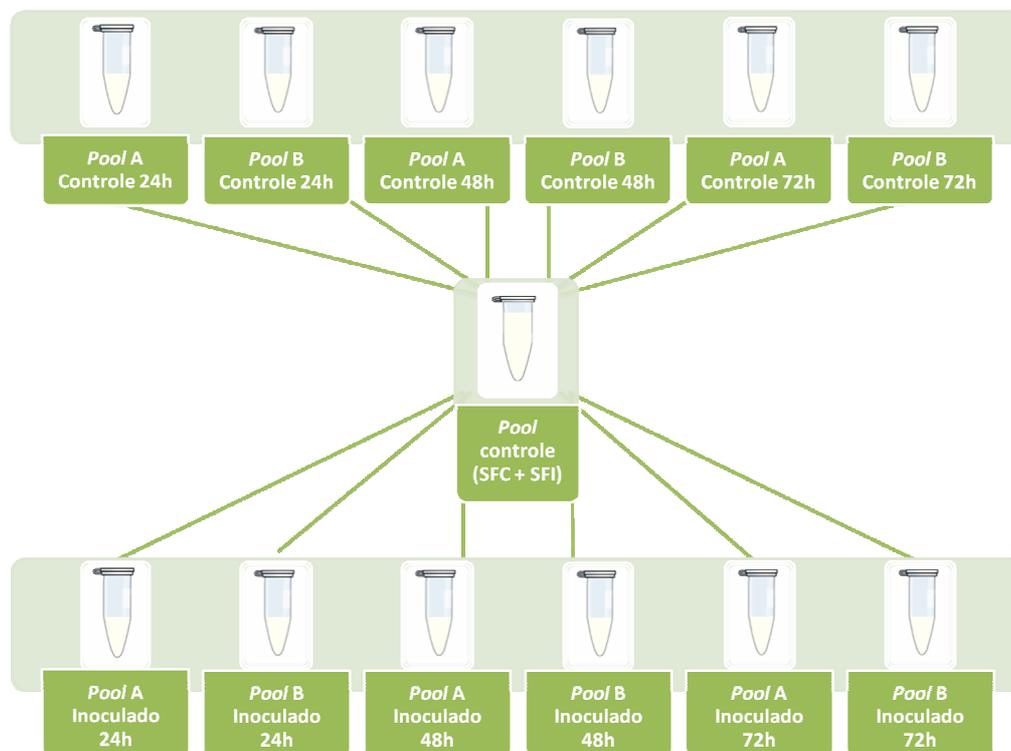


Figura 22 – Esquema ilustrativo da formação de um único *pool* contendo amostras dos *pools* A e B de todos os pontos (24, 48 e 72h) e dos dois grupos trabalhados, controle e inoculado, para análise da curva de diluição pelo RT-qPCR. SFC – *A. stenosperma* Folha Controle; SFI – *A. stenosperma* Folha Inoculada.

Análise da expressão dos genes-candidatos selecionados

A análise da expressão dos genes-candidatos (genes-alvo) selecionados foi realizada, separadamente, isto é, por ponto e por grupo, para cada amostra de cDNA, em triplicatas, todas com fator de diluição 10^{-1} . Para essa análise, a quantificação da expressão gênica foi baseada na expressão de um gene-alvo em relação a um gene de referência, previamente selecionado (Morgante, et al., 2011). Um gene de referência ideal deve ter um nível constante de expressão, no tecido escolhido, não sendo afetado por nenhum tipo de tratamento ou por alguma variabilidade inter-individual existente. Este gene é utilizado como controle endógeno no intuito de normalizar os níveis de mRNA entre as amostras testadas.

Com base em estudos realizados por Morgante e colaboradores (2011), acerca dos genes de referência para estudo de expressão gênica em organismos do gênero *Arachis*, os genes mais indicados para essa análise, baseado na espécie, órgão e tipo de estresse, estudados, foi o a *Actina* (C23_Foward – 5'TGGTCTCGGTTTCCTGAGTT^{3'} e C23_Reverse 5'AATACCACTCCAAAGCAAACG^{3'}) e a *proteína ribossomal 60S* (C43_Foward – 5'TGGAGTGAGAGGTGCATTTG^{3'} e C43_Reverse – 5'TCTTTTGACGACCAGGGAAC^{3'}).

Os dados gerados pela RT-qPCR foram analisados utilizando-se o software 7500 Real Time PCR System (Applied Biosystem) v2.0.4, compatível com o arquivo gerado pelo F7300, pelo método de quantificação de C_q comparativo. Para tanto, utilizou-se a proteína ribossomal 60S como controle endógeno. Identificou-se as amostras testadas em dois grupos biológicos a serem analisados, (grupo controle e grupo inoculado), além da indicação da amostra de referência, a qual todas as demais amostras foram comparadas, para obtenção do seu perfil de expressão. Uma análise utilizando o método C_q comparativo ou $2^{(-\Delta\Delta C_q)}$ (Livak & Schmittgen, 2001) foi realizada, em Excel, utilizando os dados de C_q gerados pela RT-qPCR, para confirmação dos resultados de expressão gênica relativa obtidos pelo 7500 Software v2.0.4 do RT-qPCR. Para que o método C_q comparativo de quantificação relativa seja validado, as eficiências de amplificação do gene-alvo e do gene de referência devem ser aproximadamente iguais.

5. RESULTADOS

5.1. Desenho experimental, inoculação e coleta das plantas

O objetivo dos ensaios de inoculação de *A. stenosperma* com o *P. personata* foi o de se obter populações de mRNA de pontos distintos da interação, ampliando a variabilidade da expressão gênica. Quatro ensaios realizados com *A. stenosperma* e *A. hypogaea* cv. IAC-Tatu foram utilizados neste trabalho (4º, 5º, 6º e 8º bioensaio). A primeira espécie citada, *A. stenosperma*, foi coletada em três pontos distintos, 24, 48 e 72h após a inoculação (HAI), escolhidos com base na cinética da infecção, para análise da expressão por meio do pirosequenciamento massal e avaliada por RT-qPCR. A espécie cultivada, conhecida pela sua suscetibilidade à mancha preta e outras doenças e pragas, serviu como controle positivo da virulência do inóculo usado, sendo possível a visualização de sintomas a partir do 24º ao 31º dia após a inoculação em placa de Petri e em casa de vegetação, respectivamente (Figura 23). Enquanto que plantas de *A. stenosperma*, não apresentaram nenhum sintoma visível, nesse mesmo período avaliado (dados não mostrados).

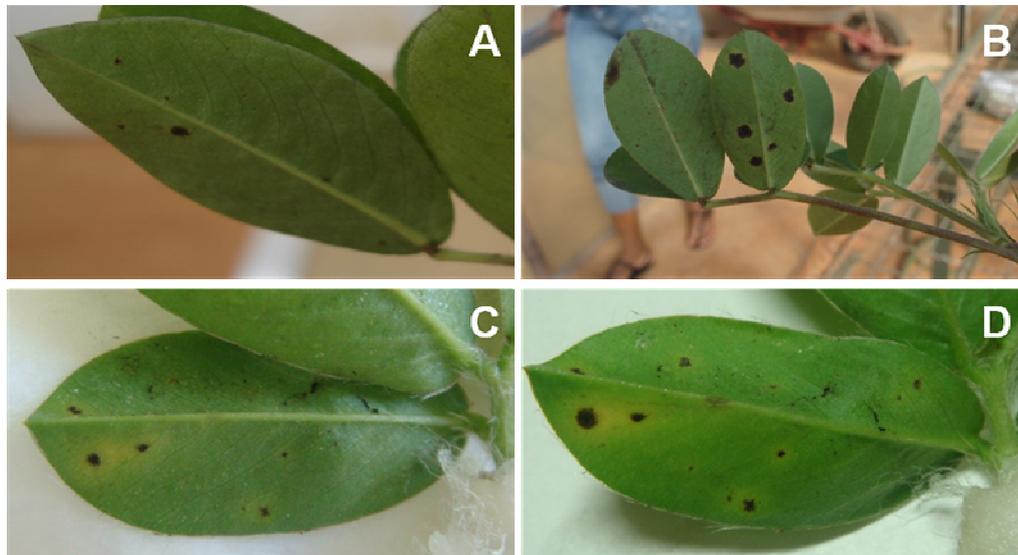


Figura 23 – Controles positivos (*A. hypogaea* cv. IAC-Tatu inoculado com *P. personata*). (A e B) Controle realizado no 6º bioensaio, com plantas mantidas durante todo o experimento até a aparição dos primeiros sintomas em casa de vegetação. (A) Sintomas registrados 31 dias após a inoculação. (B) Sintomas registrados 35 dias após inoculação. (C e D) Controle realizado no 8º bioensaio, com a técnica de folhas destacadas (Moraes & Salgado, 1982) acondicionadas em placas de Petri, mantidas em câmara de crescimento. (C) Sintomas registrados 24 dias após a inoculação dos folíolos em placa de Petri. (D) Sintomas registrados 27 dias após a inoculação.

5.2. Extração e purificação de RNA total

O isolamento de RNA de folhas de espécies de *Arachis*, incluindo *A. stenosperma*, pode resultar em pequenas quantidades de RNA total e contaminação com polissacarídeos. A aplicação do método de precipitação com LiCl, descrito no item 4.2 de Material e Métodos, permitiu o isolamento de maiores quantidades de RNA total e amostras de alta qualidade, embora com contaminação visível de DNA genômico (gDNA – Figura 24). Além disso, observou-se também a contaminação com algumas impurezas, tais como, presença de grandes quantidades de polissacarídeos, que no processo de extração podem co-precipitar com o ácido nucléico, e dos compostos fenólicos, que são facilmente oxidados, dificultando, assim, a quantificação precisa das amostras por espectrofotometria, e até mesmo inviabilizando a utilização de determinadas amostras.

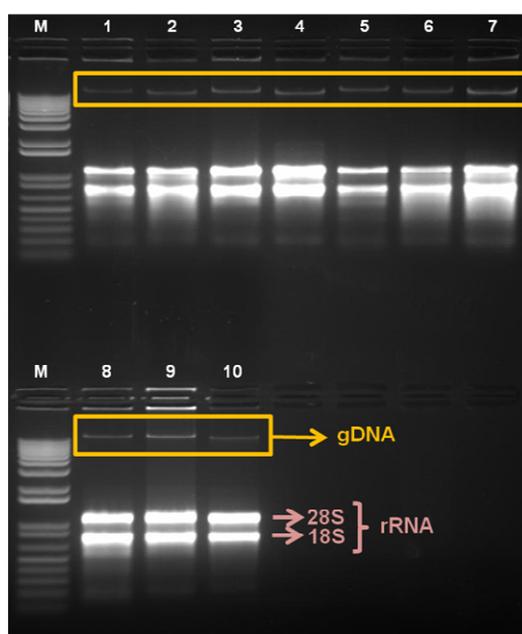


Figura 24 – Gel não desnaturante de agarose a 1,5%, corado com brometo de etídio, contendo RNA total de folha de *A. stenosperma* (1 μ L de RNA total/poço). Amostras no gel: (M) Marcador molecular (1Kb Plus – Invitrogen); (1-2) SFC 24h; (3-4) SFI 24h; (5-6) SFC 48h; (7-8) SFI 48h e (9-10) SFI 72h. Onde, SFC = *A. stenosperma*/Folha/Controle, SFI = *A. stenosperma*/Folha/Inoculada e 24, 48 e 72h refere-se aos pontos de coleta.

Visando purificar o RNA total extraído para sua utilização em RT-qPCR, utilizou-se o Kit Invisorb® Spin Plant RNA Mini (Invitek), o que provocou perda entre 30-50% da quantidade de RNA inicial, mas, gerou material de qualidade para continuidade das análises.

A qualidade da análise utilizando-se RT-qPCR depende da integridade do mRNA utilizado (Almeida & Saddi, 2007), desta forma, a avaliação da integridade do RNA é o

primeiro passo crítico na obtenção de dados de expressão gênica. Esta avaliação pode ser feita de várias formas, dentre elas, as mais usadas são o gel de eletroforese e a espectrofotometria. A presença de DNA genômico na amostra avaliada pode comprometer a medição OD 260 (densidade ótica com absorbância a 260nm) superestimando a concentração real de RNA (Fleige & Pfaffl, 2006). O resultado da purificação foi conferido tanto pela quantificação do RNA total obtido, por meio de espectrofotometria (NanoDrop – ND-1000), quanto por eletroforese em gel não desnaturante a 1,5% (Figura 25).

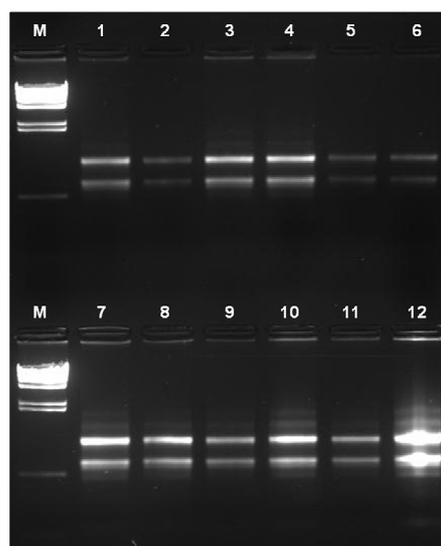


Figura 25 – Gel não desnaturante de agarose a 1,5% corado com brometo de etídio, contendo amostras combinadas de RNA total de folha de *A. stenosperma* e passadas nas colunas do Kit Invisorb® Spin Plant Mini (Invitex). Amostras no gel: (M) Marcador de peso molecular (DNA do fago Lambda digerido com Hind III); (1) SFC 24h Pool A; (2) SFC 24h Pool B; (3) SFC 48h Pool A; (4) SFC 48h Pool B; (5) SFC 72h Pool A; (6) SFC 72h Pool B; (7) SFI 24h Pool A; (8) SFI 24h Pool B; (9) SFI 48h Pool A; (10) SFI 48h Pool B; (11) SFI 72h Pool A; (12) SFI 72h Pool B. Onde SFC = *A. stenosperma*/Folha/Controle, SFI = *A. stenosperma*/Folha/Inoculada e 24, 48 e 72h refere-se aos pontos de coleta.

5.3. Análise bioinformática dos dados gerados pelo pirosequenciamento massal (sequenciamento 454)

Dois *pools* de cDNA (*pool* SFC e *pool* SFI – Figura 18) foram criados a partir do RNA total extraído de folhas de *A. stenosperma*, controle (não inoculada) e inoculadas com o *P. personata*, e seqüenciadas na CD-Genomics. Uma única corrida (placa inteira), das duas amostras em uma plataforma 454 GS FLX Titanium XLR70 (*pool* folha controle e *pool* folha inoculada), produziu um total de 362.631 leituras (*reads*), destas, 194.076 foram obtidas a partir da biblioteca de folhas de *A. stenosperma* inoculadas com *P. personata* (SFI), e 168.555 da biblioteca de folha controle (não inoculada – SFC), apresentando comprimento médio de 266 e 289 bases, em um comprimento total de,

aproximadamente, 52 e 48 Mpb, respectivamente. Após filtragem e limpeza dessas sequências, pela remoção dos adaptadores e sequências curtas, bem como pelo “mascaramento” de sequências indesejadas, tais como DNA ribossomal (rDNA), DNA mitocondrial (mtDNA) e DNA repetitivo, obteve-se 162.393 e 148.037 leituras de alta qualidade, com comprimento médio de 270 e 291 bases e mediana de 275 e 298 bases, somando um comprimento total de cerca de 42,9 e 42,8 Mpb, para as bibliotecas SFI e SFC, respectivamente (Figura 26).

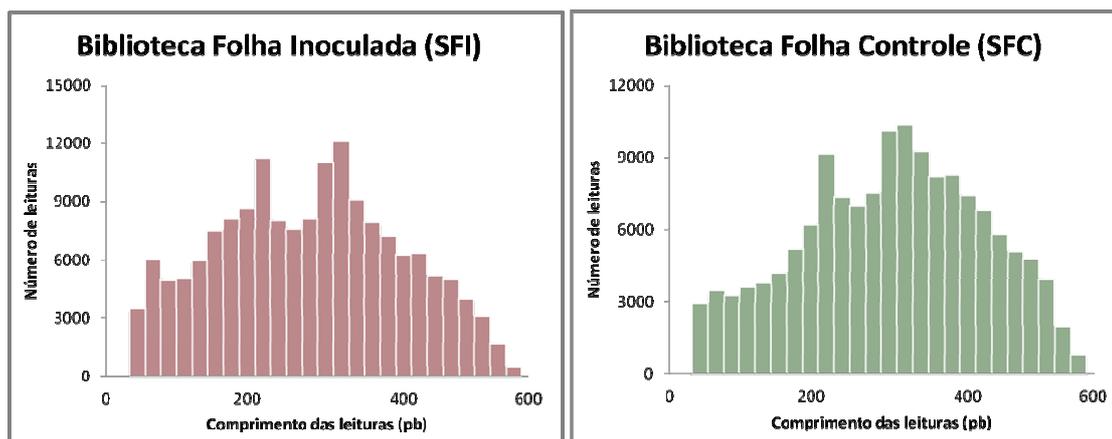


Figura 26 – Frequência dos comprimentos de leituras (pb) obtidos por meio do sequenciamento 454. À esquerda, em vermelho, dados referentes à biblioteca de folhas de *A. stenosperma* inoculadas com *P. personata*. À direita, em verde, dados referentes à biblioteca de folhas *A. stenosperma*, não submetidas à inoculação (controle).

As leituras de alta qualidade da plataforma 454 GS FLX Titanium XLR70 foram então usadas no agrupamento (“clusterização”) e montagem *de novo* das sequências usando o software MIRA3 (<http://mira-assembler.sourceforge.net>), resultando em 17.912 unigenes (*singletons* e *contigs*) das duas bibliotecas de *A. stenosperma*, onde após a eliminação dos *singletons*, foi produzido 7.723 sequências consenso de alta confiança, com cada *contig* contendo, em média, um número relativamente alto de leituras (Tabela 1). A distribuição dos *contigs* tanto pelo número de genes transcritos quanto pelo seu comprimento estão apresentados nas figuras 27 e 28. A maioria dos *contigs* foi montada a partir de 2-5 genes transcritos, com 90% dos *contigs* contendo até 30 leituras (Figura 27), com comprimento médio de 457 pb, sendo 27% dos *contigs* com mais de 500 pb (Tabela 1 e Figura 28).

A cobertura do genoma (a partir do transcrito) de *A. stenosperma* foi de 8,18 Mpb, o equivalente a 0,65% do tamanho estimado de 1.260 Mpb de uma espécie diplóide normal (Temsch & Greilhuber, 2001; Varshney, et al., 2009). Para este cálculo utilizou-se

a seguinte fórmula: $C = Unigene \times Contig$, onde C = cobertura; $Unigene$ = nº de Unigenes; $Contig$ = Comprimento médio do *contig*.

Tabela 1 – Resumo dos dados (em números) de leituras de sequências, unigenes e *contigs* obtidos pelo sequenciamento 454 GS FLX Titanium XLR70 e estimativa de cobertura genômica.

	Total
Nº de leituras de sequências	362.631
Nº de leituras após clusterização	253.134
Leituras aceitas (%)	69,8
Nº de Unigenes	17.912
Nº de <i>contigs</i>	7.723
Comprimento médio do <i>contig</i> (pb)	457
Média do nº de leituras/ <i>contig</i>	33
Cobertura genômica dos unigenes (Mpb)	8,18

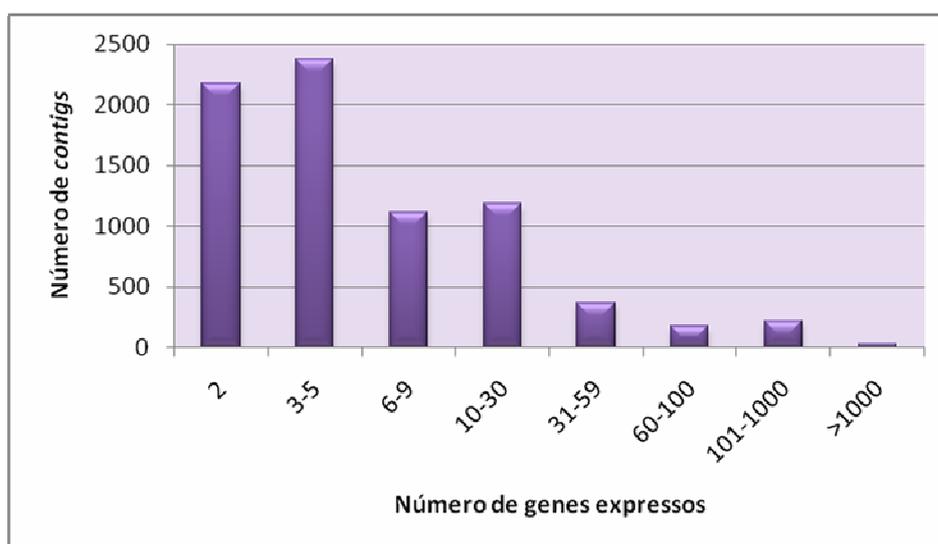


Figura 27 – Distribuição dos *contigs* por número de genes transcritos.

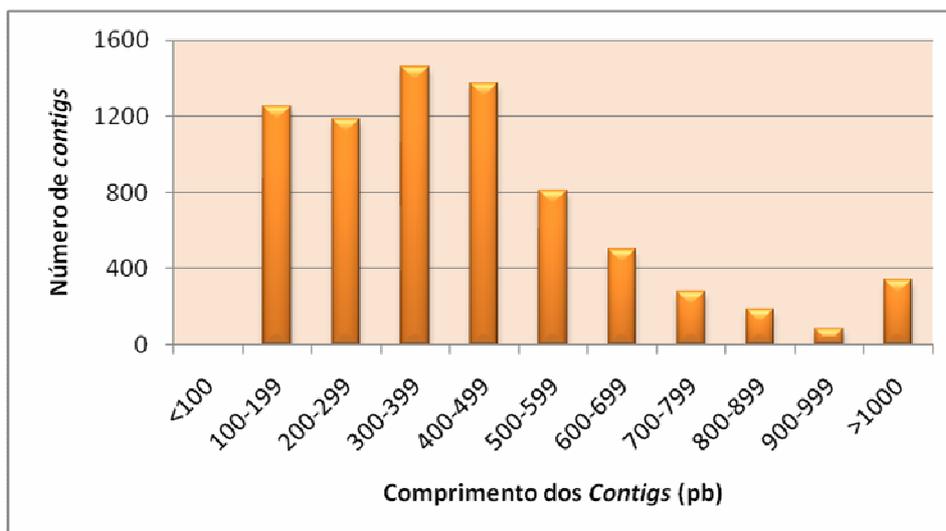
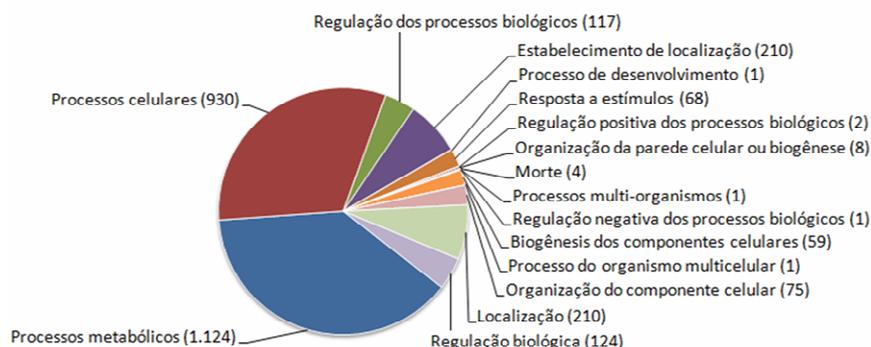


Figura 28 – Distribuição dos *contigs* por comprimento.

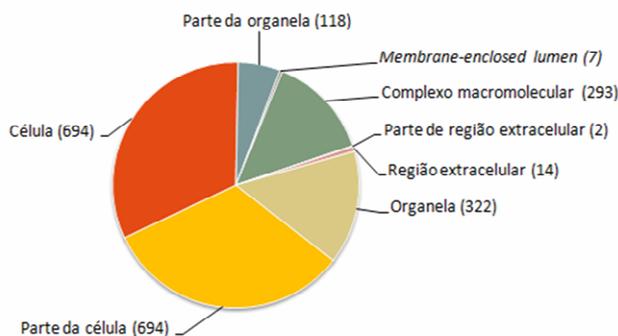
As sequências consenso de alta confiança (*contigs*) foram analisadas utilizando o BLAST2, com o algoritmo BLASTx para obtenção de proteínas/genes conhecidos (Tabela 4 – Apêndice). No total, 52% dos *contigs* poderiam ser atribuídos a supostos genes ortólogos envolvidos em várias vias e processos celulares. No entanto, devido ao escasso número de ESTs disponíveis nos bancos de dados públicos para *A. stenosperma*, aproximadamente, 32% dos transcritos mostraram similaridade com proteínas descaracterizadas, como aquelas anotadas como proteínas desconhecidas ou hipotéticas.

Para anotação funcional foi aplicado o InterProScan, para a classificação dos *contigs* em níveis de superfamília, família e subfamília, para prever a ocorrência de domínios funcionais, repetições e sítios importantes, e incluir a anotação de termos de ontologia gênica (GO – *Gene Ontology*) de assinaturas protéicas. Dos 7.723 *contigs* de *A. stenosperma*, cerca de 96% (7.396) poderia ser atribuído a uma ou mais categoria de anotação GO, sendo 2.935 (38%) atribuídos a um processo biológico, 2.144 (27,76%) a um componente celular e 2.317 (30%) a uma função molecular (Figura 29).

(A) Processos biológicos



(B) Componentes celulares



(C) Funções moleculares

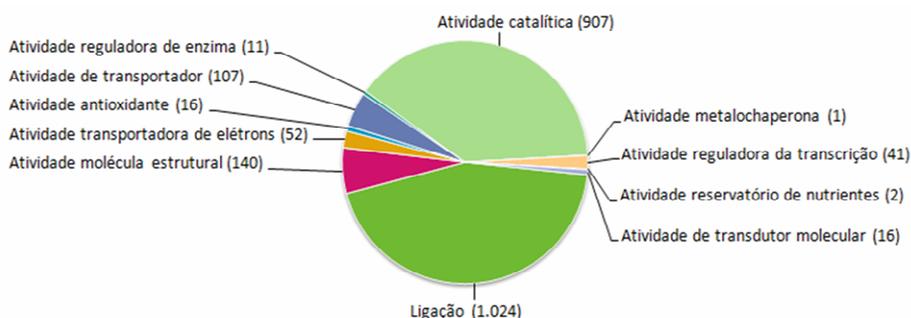


Figura 29 – Análises de InterProScan e de anotação dos termos de ontologia gênica (GO) dos *contigs* de *A. stenosperma*. (A) Número de termos de processo biológico; (B) Número de termos de componentes celulares; (C) Números de termos de função molecular.

Além disso, 68 *contigs* foram identificados na subcategoria de resposta a estímulos da ontologia gênica, o que incluiu peroxidases, catalases, quinases, glicosinases e serina-treonina quinase.

Os fatores de transcrição (FTs) constituíam menos de 1% do total das sequências consenso de confiança, e foram classificados em famílias TRF por comparação de sequência de famílias de gene de fatores de transcrição conhecidos. Para identificação das famílias de fatores de transcrição representados neste estudo, o conjunto de dados

do unigene de *A. stenosperma* foi pesquisado contra todas as sequências de proteínas de fatores de transcrição do banco de dados de fatores de transcrição de plantas (PInTFDB – *Plant transcription factor database*; <http://plntfdb.bio.uni-potsdam.de/v3.0/downloads.php>) usando BLASTx com um *e-value* de corte de 1 e-205.

Os 50 fatores de transcrição identificados foram classificados em 20 famílias, sendo os domínios zíper de leucina básico (bZIP – *basic-domain leucine-zipper*), MYB e APETALA2-Elemento de ligação a proteína responsiva ao etileno (AP2-EREBP – *Ethylene-responsive element-binding protein*), os mais representados (Figura 30).

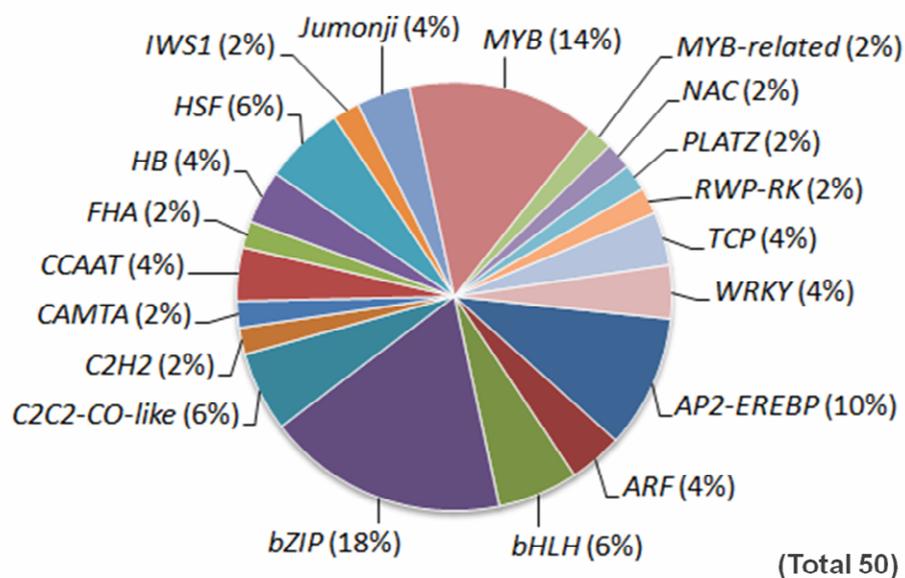


Figura 30 – Distribuição de *contigs* por famílias de fatores de transcrição em *A. stenosperma*.

Genes diferencialmente expressos in silico

O método estatístico de Audic-Claverie do programa IDEG6 identificou o perfil de 653 genes, significativamente diferencialmente expressos entre as bibliotecas de folhas de *A. stenosperma* (V10309) controle (SFC) e inoculadas com *P. personata* (SFI), dos quais 388 eram positivamente regulados (*upregulated*) e 265 negativamente regulados (*downregulated*), dentre eles, 211 e 149 genes eram exclusivos das bibliotecas SFI e SFC, com indução variando entre 28,9 a 1,1 vezes e 105,5 a 1,1 vezes, respectivamente (Tabela 4 – Apêndice). Ainda por essa análise estimativa de expressão diferencial, ≈8% dos genes positivamente regulados apresentaram alta similaridade com sequências de funções análogas relacionadas à defesa ou ao estresse, com registros encontrados em outras espécies de plantas tais como: *Arabidopsis thaliana* (L.) Heynh., *Arachis hypogaea* L., *Bruguiera gymnorhiza* (L.) Savigny, *Glycine max* (L.) Merr., *Medicago truncatula*

Gaertn., *Oryza sativa* L., *Phaseolus vulgaris* L., *Populus trichocarpa* Torr. & A. Gray e *Ricinus communis* L., são alguns exemplos. Cerca de 53% desses genes positivamente regulados são desconhecidos, com indução variando entre 272,0 a 1,1 vezes (Tabela 4 – Apêndice).

5.4. Escolha dos genes-candidatos

Os 11 primeiros genes-candidatos foram selecionados a partir do seu score nos diferencialmente expressos e sua similaridade encontrada com sequências depositadas no banco de dados público do NCBI (*National Center for Biotechnology Information*), associadas à anotação genômica de genes/proteínas relacionados à resposta de defesa ou estresse biótico (Tabela 2). Além disso, uma busca por similaridade dos 7.723 *contigs* gerados a partir da análise das duas bibliotecas (SFI e SFC) foi realizada contra o banco de análogos a genes de resistência (RGA - *Resistance Gene Analogs*) de *Arabidopsis thaliana*, gerando informações que foram representadas graficamente em 3D, para visualização da similaridade. Na figura 31 foi possível observar o agrupamento de sequências similares e de grupos com expressão diferencial, representados nesta figura por aglomerados distintos, que sobressaem na parte lateral direita do gráfico em 3D, em forma de “caliandra” (*Clusters* de 1-5), mostrando vários transcritos de uma das bibliotecas seqüenciadas (SFI ou SFC) associados a uma sequência de RGA de *A. thaliana* com alguma similaridade. Isso destacou cinco RGAs candidatos, os quais foram submetidos à avaliação da quantificação relativa de mRNA por meio do RT-qPCR. As sequências dos *primers* desenhados encontram-se na tabela 2.

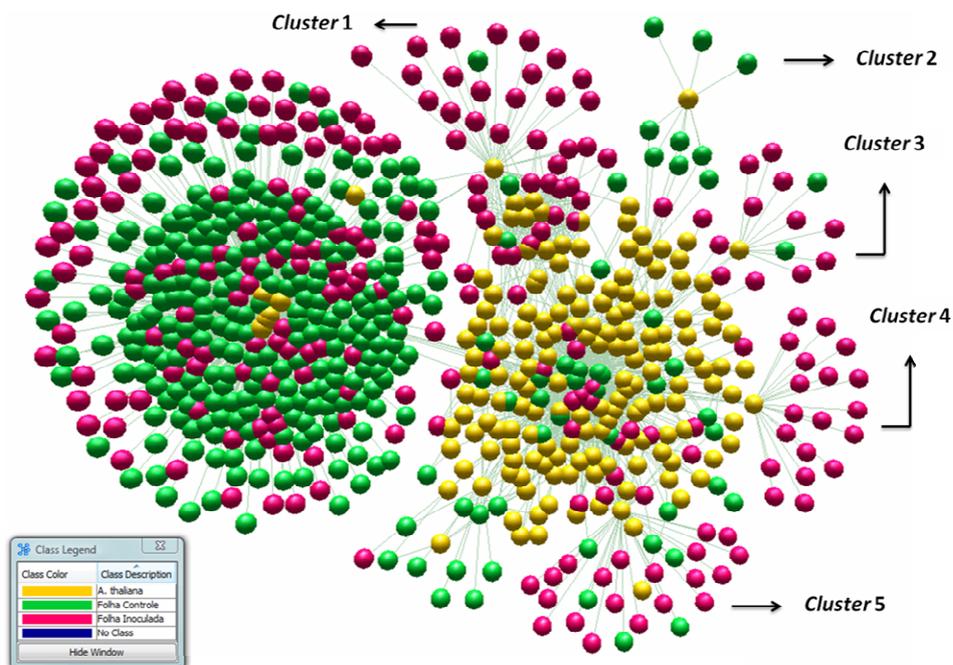


Figura 31 – Agrupamentos gerados no programa BioLayout 3D, baseado na similaridade dos unigenes das bibliotecas SFI e SFC com RGAs de *A. thaliana*, além de similaridades diferenciais de alguns grupos.

Um total de 48 sequências do tipo NBS-LRR (sítio de ligação de nucleotídeo rico em repetições de leucina), foram identificadas nas bibliotecas de *A. stenosperma*, sendo que cinco deles foram escolhidos para análise em RT-qPCR, justamente por apresentarem uma expressão diferencial *in silico*.

Os *primers* citados na tabela 2 foram desenhados para sua utilização em RT-qPCR e para validação da expressão gênica dos genes-alvo selecionados.

Tabela 2 – Primers desenhados para os genes candidatos selecionados em *A. stenosperma*. Valores de *e-value* obtidos pelo BLASTx.

Nº	ID primer	Função putativa	Abreviação	<i>e-value</i>	Sequência dos primers Foward/Reverse (5'–3')
1	4As7	Proteína Dissulfeto Isomerase	PDI	$4e^{-70}$	TACAGTGCTGCTCGCTCATC/ ACTGTTTCCGAGGCTCTTGA
2	4As8	Citocromo P450	P450	$3e^{-04}$	CAAAGGCCTTCATATAGCTCCA/ ATGAGTGAAGGGTTAGGTGTGG
3	4As11	Redutase da Adenosina 5' – fosfosulfato	RAPS	$1e^{-52}$	GGCCATGAGAACATTCGTTT/ GAATGCCCTCAGGTGATGAT
4	4As13	Homeodomínio Zíper de Leucina	HD-Zip	$7e^{-36}$	GATGTCTTCGTGTGGTGGTG/ AGGGTTGGTGCTTGTTTCATC
5	4As20	Proteassoma	Proteassoma	$3e^{-101}$	TGGGTGGTCTTGACAATGAA/ CAGCCAATGGAGTAACAGCA
6	4As22	Domínio transmembrana da proteína quinase rica em repetições de leucina	LRR-TM-PK	0,32	CAGTTTACATGGCCTCTCCTG/ GAGTGGTATTTGGGATGTTCT
7	4As23	Proteínas de EF-hand ligadoras de Ca^{2+}	Bomba Ca^{2+}	$1e^{-12}$	GCCTCTAGAGATTCTTCCCACA/ AGGCTTTCTATGGAAGAAGACG
8	4As24	Fator de ribosilação da ADP	ARF	$6e^{-131}$	TTTGTGGCAAACACCAAGA/ GGATGTCCGAGGTCAAGAGA
9	4As26	Localização de Proteínas Nucleares	NPL	$4e^{-84}$	ATCAAGGACGACGATGATCC/ CGTTCTCAATGGGAAATGT
10	4As27	DNA dependente de RNA polimerase II	DNA-RNA Pol	$5e^{-97}$	AAGCGATGATGGTGGAGAAC/ TTCAGAAGCAGGGATGGAGT
11	4As31	Proteína contendo domínio BURP	BURP	$7e^{-10}$	CTGGGTTCCCGGATCTTACT/ TTGATCCCCATACATCAAACC
12	RGA_c11	Aumento de resistência à doença	EDR1	0.0	GGCAAAGGTGATGCTACCTG/ TTCCCAAGGAAGTTGATCCT
13	RGA_c122	Glicogênio Sintase Quinase	GSK	0.0	TTGCACCAACTTCTGGTTTG/ CAATGATATGCCCGTCTCT
14	RGA_c256	Proteína Calcineurina tipo B (CBL) interagindo com proteínas quinases	CIPKs	0.0	AGCTTCACGGTGAGGAAGAA/ ATTCCGTGTTGCAGAACCTC
15	RGA_c390	Proteína Serina-Treonina Quinase	STPKs	$3e^{-73}$	CAGAACAACGAGAACAATCG/ TTACCGTCATCCGAAACTTG
16	RGA_c391	Receptores ricos em cisteína	CRKs	$2e^{-113}$	TCTAGAAAGCGAGGAGCTCAA/ TGGATTGACTGGGATATGCTC

5.5. Análise da qualidade do cDNA

A qualidade do cDNA sintetizado de cada um dos diferentes *pools* de RNA total de folha de *A. stenosperma* (controle e inoculada, dos pontos 24, 48 e 72h) foi avaliada pela técnica de RT-PCR. Os pares de *primers* leg066 (*forward* e *reverse*) foram utilizados nesta etapa, por amplificarem, no DNA genômico (gDNA), uma região que inclui um intron, permitindo a distinção entre os produtos de PCR amplificados a partir do gDNA (≈ 570 pb) (Figura 32A) e os amplificados do cDNA (≈ 250 pb) (Figuras 32A e B), com isso foi possível determinar se o cDNA sintetizado apresentava ou não contaminação com gDNA. Essa etapa de confirmação é essencial, pois isso pode comprometer a análise por RT-qPCR, possibilitando a geração de falsos positivos. Conforme a figura 32B, foi possível identificar a eficiência do tratamento com DNase das amostras sintetizadas, onde nenhuma contaminação foi verificada.

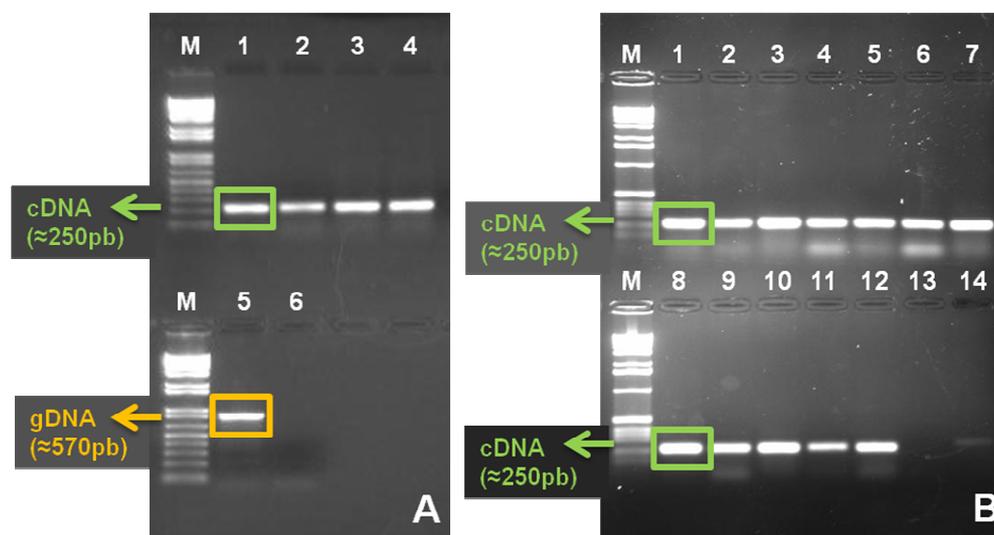


Figura 32 – Gel não desnaturante de agarose a 1,5% corado com brometo de etídio. (A) Produto da RT-PCR onde foram avaliadas as primeiras amostras de cDNA sintetizadas; (1-4) Produto que teve como molde cDNA (SFI 24h *Pool* A, SFI 24h *Pool* B; SFI 48h *Pool* A e SFI 72h *Pool* A); (5) Produto que teve como molde DNA genômico (*A. hypogaea* cv. IAC-Tatu – controle positivo) e (6) Controle negativo da reação, sem cDNA molde. (B) Produto da RT-PCR dos cDNAs sintetizados de todas as amostras. (1) SFC 24h *Pool* A; (2) SFC 24h *Pool* B; (3) SFI 24h *Pool* A; (4) SFI 24h *Pool* B; (5) SFC 48h *Pool* A; (6) SFC 48h *Pool* B; (7) SFI 48h *Pool* A; (8) SFI 48h *Pool* B; (9) SFC 72h *Pool* A; (10) SFC 72h *Pool* B; (11) SFI 72h *Pool* A e (12) SFI 72h *Pool* B; (13) Controle negativo (sem cDNA molde) e (14) Controle positivo (cDNA de *A. stenosperma*). Marcador de peso molecular (M) utilizado foi 1Kb Plus (Invitrogen) SFC = *A. stenosperma*/Folha/Controle, SFI = *A. stenosperma*/Folha/Inoculada, onde e 24, 48 e 72h referem-se aos pontos de coleta.

5.6. Validação dos genes-candidatos por RT-qPCR

Os pares de *primers* desenhados para análise da expressão gênica, por RT-qPCR foram analisados pelo estabelecimento de curva-padrão de diluição seriada (Figura 33), a

fim de avaliar sua eficiência bem como certificar-se de que apenas um produto de amplificação foi gerado. Para tanto, um *pool* formado pela mistura equimolar dos três pontos de coleta dos grupos controle e inoculado, foram diluídos a 10^{-1} , 10^{-2} e 10^{-3} . Os valores médios de C_q foram obtidos para cada diluição e para cada par de *primers* como a média aritmética dos C_q das triplicatas. Aquelas médias com valores de desvio padrão superiores a 0,5 eram rejeitadas, porém não antes de repetir a reação de RT-qPCR, para confirmar a persistência desse desvio. A melhor diluição do cDNA a ser utilizada nas reações de quantificação relativa foi escolhida como sendo aquela que estava dentro de um intervalo de valor média de C_q entre 18-25. Dois genes de referência (*Actina* e *60S ribossomal*), sugeridos por Morgante e colaboradores (2011) também foram avaliados pela curva de diluição seriada.

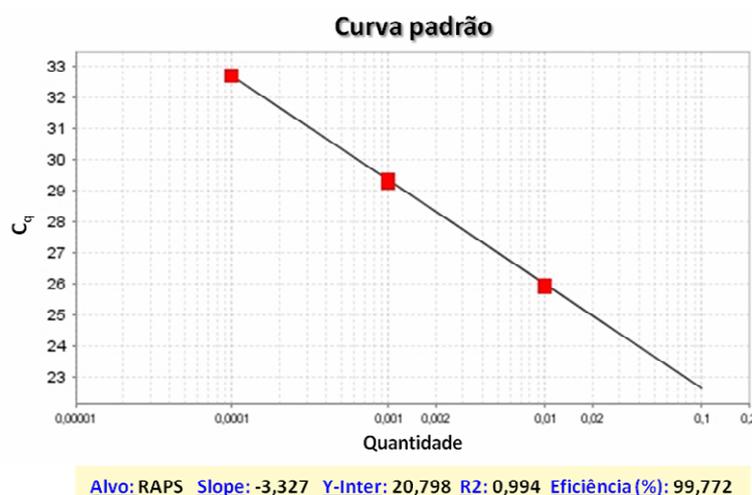


Figura 33 – Exemplo de uma curva padrão de diluição seriada. Dados obtidos pela amplificação dos *primers* RAPS, utilizando-se como molde o cDNA dos *pools* de folhas controle e inoculadas de *A. stenosperma*, diluído a 10^{-1} , 10^{-2} e 10^{-3} .

A partir da regressão linear da curva-padrão da diluição seriada, foi determinado o coeficiente angular da reta (*slope*) usado no cálculo da eficiência de amplificação pela fórmula: Eficiência (\mathcal{E}) = $[10 (-1/\text{coeficiente angular da reta})]$. Valores de *slope* (exemplo figura 33) iguais a -3,32 indicam uma reação com 100% de eficiência de amplificação, onde o número de moléculas amplificadas sobra a cada ciclo da PCR. O coeficiente de correlação (R^2) com valores menores que 0,99 podem indicar problemas de precisão técnica, como problemas na pipetagem. Os pares de *primers* que apresentaram eficiência entre 90-110% e apenas um pico na curva de dissociação foram considerados válidos para a utilização na análise de expressão (Tabela 3 e Figura 33).

Tabela 3 – Dados obtidos pela curva-padrão de diluição seriada dos pares de *primers* selecionados neste estudo. ID = Identidade do par de *primers*; Indicativo dos valores do *Slope*, R^2 (coeficiente de correlação) e média dos C_q das triplicatas na diluição 10^{-1} .

Nº	IDs	Função putativa	<i>Slope</i>	R^2	Eficiência (%)	C_q de 10^{-1}
1	4As7	<i>PDI</i>	-3,238	0,971	103,64	26,7
2	4As8	<i>P450</i>	-3,292	0,998	101,27	24,9
3	4As11	<i>RAPS</i>	-3,327	0,994	99,77	22,2
4	4As13	<i>HD-Zip</i>	-3,341	0,994	99,19	22,9
5	4As20	<i>Proteassoma</i>	-3,273	0,998	102,09	20,4
6	4As22	<i>LRR-TM-PK</i>	-3,414	0,983	96,30	24,3
7	4As23	<i>Bomba de Ca²⁺</i>	-3,476	0,994	93,95	22,4
8	4As24	<i>ARF</i>	-3,547	0,993	91,38	23,7
9	4As26	<i>NPL</i>	-3,379	0,988	97,66	25,3
10	4As27	<i>DNA-RNA pol</i>	-3,188	0,987	105,91	26,8
11	4As31	<i>BURP</i>	-3,453	0,992	94,80	18,5
12	RGA_c11	<i>EDR1</i>	-3,425	0,996	95,86	25,9
13	RGA_c122	<i>GSK</i>	-3,464	0,979	94,41	25,4
14	RGA_c256	<i>CIPKs</i>	-3,500	0,981	93,09	24,7
15	RGA_c390	<i>STPKs</i>	-3,213	0,999	104,76	26,9
16	RGA_391	<i>CRKs</i>	-3,419	0,991	96,09	24,9
17	C23	<i>Actina</i>	-3,499	0,981	93,10	20,5
18	C43	<i>60S</i>	-3,392	0,997	97,15	19,8

Para todos os 18 pares de *primers* analisados, um único pico na curva de dissociação foi observado, indicando a presença de um *amplicon* específico e ausência de dímeros de *primer* (Figura 34).

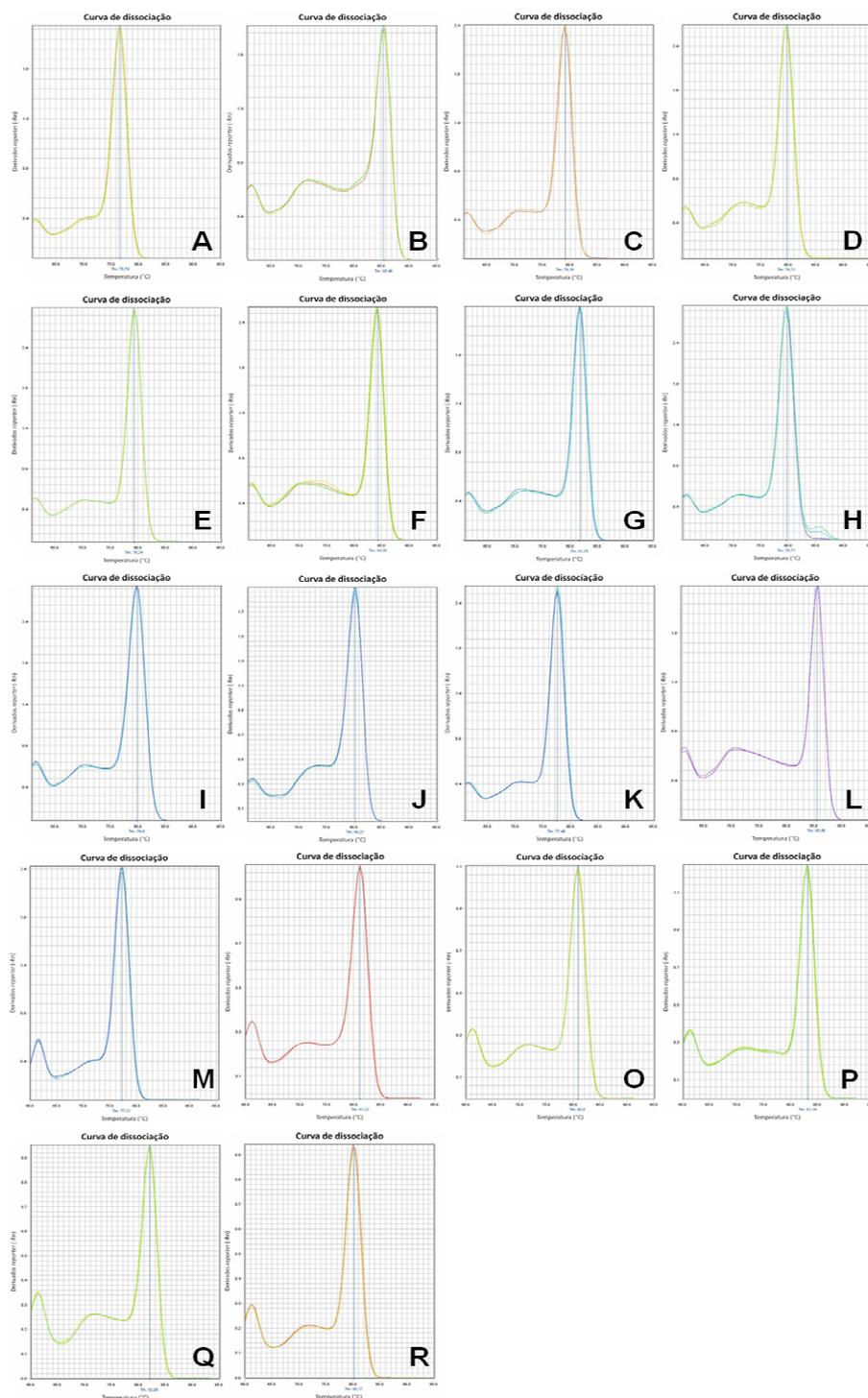


Figura 34 – Curva de dissociação obtida para cada par de *primers* selecionados neste estudo, mostrando a formação de um único pico, referente à amplificação de um único fragmento específico. (A) *Actina*; (B) *60S*; (C) *PDI*; (D) *P450*; (E) *RAPS*; (F) *HD-Zip*; (G) *Proteassoma*; (H) *LRR-TM-PK*; (I) *Bomba de Ca²⁺*; (J) *ARF*; (K) *NPL*; (L) *DNA-RNA pol*; (M) *BURP*; (N) *EDR1*; (O) *GSK*; (P) *CIPKs*; (Q) *STPKs* e (R) *CRKs*. Tm = Temperatura de dissociação (*melting temperature*). OBS.: O pico à esquerda observado em todas as imagens refere-se a um problema gerado pela calibração do equipamento F7300 Real-Time PCR System, da Applied Biosystem.

Desse modo, a análise da curva-padrão de diluição seriada dos pares de *primers* dos genes de referência e dos genes-alvo pela RT-qPCR permitiu, inicialmente, a determinação da viabilidade desses *primers* para o estudo da expressão gênica, bem como a seleção do melhor gene referência, isto é, aquele em que a estabilidade de expressão nos diversos pontos tenha sido a melhor. Os valores de C_q médio obtidos para os genes de referência *Actina* e *60S* nos tratamentos controle e inoculado foram de 21,86 e 21,80; 19,88 e 19,84, respectivamente. Isso mostra que, embora a eficiência de amplificação do cDNA tenha variado entre os pares de *primers* utilizados (93,10% para *Actina* e 97,15% para *60S*; Tabela 3), sua proporção foi mantida durante a amplificação, não favorecendo ou amplificando, diferencialmente, um outro cDNA nas amostras analisadas.

Em seguida, dois *pools* (controle – SFC e inoculado – SFI) contendo, de forma equitativa, amostras dos três pontos (24, 48 e 72hs) do sexto bioensaio, foram submetidos à análise por RT-qPCR, a fim de quantificar a expressão relativa dos 11 genes de *A. stenosperma* (V10309), candidatos a serem expressos diferencialmente na interação planta-fungo (Tabela 2). Para esta análise, utilizaram-se os genes de referência *Actina* e *60S*, conforme recomendação de Morgante e colaboradores (2011). A análise dos níveis relativos de mRNA ou quantificação relativa (RQ) analisada por RT-qPCR é dada pela razão entre o gene-alvo (genes-candidatos) e o gene de referência (*Actina* ou *60S*) no valor de 1. Dessa forma, 1 é o valor inicial comparativo, e portanto, todo valor encontrado por esta razão (SFI/SFC), acima de 1 entende-se que houve uma regulação positiva da expressão gênica, enquanto que um valor abaixo de 1, significa que houve uma regulação negativa da expressão gênica.

A quantificação relativa desses 11 primeiros genes-candidatos mostrou uma pequena expressão diferencial entre as plantas inoculadas e seus controles (inoculado/controle) para a maioria dos genes avaliados (Figura 35). O gene *LRR-TM-PK* (*Domínio transmembrana da proteína quinase rica em repetições de leucina*) destacou-se por ser o único a apresentar uma regulação positiva inoculado/controle (razão acima de 1). Dentre os genes regulados negativamente, os genes *Citocromo P450*, *Homeodomínio de zíper de leucina (HD-Zip)*, *Localização de proteínas nucleares (NPL)* e *proteína contendo domínio BURP*, foram aqueles que apresentaram a maior diferença de expressão relativa (0,74; 0,83; 0,99 e 0,67, respectivamente, normalizados com o gene de referência *Actina* e 0,66; 0,74; 0,89 e 0,59, respectivamente, normalizados com o gene de referência *60S*) (Figura 35).

Os resultados da expressão relativa dos 11 primeiros genes-alvo avaliados foram próximos para dois genes de referência (*Actina* e *60S*) utilizados como normalizadores (Figura 35), indicando que ambos possuem uma expressão altamente estável entre as amostras. Assim, para as análises posteriores, optou-se por usar somente o gene de referência *60S*, dispensando a utilização de um segundo gene de referência, neste caso, o *Actina* (Figura 35), como normalizador.

6º bioensaio

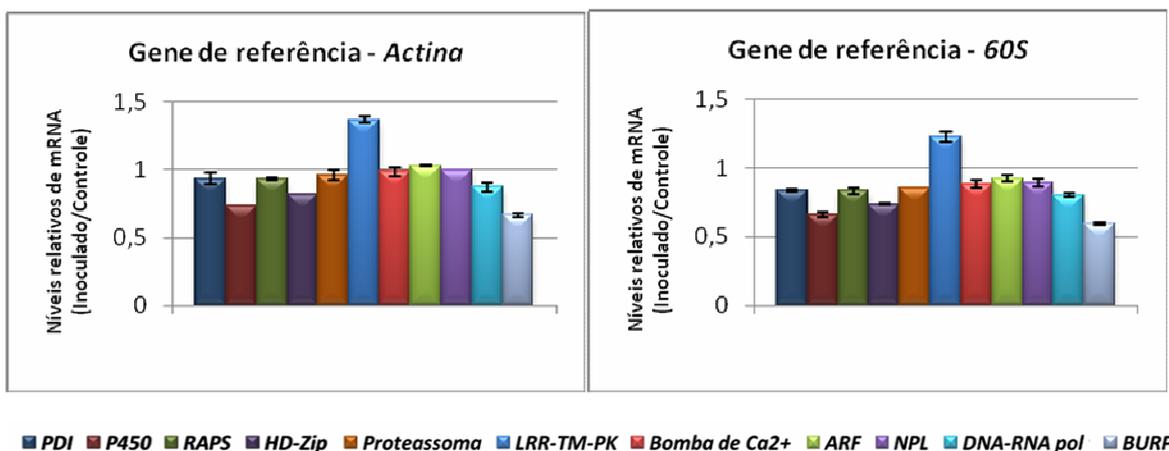


Figura 35 – Análise da quantificação relativa (RQ) por RT-qPCR para os 11 primeiros genes-candidatos, a partir de dois pools de *A. stenosperma* (SFC e SFI) do 6º bioensaio. Os níveis relativos de mRNA (razão entre SFI/SFC) estão representados por barras e o desvio-padrão das triplicatas pelas linhas verticais. (A) Dados normalizados com o gene de referência *Actina*. (B) Dados normalizados com o gene de referência *60S*.

A partir da análise de expressão relativa dos 11 primeiros genes-alvo, os cinco que apresentaram as maiores diferenças de expressão relativa entre as plantas inoculadas e seus controles (inoculado/controle) foram selecionados para uma análise mais refinada de expressão relativa. São eles: *LRR-TM-PK*, *P450*, *HD-Zip*, *NPL* e *BURP* (Figura 36 e Tabela 2). Para validar essa primeira seleção realizada com amostras do 6º bioensaio, os cinco genes-alvo selecionados foram posteriormente avaliados quanto à sua expressão relativa com amostras de cDNA, oriundas dos 4º, 5º e 8º bioensaios, e avaliadas por ponto de coleta e por tratamento. Nesta avaliação, cada bioensaio foi considerado como uma repetição biológica e a média entre os níveis relativos de mRNA encontrados nos três bioensaios foi avaliada separadamente por ponto de coleta e por gene, a fim de analisar o perfil de expressão global entre os ensaios (Figura 36).

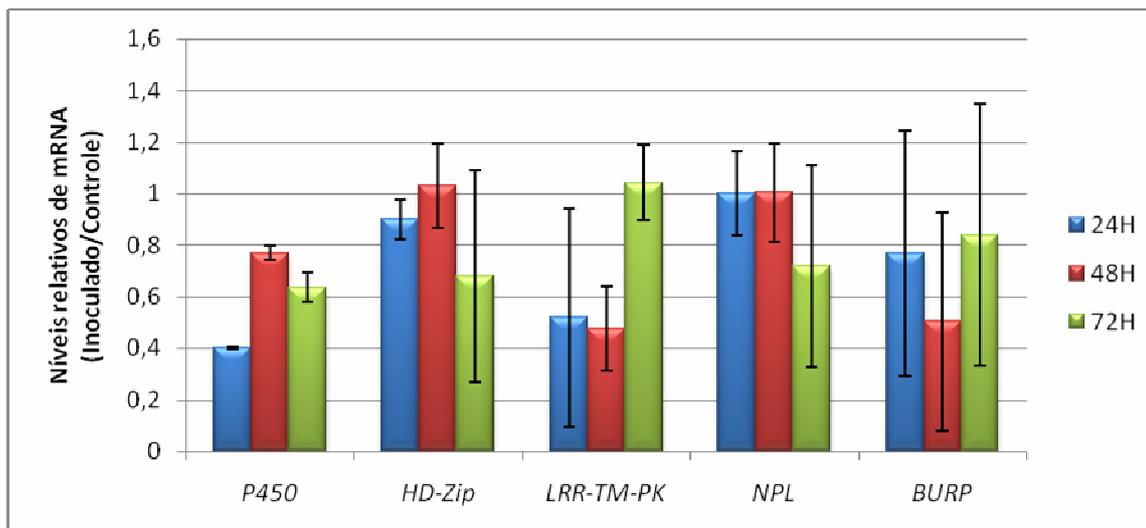


Figura 36 – Análise da RQ por RT-qPCR para os genes-alvo *P450*; *HD-Zip*; *LRR-TM-PK*; *NPL* e *BURP*. Os níveis relativos de mRNA (razão SF1/SFC) estão representados por barras e o desvio-padrão das médias entre os mesmos pontos de coleta (24, 48 e 72h) do 4º, 5º e 8º bioensaios pelas linhas verticais.

A média de quantificação relativa obtida entre os três bioensaios mostra que todos os cinco genes candidatos acima avaliados, por ponto de coleta, apresentaram um perfil de expressão negativamente regulado (*down-regulated*), inclusive o gene *LRR-TM-PK*, que na avaliação global de expressão havia mostrado uma regulação positiva (Figura 35). Os demais genes confirmaram o perfil anteriormente observado.

Para cada gene avaliado, observa-se um perfil de comportamento de expressão distinto ao longo do bioensaio (*time-course*) (Figura 36). Alguns genes (como *HD-Zip* e *NPL*) começam com uma expressão muito próxima (em torno de 1) entre material inoculado e controle nas primeiras horas após a infecção (24 e 48 HAI) para ter uma queda brusca de expressão relativa no último ponto de coleta (72 HAI). Já o gene *LRR-TM-PK* possui um comportamento completamente oposto, ou seja, começa com uma expressão muito negativa (em torno de 0,5) entre o material inoculado e controle nas primeiras horas após a infecção (24 e 48 HAI) para igualar ao controle no último ponto de coleta (72 HAI). O gene *BURP* mostra um comportamento irregular ao longo do bioensaio, enquanto que o gene *P450* tem um comportamento próximo ao *LRR-TM-PK*, isto é, começa com uma expressão muito negativa (em torno de 0,4) no início do bioensaio (24 HAI) para em seguida chegar próximo ao controle nas horas posteriores (48 e 72 HAI).

Entretanto, há uma variação dos dados dependendo de como eles são analisados. Observa-se um padrão de expressão com pequeno desvio-padrão entre as

amostras quando as mesmas são comparadas entre si (triplicatas técnicas) dentro de um mesmo bioensaio (análise intra-bioensaio – dados não mostrados) e um grande desvio padrão na análise que leva em consideração as repetições biológicas (análise inter-bioensaios – Figura 36). Essa grande variação entre os bioensaios, já esperada, deve-se, possivelmente, a fatores não controláveis, aos quais esses bioensaios foram expostos, uma vez que os mesmos foram realizados em casa de vegetação com condições parcialmente controláveis de temperatura e umidade, com grupos de plantas distintos e com diferentes inóculos de *P. personata*.

Em paralelo à análise dos 11 genes-candidatos, procedeu-se a análise por RT-qPCR dos cinco genes RGAs-candidatos (Tabela 2), a fim de quantificar sua expressão relativa. Essa análise foi realizada com dois *pools* (controle – SFC e inoculado – SFI) contendo, de forma equitativa, amostras dos três pontos (24, 48 e 72h) do 6º bioensaio. Para essas análises, utilizou-se somente o gene referência *60S*, pois o mesmo mostrou-se altamente estável nas análises de RT-qPCR anteriores, dispensando a utilização de um segundo gene de referência para a normalização dos níveis de mRNA entre as amostras testadas.

Todos os cinco RGAs candidatos analisados por RT-qPCR mostram-se positivamente regulados quando comparado ao seu controle (razão inoculado/controle, acima de 1) (Figura 37). Os genes de *Aumento de resistência a doença – 1 (EDR1)*, *Glicogênio sintase quinase (GSK)* e *proteína calcineurina tipo B interagindo com proteínas quinases (CIPKs)* apresentaram expressão relativa maior ou igual a 1,2.

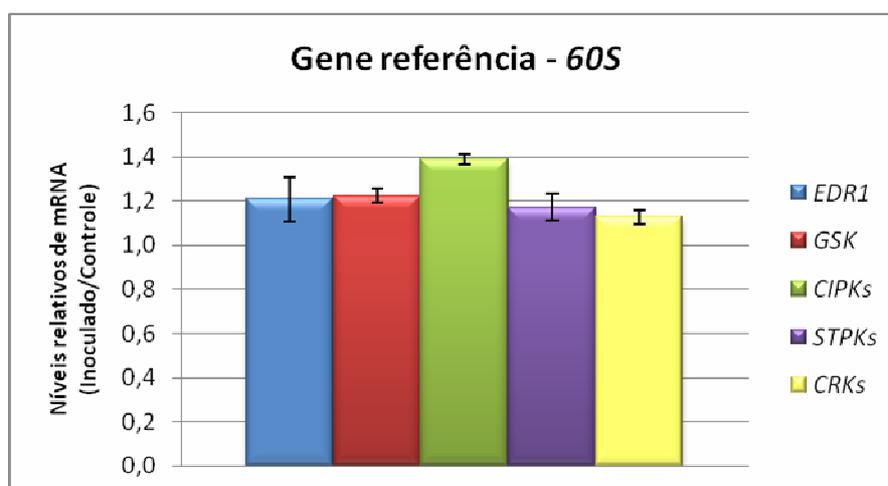


Figura 37 – Análise da quantificação relativa (RQ) para os cinco RGAs-candidatos, a partir de dois *pools* de *A. stenosperma* (SFC e SFI) do 6º bioensaio. Os níveis relativos de mRNA (razão entre SFI/SFC) estão representados por barras e o desvio-padrão das triplicatas pelas linha verticais. Dados normalizados com o gene referência *60S*.

6. DISCUSSÃO

A mancha preta, causada por *P. personata*, é uma das doenças foliares mais graves e economicamente importantes do amendoim (*A. hypogaea*). Vários trabalhos têm relatado a importância dessa doença no mundo e muitos esforços têm sido realizados para determinar meios mais eficientes de controle do patógeno, além de tentar identificar características morfológicas, bioquímicas, genéticas e moleculares da suscetibilidade e resistência envolvidas na resposta à infecção desse fungo em espécies cultivadas e silvestres do gênero *Arachis*.

As plantas utilizam tanto de barreiras pré-formadas quanto induzidas para se proteger contra invasão de patógenos. No entanto, quando essas barreiras pré-formadas são superadas pelo patógeno, o mesmo é reconhecido, induzindo respostas precoces ou tardias de defesa, a fim de inibir a invasão ou lidar com o estresse ocasionado. Apesar dos dois tipos de respostas serem importantes na defesa do hospedeiro, os genes de resposta precoce constitui a chave de percepção e amplificação de sinais de diferentes estresses, além de serem responsáveis pela indução da expressão de genes mais a jusante (Kumar & Kirti, 2011).

Trabalhos anteriores incluindo *P. personata* e outras espécies de *Arachis* (Nobile, et al., 2008; Kumar & Kirti, 2011), demonstraram que entre as primeiras horas de interação planta-patógeno ocorrem eventos cruciais para o estabelecimento ou não da doença. Estes dados foram confirmados por microscopia ótica e microscopia eletrônica de varredura, para *A. hypogaea* e *A. stenosperma*, demonstrando que após a adesão do conídio na superfície foliar, entre o período de 12-72h, já havia o início do crescimento dos tubos germinativos bem como a ocorrência da penetração via estomatal, apenas na espécie suscetível, *A. hypogaea* (Nobile, et al., 2008; Leal-Bertioli, et al., 2010).

Estudos anteriores, incluindo a análise do transcrito de *A. hypogaea* (Nobile, et al., 2008) e *A. diogeni* (Kumar & Kirti, 2011), também observou-se o intervalo de 6-72h e 0-48h, respectivamente, para identificação de genes regulados positivamente (*upregulated*) durante a interação entre *A. hypogaea*/*A. diogeni* e *P. personata*. Dessa forma, neste estudo, optou-se em realizar o experimento utilizando os pontos 24, 48 e 72h, visando identificar genes de resposta precoce que estariam envolvidos na interação entre *A. stenosperma* e *P. personata*.

Até o momento existem três relatos envolvendo a identificação de genes envolvidos na interação *Arachis*–*P. personata*, sendo dois deles com a espécie cultivada, *A. hypogaea* (Luo, et al., 2005b; Nobile, et al., 2008), e apenas um com uma espécie silvestre, *A. diogeni* (Kumar & Kirti, 2011). Dessa forma, a análise do transcrito da interação *P. personata* e *A. stenosperma* será de grande importância para a melhor compreensão das bases moleculares desta resistência. Além disso, *A. stenosperma* é um dos parentais da população de mapeamento de referência de *Arachis* (Moretzsohn, et al., 2005), na qual já foram identificados *loci* de caracteres quantitativos (QTLs – *Quantitative trait loci*) associados à resistência a doença (Leal-Bertioli, et al., 2009). A identificação de genes-candidatos e seu mapeamento neste mapa de referência contribuirão para a identificação de regiões associadas à resistência e, conseqüentemente, a seleção assistida por marcadores para a produção de amendoim mais resistente.

A obtenção de RNA de alta qualidade em altas concentrações a partir de folhas de espécies silvestres de *Arachis* é bastante trabalhosa, possivelmente pela grande concentração de polissacarídeos que podem ter se agrupado ao RNA e co-precipitado, tornando a amostra excessivamente viscosa. Essa viscosidade foi encontrada em várias espécies de *Arachis*, tais como *A. magna* (KG30097), *A. cardenasii* (GKP10017) e *A. duranensis* (K7988) (Brasileiro, et al., 2008; Martins, 2008; Morgante, et al., 2011), além de outras espécies de plantas, tais como *Cinnamomum tenuipilum* – um tipo de canela (Zeng & Yang, 2002), *Pinus densiflora* – pinho, *Picea abies* – Espruce-da-Noruega, *Ginkgo biloba* – Ginkgo (Suzuki, et al., 2008), *Fragaria x ananassa* – morango (Mazzara & James, 2000), *Jatropha curcas* – pinhão manso (Sangha, et al., 2010), dentre outras.

A utilização do Kit Invisorb® Spin Plant RNA Mini (Invitex) possibilitou a redução da viscosidade encontrada no RNA total de folha de *A. stenosperma*, como já descrito em outras plantas (Czechowski, et al., 2005; Vrebalov, et al., 2009; Artico, et al., 2010; Terefe-Ayana, et al., 2011). A tecnologia empregada nesse kit combina as propriedades de ligação seletiva de uma membrana à base de sílica com a velocidade da tecnologia “microspin”. O isolamento de RNA total combina uma lise eficiente e uma homogeneização do material inicial na presença de um tampão altamente desnaturante, contendo isotiocianato de guanidina que inativa imediatamente as RNases, garantindo a purificação de um RNA mais intacto, com remoção do gDNA. O gDNA é fixado na superfície de partículas minerais transportadoras de ácido nucléico, contido nas soluções de lise (DCT ou RP), simultaneamente durante a lise. O tampão de lise DCT é otimizado para materiais vegetais com alta concentração de polissacarídeo, enquanto que o RP é adaptado a materiais vegetais com alto teor de fenol. Após a lise, as partículas minerais transportadoras carregadas com o gDNA são removidos por centrifugação. A adição de

etanol ao sobrenadante filtrado favorece condições adequadas de ligação do RNA total a membrana contida no filtro RTA, enquanto que os contaminantes são lavados em etapas repetitivas. Assim, o RNA total purificado pode ser eluído e armazenado a -80°C.

Muitos projetos com organismos não modelos publicados, tem-se utilizado da plataforma de pirosequenciamento 454 da Roche, pois o sequenciamento *de novo* do transcrito depende de uma tecnologia de sequenciamento com comprimentos maiores de leituras, como as geradas por esta tecnologia, por meio da química GS FLX Titanium XLR70 (atualmente, cerca de 500-1.000pb, são gerados com a nova tecnologia GS FLX Titanium XL+), facilitando, assim, a montagem dos genes transcritos sem um genoma de referência (Kumar & Blaxter, 2010; Bräutigam, et al., 2011; Roche, 2011). Embora ainda seja um desafio montar um genoma complexo e completamente novo usando tecnologias de sequenciamento de próxima geração (454/Illumina), com a montagem *de novo* do transcrito tem sido facilitado pelo aumento da profundidade de cobertura, devido ao número muito menor de nucleotídeos no transcrito do que em todo o genoma, além da quantidade de DNA repetitivo ser menor quando comparado com regiões codificantes (Parchman, et al., 2010).

Além de proporcionar um sequenciamento em larga escala a um custo, relativamente menor, os dados gerados pelo sequenciamento de genes transcritos possibilitam o descobrimento de novos genes (Garg, et al., 2011), análise de expressão em grande escala (Bräutigam, et al., 2011), estabelecimento de relações filogenéticas (Tangphatsornruang, et al., 2010), estudos sobre a dinâmica evolutiva do genoma (Buggs, et al., 2010), desenvolvimento de marcadores moleculares (Parchman, et al., 2010), identificação de sequências simples repetitivas (SSRs) (Kaur, et al., 2011) e polimorfismo de única base (SNPs) (Moe, et al., 2011).

Apesar da existência de vários programas de montagem de sequências oriundas de NGS, não há um consenso e nem uma padronização indicando qual o melhor programa a ser usado. Segundo Kumar e Blaxter (2010), o MIRA, em termos de recursos e flexibilidade é o montador mais abrangente, apesar de mais lento. Além disso, Garg e colaboradores (2011) testaram cinco montadores e dentre eles, o MIRA foi o que produziu maiores montagens em termos de número de bases (≈ 52 Mb). *Contigs* com sequências maiores permitem uma melhor comparação com sequências já depositadas em bancos de dados públicos, aumentando a identidade entre o alinhamento das sequências comparadas.

Neste trabalho, essa tecnologia permitiu, em uma única corrida, a cobertura do genoma de até 85Mbp de leituras de alta qualidade para o genoma diplóide de *Arachis*

estudado. Com a produção de 10.189 *singletons*, que em comparação a outros estudos com legumes, pode ser considerada uma pequena quantidade (Tangphatsornruang, et al., 2010; Dubey, et al., 2011; Kaur, et al., 2011). O comprimento médio do *contig* foi de 457 bases. Um tamanho médio, parecido, de *contig* com 454 bases foi encontrado no trabalho de Franssen e colaboradores (2011) com ervilha. Valores próximos foram encontrados nos trabalhos com Artemísia (334 bases) (Wang, et al., 2009) e com pinheiro (*Pinhus contorta* Douglas ex Loudon – 500 bases) (Parchman, et al., 2010), porém em um trabalho com lentilha produziu *contigs* com o tamanho médio de 770 bases (Kaur, et al., 2011). A montagem de *contigs* maiores é devido à nova química GS FLX Titanium XLR70, permitindo a sequências de leituras de até 300-600pb, e, portanto a montagem de *contigs* maiores (Kaur, et al., 2011).

Dos 7.723 *contigs*, 52% apresentou *hits* no BLASTx com similaridade às sequências de peptídeos no domínio público. Este valor, em porcentagem, é o mesmo encontrado em trabalhos com espécies que não tem genoma referência disponível, tais como, o apresentado por Russel e colaboradores (2011) em seu trabalho com groselha (*R. nigrum*). Esse valor foi maior que os relatados para eucalipto (Novaes, et al., 2008) e pinheiro (Parchman, et al., 2010), por exemplo.

A falta de um genoma de referência completamente seqüenciado e anotado, torna muito difícil a estimativa da cobertura do transcrito alcançada neste estudo para *A. stenosperma*, tendo-se em vista que mais de um *contig* ou *singleton* pode ter origem de um único gene, isso devido a não sobreposição de uma sequência de leitura ou por altos níveis de erros em uma única sequência de leitura (Kaur, et al., 2011), podendo causar uma superestimação do número de *contigs* produzidos num estudo.

Dentre as sequências com similaridade encontrada pelo BLASTx, alguns transcritos foram identificados como fatores de transcrição (FTs). As famílias FTs são altamente conservadas entre organismos eucarióticos, especialmente em plantas. Por meio de trabalhos com transcrito de plantas durante a interação com algum patógeno, vários FTs têm sido encontrados entre os genes diferencialmente expressos, implicando-os em processos de regulação de desenvolvimento específico ou respostas ao estresse biótico e abiótico (Udvardi, et al., 2007; O'Brian & Vance, 2007).

Alguns dos FTs aqui encontrados também foram relatados em estudos anteriores em outras leguminosas: *Lotus japonicus* (Regel) K. Larsen, *Medicago trunculata* Gertn., *Glycine max* (L.) Merr. (Wang, et al., 2010) e *A. diogeni* Hoehne (Kumar & Kirti, 2011), com bZIP (18%), MYB (14%) e AP2-EREBP(10%) mais representados. A família com maior representação foi a pertencente à família de domínio zíper de leucina básico (bZIP),

sendo associados com reguladores de processos de desenvolvimento e fisiológicos, por exemplo, via de sinalização de hormônios, bem como a mecanismos de defesa ao estresse biótico e abiótico (Niggeweg, et al., 2000; Corrêa, et al., 2008; Govind, et al., 2009). A expressão de bZIP relacionada a resposta de defesa já foi descrita, por exemplo, na interação entre trigo e *Puccinia striiformis* f. sp. *tritici* (Zhang, et al., 2008).

A segunda família de FTs com maior representação durante a interação *A. stenosperma* x *P. personata*, foi a de domínio MYB, um domínio altamente conservado entre animais e plantas. Em plantas existem boas evidências para distintas funções para as diferentes proteínas MYB, sendo algumas de controle do metabolismo secundário, de regulação da morfogênese celular e algumas servindo na via de transdução de sinal a resposta a reguladores de crescimento vegetal. Dentro do controle do metabolismo secundário, o MYB atua no metabolismo dos fenilpropanóides, modificando os compostos derivados a partir da fenilalanina, além de ser responsável pela produção de antocianinas (maior grupo de pigmentos plantas). Como regulador da morfogênese celular, ele é essencial em *Petunia* para o desenvolvimento da forma cônica das células epidérmicas da pétala e em *Arabidopsis*, é essencial para diferenciação das células pilosas das folhas e caule. E como sinalizador de transdução, ele atua em resposta hormonal durante o desenvolvimento e germinação (Martin & Paz-Ares, 1997). Em *Arabidopsis* esse fator de transcrição foi identificado sendo ativado durante as primeiras etapas a resposta de hipersensibilidade (HR) em resposta a *Xanthomonas campestris* pv *campestris* (Xcc) (Daniel, et al., 1999).

E por último, a terceira família de FTs, também com boa representação de transcritos foi a APETALA2-Elemento de ligação a proteína responsiva ao etileno (AP2/EREBP), estando envolvido em vários aspectos de crescimento e desenvolvimento da planta, tal como o desenvolvimento floral, transdução de sinais hormonais e diferenciação celular, bem como em resposta ao estresse biótico e abiótico (Sharoni, et al., 2010). Alguns genes tais como TERF1 de arroz, HvRAF de cevada e TSRF1 de tomate estão envolvidos na regulação da tolerância tanto ao estresse biótico quanto o abiótico (Kim, et al., 2004; Jung, et al., 2007; Li, et al., 2011).

Apesar dos fatores de transcrição funcionar como reguladores de diversos processos celulares, muito ainda se tem a ser elucidado, principalmente acerca do papel funcional ao qual estão associados, bem como a sua dinâmica na participação em respostas de defesa a patógenos.

As plantas têm desenvolvido estratégias notáveis para detectar e se adaptar a estímulos ambientais pelo uso de receptores na superfície celular e dentro da célula

(Gómez-Gómez & Boller, 2002), gerando uma cascata de transdução de sinais, que inclui a percepção de um sinal ou um conjunto de sinais primários, extracelular, que alteram a interação do receptor associado, transmitindo sinais secundários pela membrana plasmática, induzindo cascatas de fosforilação e desfosforilação específicas, ocasionando mudanças no metabolismo vegetal e assim, ativando genes envolvidos em respostas de defesa (Gachomo, et al., 2003).

Os genes ativados na resistência a doenças em plantas podem ser divididos em três classes: 1) genes que codificam fatores de transcrição; 2) genes que codificam enzimas que estão envolvidas nas vias metabólicas secundárias e; 3) genes que são induzidos em uma fase tardia da resposta de defesa, não somente de forma local, mas também sistematicamente (Takemoto, et al., 1999). Sendo assim, os genes-candidatos selecionados enquadravam-se na primeira classe, acima citada. Para estes genes procurou-se, também, analisar o perfil de expressão para cada um pela técnica de RT-qPCR.

O sistema de monooxigenases, conhecido como *citocromo P450* (superfamília *CYP450*) é um grupo de proteínas redox que cataliza várias reações oxidativas (Isin & Guengerich, 2007) de compostos endógenos (via biossintética) e exógenos tóxicos (via desintoxicante), propondo assim, que o *CYP450* media a síntese e metabolismo de muitos compostos fisiologicamente importantes, como compostos primários e secundários que atuam como agentes de defesa vegetal contra uma ampla diversidade de fitopatógenos e pragas, sendo alguns deles, esteróides, ácidos graxos, ligninas, terpenos, alcalóides, fenilpropanóides, glicosídeos cianogênicos e fitoalexinas (Kim & Tsukaya, 2002; González-Mendoza, 2007; Ohnishi, et al., 2009; Pinot & Beisson, 2011), bem como mediando à desintoxicação herbicida (Werck-Reichhart, et al., 2000). No que se atem as resposta de defesa pode-se incluir a resposta de hipersensibilidade (Glazebrook, 2005) e a inibição do crescimento de um determinado patógeno (Kliebenstein, et al., 2005). Centenas de genes *CYP450* têm sido identificadas em diferentes espécies de planta. Como, por exemplo, em *Arabidopsis thaliana*, onde 290 genes *CYP450* estão representados em 69 subfamílias, de 44 de um total de 59 famílias *CYP450* existentes, em todas as espécies de planta (Kim & Tsukaya, 2002; Schuler & Werck-Reichhart, 2003).

Dada a participação de genes *CYP450s* em diversas vias bioquímicas, não é de se estranhar que os níveis de transcrição possam variar significativamente em diferentes tecidos e sob diferentes condições de estresse (Schuler & Werck-Reichhart, 2003). Por exemplo, para *t*-CAH no *core* da via de fenilpropanóides, são expressos em um nível

constitutivo significativo em muitos tecidos e moderadamente induzido em resposta a alguns estímulos, incluindo os bióticos, outros genes de hidroxilase de ácidos graxos são expressos em níveis insignificantes, induzidos por ferimentos ou tratamentos químicos (Frank, et al., 1996; Pinot & Beisson, 2011).

Oh e colaboradores (1999) verificaram, o nível de expressão do gene *PepCYP*, um CYP450, é maior em interações incompatíveis do que em compatíveis entre pimentão (*Capsicum annuum* L.) e o fungo causador da antracnose (*Colletotrichum gloeosporioides*). O mesmo ocorreu com o gene *CYP736B* durante interações compatíveis e incompatíveis entre videira e *Xylella fastidiosa* (Cheng, et al., 2010). Além disso, Takemoto e colaboradores (1999) observaram que a expressão de *CYP82E1* aumentada dentro de 6 horas após a inoculação de folhas de fumo com *Pseudomonas syringae* pv. *glycinea*, mostrando um pico de expressão ainda maior às 12 horas após a inoculação (HAI), porém, posteriormente havia um decréscimo de expressão, mas na área ao redor do local aonde o fungo foi inoculado, não foi detectado nenhuma indução do *CYP82E1*. Esse resultado foi contrastante com os baixos níveis e com a expressão tardia observados em folhas de fumo inoculadas com *P. syringae* pv. *tabaci*, sugerindo que a expressão do gene *CYP82E1* seja regulada pelo(s) sinal(is) envolvido(s) na resistência a doença e que o produto desse gene possa estar envolvido na resistência a doença. Em outro trabalho, analisando a indução de um agrupamento do gene *CYP72A* pelo *Magnaphorte grisea* em arroz, no início da infecção (48HAI), Wang e colaboradores (2004), verificaram que dentre os sete genes *CYP72A* detectados, quatro (*CYP72A18*, *CYP72A19*, *CYP72A22* e *CYP72A23*) foram regulados pelo *M. grisea*, mostrando expressão diferencial na interação compatível e na interação incompatível. Na interação compatível (H7S), *CYP72A18* foi negativamente regulado entre 4-8HAI, se recuperando às 12HAI, e tornando a baixar a partir das 24HAI. Na interação incompatível (H7R), este gene tinha padrão de expressão similar, porém sem recuperação até as 24HAI após inoculação, indicando que este gene foi ritmicamente regulado pela infecção do patógeno, sendo considerada uma característica incomum para genes relacionados á defesa. *CYP72A19* foi positivamente regulada em H7R 24HAI, mas sem regulação aparente em H7S. *CYP72A22* teve sua expressão um pouco reprimida as 12 e 24HAI em H7R. *CYP72A23* foi positivamente regulada as 24HAI em H7R, mas negativamente regulada em H7S. Esses resultam significam que esses quatro genes *CYP72A* funcionam, provavelmente, de forma diferente na resposta de defesa no arroz. A diferenciação de expressão encontrada entre os membros de *CYP450* pode justificar o perfil de expressão negativamente regulado encontrado neste trabalho, para o gene putativo *P450*, durante a interação incompatível entre *A. stenosperma* e *P. personata*,

isso porque vários *CYP450s* em subfamílias diferentes respondam de forma independente nesses estresses, podendo ser positiva ou negativamente regulados (Schuler & Werck-Reichhart, 2003).

O *HD-Zip* é uma das famílias (superfamília) de fatores de transcrição chamada homeodomínio, presente apenas em plantas (Javelle, et al., 2011). Essa superfamília é composta por quatro subfamílias, as quais se diferenciam uma da outra por apresentarem combinações distintas do *HD-Zip* a outros domínios. *HD-Zip* subfamília I (*HD-Zip-I*) apresenta uma distribuição comum de introns/éxons de acordo com suas relações filogenéticas, seu homeodomínio é altamente conservado, enquanto que seu domínio *zíper de leucina (LZ)* é menos conservado (Ariel, et al., 2007; Arce, et al., 2011). *HD-Zip* subfamília II (*HD-Zip-II*) possui um motivo *CPSCE* na extremidade C-terminal, responsável pela regulação redox da atividade da proteína (Ariel, et al., 2007), e um motivo *ZIBEL* na extremidade N-terminal (Mukherjee, et al., 2009). *HD-Zip* subfamílias III e IV (*HD-Zip-III* e *HD-Zip-IV*) são definidas pela presença de mais dois domínios: proteína esteroidogênica regulatória aguda relacionada à transferência de lipídios (*START*) (Ponting & Aravind, 1999) e o domínio adjacente ao *START* (*SAD*), podendo ser distinguível entre si, pelo quinto domínio, presente na família *HD-Zip-III* de um motivo *MEKHLA* na extremidade C-terminal (Mukherjee & Bürglin, 2006). Membros do *HD-Zip-I* estão geralmente envolvidos em respostas ao estresse abiótico, como estresse hídrico e estresse a luz (Harris, et al., 2011). Genes da subfamília *HD-Zip-II* estão envolvidos principalmente no fototropismo e resposta a auxina (Elhiti & Stasolla, 2009). Membros do *HD-Zip-III* possuem papel importante durante a morfogênese, enquanto que a expressão de vários membros do *HD-Zip-IV* é muitas vezes restrita nas células externas dos órgãos vegetais, regulando processos como destino da epiderme, formação de tricomas e acumulação de antocianina (Elhiti & Stasolla, 2009). Poucos são os trabalhos que correlacionam, diretamente, a resposta de defesa vegetal contra um patógeno, a alguma proteína *HD-Zip*. No entanto, Park e colaboradores (2011), verificaram que os genes *ATHB7* e *ATHB12* tiveram a expressão induzida pela infecção em *A. thaliana* com o BSCTV (*Beet severe curly top virus*) e que o desenvolvimento de sintoma induzido por BSCTV estava fortemente correlacionado com a ativação transcricional de *ATHB12*. Em outro estudo, verificou-se que o gene *HAHB4*, uma proteína *HD-Zip*, participa na regulação de respostas ao estresse biótico e a ferida em girassol (*Helianthus annuus* L.), sendo que essa indução depende de elementos presentes na região promotora de *HAHB4* e também do ácido metila jasmônico (MeJA) e do etileno (ET). A superexpressão de *HAHB4* no girassol resulta no aumento do nível de transcritos e na atividade da proteína de vários genes associados com as respostas ao estresse biótico e ao

ferimento, sugerindo que este gene desempenha um papel na integração e na ativação cruzada desses fitohormônios durante essas respostas (Manavella, et al., 2008). Um número significativo de genes relacionados às respostas de defesa foi identificado como diferencialmente expresso, muitos deles negativamente regulados, em plantas transgênicas expressando constitutivamente o gene *HAHB10*, contudo, este gene foi induzido após infecção de folhas de girassol com uma estirpe virulenta de *P. syringae*, mostrando um pico às 48HAI. Além disso, 24HAI, um aumento significativo nos níveis de SA foi observado em plantas transgênicas, comparadas ao seu controle. Com 48HAI os níveis foram reduzidos aos níveis do controle, isso enquanto os níveis de JA não tiveram alteração significativa durante a infecção (Dezar, et al., 2011). Diferente do que foi relatado nos trabalhos acima mencionados, o gene *HD-Zip* em *A. stenosperma*, foi negativamente regulado, onde às 24HAI ele apresentou um ligeiro decréscimo de expressão, retomando a ativação de expressão às 48HAI e voltando a cair às 72HAI, quando comparada ao seu controle.

As *NPLs* (*Nuclear protein localization*) contêm uma sequência de localização nuclear (NPS – *Nuclear localization sequence*), sequência de aminoácidos capazes de redirecionar proteínas não nucleares, permitindo o reconhecimento da proteína pelo núcleo, sendo assim transportada através do poro nuclear de forma ATP-dependente (Sadler, et al., 1989; Silver, 1991; Bossie, et al., 1992).

Alguns mutantes *NPL1* mostram deficiências na retenção de proteínas nucleares e, podendo ser importante para a translocação de proteínas através da membrana do retículo endoplasmático. Suspeita-se que *Npl6* também possa estar envolvido com processo de translocação (Osborne & Silver, 1993).

Npl3 codifica uma proteína nuclear com similaridade a uma classe de proteínas que contém um motivo de ligação ao RNA, propondo um papel na exportação de mRNA, importação de proteína e processamento do rRNA (RNA ribossomal) (Bossie, et al., 1992; Osborne & Silver, 1993; Henry, et al., 1996).

Os genes *Npl4* estão envolvidos nos mecanismos de transporte nucleocitoplásmico, além de codificar um componente poro nuclear e estar implicado com na biossíntese dos ácidos graxos insaturados. A deficiência desse gene implica em falhas no transporte de importação de proteínas para o núcleo, exportação do mRNA, defeitos na estrutura nuclear, separação das membranas interna e externa nuclear, e protruções na membrana contendo grandes acumulações de mRNA (DeHoratius & Silver, 1996). Foi verificado que o gene *HRD4* era idêntico ao *NPL4*, o qual é necessário para diversos substratos para degradação associada ao retículo endoplasmático (ERAD). *Hrd4/Npl4*

não interfere nem na ubiquitinação nem na função proteassomal, mas parecem exercer um papel interessante entre essas duas etapas, ou anterior a ubiquitinação, como a construção ou edição de cadeias poliubiquitinadas adicionadas às proteínas-alvo, que será reconhecida pelo complexo proteassoma 26S (Bays, et al., 2001; Hitchcock, et al., 2001).

O gene *NPL* em *A. stenosperma* na presença de *P. personata*, foi expresso constitutivamente, mas mostrando uma regulação negativa às 72HAI. Provavelmente sua expressão não foi alterada no início do processo de infecção (24 e 48HAI), porque a quantidade ubiquitínica presente na célula, e frequentemente produzida, não precisaram oscilar para manter os mecanismos que estavam sendo ativados. Contudo, a redução de expressão, diante do pouco que se sabe sobre esse gene, pode sugerir que a desativação desse gene possa estar relacionada à ativação ou desativação de algum mecanismo celular, ainda não conhecido. Pouco se sabe sobre os genes *NPLs*, dificultando a inferência de sua atividade em processos de resposta a estímulos abióticos ou bióticos.

A família de genes que contém proteína com domínio *BURP*, composto pelos motivos de sequência de aminoácidos inicialmente identificados em *BNM2* (*Brassica napus*), *USP* (*Vicia faba* L.), *RD22* (*A. thaliana*) e *PG1 β* (*Lycopersicon esculentum* Mill., subunidade b de poligalacturonase isoenzima 1), é, aparentemente, exclusiva ao reino vegetal, e está localizado na região C-terminal de cada proteína. Este domínio contém vários resíduos altamente conservados, um peptídeo de transporte hidrofóbico e uma região variável interna que pode ser de natureza repetitiva. Embora as proteínas contendo o domínio *BURP* compartilhem características estruturais primárias, seus padrões de expressão e as condições em que são expressas diferem, podendo assim desempenhar diversos papéis na célula (Hattori, et al., 1998; Granger, et al., 2002).

Apenas uma única proteína com domínio *BURP* tem função conhecida, a *PG1b*, que é uma glicoproteína não catalítica que forma um complexo com as isoenzimas poligalacturonase catalítico, *PG2* (Hattori, et al., 1998).

Shunwu e colaboradores (2004) verificaram que o gene *BnBDC1*, homólogo ao *RD22*, em *B. napus* era especificamente expresso na parte aérea, mas não nas raízes, sugerindo ser órgão específico. Além disso, a expressão desse gene foi positivamente regulado por ABA e estresse osmótico, como em tratamento com manitol e NaCl, negativamente regulado por radiação UV e SA, enquanto que o tratamentos com frio e H₂O₂ não mostraram uma tendência clara no padrão de expressão, sugerindo o envolvimento desse gene tanto no estresse abiótico quanto no biótico. Em outro estudo,

um gene com domínio *BURP* foi positivamente regulado em raízes de tomate (*L. esculentum* cv. Kyouryoku-Beijyu) infectadas com *Globodera rostochiensis* (Woll.) (Uehara, et al., 2007). Contudo, e reafirmando que os padrões de expressão de genes contendo o domínio BURP e as condições em que eles são expressos diferem, Xu e colaboradores (2010), avaliaram o padrão de expressão em oito tecidos e órgãos diferentes (raiz, caule, folha, flor, epicótilo, hipocótilo, cotilédone e semente) de soja cv. Zhonghuang13, e verificaram que os genes GmBURP variam muito em suas especificidades e nos níveis de expressão. O envolvimento de vários hormônios vegetais, bem como outras respostas de defesa ativadas, pela presença de *P. personata* em *A. sternosperma* pode ter influenciado a expressão negativamente regulada desse gene, apresentando uma queda na expressão desse gene às 48HAI, sendo retomada, logo em seguida, às 72HAI. Pode ser que haja um pico de produção, por exemplo, do ácido salicílico (SA), que já foi identificado como regulador negativo desse gene, indicando, provavelmente um reforço na indução da produção de proteínas PR. O perfil de expressão observado na infecção fúngica em espécie silvestre de *Arachis* difere da expressão verificada por Uehara e colaboradores, durante a interação tomate (*L. esculentum* cv. Kyouryoku-Beijyu) e o nematoide *Globodera rostochiensis*, onde o gene com o domínio BURP foi positivamente regulado (Uehara, et al., 2007).

Enfim, pouco se sabe acerca das funções desempenhadas pelas proteínas com domínio BURP, principalmente, acerca da sua relação com infecções fúngicas. A maioria dos trabalhos publicados com o estudo de algum gene pertencente ao domínio BURP estão relacionados a respostas ao estresse abiótico ou processos de desenvolvimento vegetal (Hattori, et al., 1998; Granger, et al., 2002; Shunwu, et al., 2004; Ding, et al., 2009; Son, et al., 2009; Gan, et al., 2011).

Além das respostas de defesa acima citados, onde é possível observar o desencadeamento de cascatas de sinalização, entrelaçando o funcionamento de vários genes ativados, que objetivam a retenção ou o impedimento da invasão e dispersão do patógeno, uma segunda linha de defesa é também ativada, caso o patógeno consiga ultrapassar essa barreira inicial, pelo reconhecimento específico da molécula elicitora do patógeno, resultando em uma imunidade desencadeada pelo efector (ETI – *effector-triggered immunity*) (Jones & Dangl, 2006; Göhre & Robatzek, 2008; Greenshields & Jones, 2008). A ETI é muitas vezes expressa em alguns, mas não todos os genótipos dentro de uma espécie vegetal, que no caso de uma interação incompatível, o patógeno é detectado pela planta contendo um gene R, sendo essa planta resistente, evidenciando assim, que os genes R são determinantes polimórficos de resistência do hospedeiro contra patógenos específicos (Xiao, et al., 2008). Os genes R possuem um ou mais

motivos com sequências de aminoácidos altamente conservados, facilitando não somente a sua identificação, como também a identificação de análogos a genes de resistência (RGAs – *Resistance gene analogs*), os quais são supostos genes de resistência a doença, que são identificados com base na similaridade de suas sequências/estrutura, sendo eficazes na identificação e isolamento de genes R (McDowell & Woffenden, 2003; Khan, 2009). Estudos têm mostrado que muitos genes R putativos estão agrupados, e foram submetidos à duplicação e evolução devido à diversificação de seleção (Lucas, 2011). Neste estudo, a partir de 48 diferentes RGAs identificados a partir do transcrito pirosequenciado, quando comparados contra o banco de RGAs de *A. thaliana*, cinco RGAs se destacaram, formando grupos com expressão diferencial quando desafiados por *P. personata*, quando comparadas a expressão basal nas amostras controle.

A classe *LRR-TM-PK* ou *eLRR-TM-Quinase* é oriunda de uma combinação entre três diferentes domínios protéicos. Os genes desta classe codificam uma proteína quinase citoplasmática, que possui um domínio serina/treonina quinase (PK – *Protein kinase*), um domínio transmembrânico (TM – *transmembrane*) e um domínio extracelular com repetições ricas em leucina (LRR – *Leucine-rich repeats*), onde o domínio LRR extracelular estaria apto a reconhecer a presença de elicitores produzidos pelos patógenos e o domínio quinase, no citoplasma, seria o ativador de cascatas de transdução de sinais (Ellis & Jones, 1998; Hammond-Kosack & Jones, 1997; Wang, et al., 2005). De acordo com Song e colaboradores (1995), para o gene *Xa21* de arroz, que confere resistência a *Xanthomonas oryzae* pv. *oryzae* (Xoo), pertencente à classe LRR-TM-PK, sugere-se que a ação deste gene ocorra no reconhecimento da superfície celular do patógeno, com subsequente ativação de respostas de defesa intracelulares. Porém, diferente das proteínas R associadas à classe NBS-LRR, a qual não apresenta outras funções, a não ser a já conhecida resistência a doenças, a família de proteínas da classe LRR-TM-PK, que não possui o domínio de sítio de ligação a nucleotídeo (NBS – *nucleotide binding site*) comporta-se de forma polivalente, incluindo, por exemplo, proteínas de desenvolvimento, proteínas receptoras de hormônios, proteínas de imunidade nata e receptores extracelulares de sinais derivados do patógeno (Ellis & Jones, 1998; Vleeshouwers, et al., 2001; Ellis & Jones, 2003; Diener & Ausubel, 2005).

Ponciano e colaboradores (2006) verificaram que o gene *Xa21* media à resposta de resistência, induzindo a expressão de outros genes (*OsPR1a*, *OsPR1b*, *OsPR1c* e *OsPR10a*) em resposta a inoculação de Xoo raça 6, em arroz. Enquanto que Gómez-Gómez e Boller (2000) indicaram que o gene *FLS2*, isolado a partir de *Arabidopsis*, está envolvido no reconhecimento da flagelina como um sinal da presença de bactérias e leva

à ativação de respostas de defesa, porém sua expressão é induzida por meio da sinalização de etileno (Boutrot, et al., 2010; Mersmann, et al., 2010; Lopez-Gomez, et al., 2011; Zhao & Guo, 2011).

O gene mutante *edr1* (*enhanced disease resistance 1*), isolado de uma *Arabidopsis* mutante, tipo selvagem, confere resistência aumentada a fungos e bactérias, e apresentam indução em várias respostas de defesa, entre elas a morte celular. Esse gene codifica uma classe de proteínas chamada *MAPKKK* (Proteína quinase quinase quinase ativada por mitógenos), com função de emitir sinais ativando e desativando outros genes, podendo funcionar como um regulador negativo no topo da cascata MAPK em plantas do tipo selvagem, na resposta de defesa dependente de ácido salicílico e em resposta ao etileno, enquanto que a função normal de *EDR1* seria prevenir a iniciação inadequada de morte celular, inibição do crescimento e senescência (Frye, et al., 2001; Suzuki K. , 2002; Niu, et al., 2005; Tang, et al., 2005). Segundo Wawrzynska e colaboradores (2008), *edr1* pode mediar à morte celular por um mecanismo de direcionamento de fatores de transcrição para o proteassoma, similar ao de regulação do etileno pelo *CTR1* (*Constitutive triple response 1*). O gene *edr1*, parece ser normal na ausência de patógenos, mas apresenta maior resistência quando em contato com o patógeno. Além de regular as respostas a patógenos, *EDR1* também regula as respostas a estresses abióticos, como a seca (Frye, et al., 2001; Niu, et al., 2005; Tang, et al., 2005). Outra evidência recém relatada por Gu e Innes (2011) é o funcionamento em conjunto entre os genes *EDR1* e *KEG* (*Keep on going*) para a regulação do tráfico endocítico e/ou formação do complexo de sinalização de vesículas a rede trans-golgi para endossomo precoce (TGN/EE – *trans-Golgi network/early endosome*). Christiansen e colaboradores (2011) tentando entender como ocorre a resistência mediada por *edr1*, analisaram o transcrito de plantas do tipo selvagem e plantas *edr1* inoculadas com o fungo patogênico *Golovinomyces cichoracearum* (oídio) e observaram que a expressão de muitos genes suposta ou conhecidamente associados à defesa foi mais rapidamente induzida e com níveis mais elevados em plantas *edr1* em relação ao tipo selvagem. Além disso, por meio de microscopia confocal mostrou que uma fração significativa de *EDR1* foi localizada no núcleo, sugerindo que este gene poderia interagir com fatores de transcrição no núcleo.

Os *GSKs* são uma família de serina/treonina quinases citoplasmática, multigênica, altamente conservadas, que pertencem à superfamília das proteínas quinase mitógeno-ativadas e são encontradas em animais, fungos e plantas (Tavares, et al., 2002), sendo evolucionariamente conservada entre os eucariotos (Wrzaczek, et al., 2007). Todos *GSKs* contêm um resíduo tirosina totalmente conservado entre os domínios VII e VIII que

é essencial para atividade quinase (Hughes, et al., 1993). Apesar do nome, sabe-se que *GSK-3* é uma proteína serina-treonina quinase, não receptora, multifuncional que atua como regulador em várias vias de sinalização. Abordagens genéticas e bioquímicas indicam que os diferentes *GSKs* de plantas estão envolvidos em diversos processos, incluindo a sinalização, resposta de desenvolvimento e estresse (Jonak & Hirt, 2002; Yoo, et al., 2006; Wrzaczek, et al., 2007). Nas células em repouso, *GSK3* é uma quinase constitutivamente ativa que fosforila uma ampla gama de substratos de proteína para inibir diretamente suas atividades bioquímicas. Sinais de estresse ativam o *GSK* por meio de diferentes formas e substratos são fosforilados e a atividade é regulada. Como exemplo, o produto do gene *WIN* (um *GSK-3* induzido por ferida – *wound-induced GSK-3*), membro da família gênica *GSK3* de alfafa (*Medicago sativa* cv. Europa) foi rapidamente induzido pela resposta a injúria mecânica (Jonak, et al., 2000), o gene *MsK1* de *M. sativa* mostra uma mudança na atividade em resposta aos elicitores de defesa da planta, sendo regulado pelo elicitor celulase, o qual provocou a degradação da proteína *MsK1* de maneira proteassoma dependente, mostrando que a superexpressão de *MsK1* refletia a suscetibilidade a *Pseudomonas syringae* (Wrzaczek, et al., 2007), o gene *TmGSK1* mostrando-se como regulador positivo, comumente envolvido na tolerância ao sal, aos danos mecânicos, ao hormônio ácido abscísico (ABA), ao frio e a seca em trigo einkorn (*Triticum monococcum* L.) (Xian-Guang, et al., 2011), *ASK0* é um fator regulatório na sinalização de brassinosteróides (BRs) (De Rybel, et al., 2009; Rozhon, et al., 2010) (Rozhon, et al., 2010).

As proteínas sensoras de cálcio do tipo calcineurina B (*CBL* – *calcineurin B-like proteins*) são pequenas proteínas que contêm vários domínios de ligação ao Ca^{2+} , resultando na ativação de quinases envolvidas em respostas a estresses por excesso salino e em sinalização por ABA (Luan, et al., 2002; Hrabak, et al., 2003). Essas proteínas são similares tanto com a subunidade B regulatória de calcineurina quanto o sensor neural de Ca^{2+} (NCS – *neural Ca²⁺ sensor*), em animais (Rusnak & Mertz, 2000). Vários sinais extracelulares podem provocar mudanças na concentração de Ca^{2+} no citoplasma de células vegetais, como: hormônios vegetais, luz, fatores de estresse, elicitores simbióticos e patogênicos (Shi, et al., 1999). Shi e colaboradores (1999) identificaram uma família de proteínas quinases (CIPKs) como proteínas-alvo para *AtCBLs/SOS3*, onde *AtCBL* é altamente induzível por sinais de estresse e *SOS3* (*Salt overly sensitive 3*) desempenha papel de resistência salina em *Arabidopsis*, podendo, CIPK1 e *AtCBL* se associarem de forma Ca^{2+} dependente. O cálcio, assim como qualquer outro mensageiro secundário, pode ser ativado e compartilhado, ao mesmo tempo, por diferentes estresses e vias de sinalização (Batistic & Kudla, 2004). Muitos são os estudos

acerca do complexo *CBL-CIPK* relacionado ao seu papel desempenhado em condições de estresse abiótico, tais como: deficiência de potássio, resposta ao nitrato, salínico, seca, injúria mecânica (Amtmann & Armengaud, 2007; Hu, et al., 2009; Tripathi, et al., 2009; Kabir & Wang, 2010; Piao, et al., 2010).

Proteínas quinase são elementos chave na regulação de vários eventos celulares e tentando tornar a célula mais adaptável as mudanças ambientais (Rajkarnikar, et al., 2006). A proteína serina/treonina quinase (*STK* ou *STPK*) é uma das proteínas mais importantes responsável pela transdução de sinal de defesa (Xu & Deng, 2010), pois podem desempenhar o papel de receptores serina treonina quinase (*RSTK*) um tipo de receptor proteína quinase (*RLK*), interagindo com outras proteínas e participando de vários processos, como a resistência a doenças e a regulação do desenvolvimento (Afzal, et al., 2008), em bactérias, fungos, animais e plantas (Liang, et al., 2011; Zorina, et al., 2011). O domínio *STK* é o maior constituinte do gene *Pto* de tomate que não só interage com as proteínas de avirulência de *Pseudomonas syringae*, como também funciona como um mediador de transdução de sinal (Martin, et al., 1993). Além disso, esse domínio *STK*, também está contido no gene *Xa21* de arroz, conferindo resistência a *X. oryzae* pv. *oryzae* (Song, et al., 1995). Como já se espera, esses domínios *STKs* são conservados entre as espécies de plantas, e tem sido isolado em várias espécies de plantas. Uma análise comparativa entre esses RGAs mostrou que eles estão sob uma rápida evolução e são polimórficos entre genótipos suscetíveis e resistentes (Xu, et al., 2008). Xu e colaboradores (2010), encontraram um novo tipo de gene *STK* de cópia única e com menos de 40% de similaridade com qualquer outra classe de *STK*. Cao e colaboradores (2011) identificaram uma proteína serina-treonina quinase putativa, *Stpk-V*, que está localizado no locus *Pm21* do braço curto do cromossomo 6V de *Haynaldia villosa* (L.) Schur, o qual confere resistência ao oídio do trigo. Esta proteína *Stpk-V*, compartilha um espectro de resistência ao oídio semelhante ao gene *Pm21* transferido ao trigo, porém não se sabe, ainda, se *Stpk-V* é o único gene necessário para conferir o fenótipo de resistência.

Genes *STKs* polimórficos combinado com outros genes de imunidade da planta, fornecem muitos recursos a serem utilizados para defesa contra o ataque de patógenos (Xu & Deng, 2010). Semelhanças generalizadas têm derivado tanto da conservação evolutiva quanto da convergência, o que sugere eventos comuns de sinalização, constitui a base de cascatas de defesa (Afzal, et al., 2008).

Outro subgrupo da família de *RLKs*, os receptores ricos em cisteína (*CRKs*), é composto por pouco mais de 40 membros. Um deles é o DUF26 (Domínio de função

desconhecida 26 ou *domain of unknown function 26*), com motivos ricos em repetições de cisteína. Esses resíduos conservados de cisteína podem funcionar na manutenção da estrutura tridimensional e formar um motivo *zinc finger* para mediar interações proteína-proteína ou a percepção de alterações redox no espaço extracelular durante respostas de defesa vegetal (Chen Z. , 2001; Zhang, et al., 2009; Wrzaczek, et al., 2010). Os *CRKs* são transcricionalmente induzidos pelo estresse oxidativo, ataque de patógenos e aplicação de ácido salicílico (SA) (Czernic, et al., 1999; Chen Z. , 2001; Chen, et al., 2004). A superexpressão constitutiva de *CRK5*, em plantas transgênicas, levou ao aumento de resistência a *Pseudomonas syringae* pv. tomato DC3000, além de reforçar o crescimento das folhas. Porém, ainda em plantas transgênicas, uma alta expressão de *CRK5* sob controle do promotor induzido por esteróides, desencadeou uma resposta de hipersensibilidade e morte celular (Chen, et al., 2003). Acharya e colaboradores (2007), sugerem que o gene *CRK13*, outro membro da família *RLK*, positivamente regulado, leva a resposta de hipersensibilidade associada à morte celular e induz a defesa contra patógenos por causar o acúmulo de SA.

O perfil de expressão de cinco RGAs-candidatos, aqui avaliados, mostrou um padrão de expressão gênica, positivamente regulado quando comparados ao seu controle. Com exceção do gene *LRR-TM-PK*, que teve seu par de *primers* desenhado de forma diferente dos demais RGAs analisados, não levando em consideração o desenho fora de região de domínio conservado, todos os outros cinco RGAs (*EDR1*, *GSK*, *CIPKs*, *STPKs* e *CRKs*) foram positivamente regulados, comparado com a expressão basal de suas amostras controle (Figura 37). Entretanto, os níveis de diferença de expressão são muito sutis, embora significativos, variando de 1,1 a, no máximo, 1,4. Este resultado está de acordo com dados da literatura que mostram que muitos genes R são expressos constitutivamente em um nível baixo, estando presentes, constantemente, em plantas não-desafiadas por um patógeno, isto é, saudáveis, facilitando uma resposta mais rápida ao ataque microbiano (Hammond-Kosack & Jones, 1997).

Embora pareça concebível que as plantas tiveram que desenvolver uma estratégia evolutiva e adaptativa as adversidades ambientais, como estresses bióticos e abióticos, ainda não é claro como as plantas conseguem respostas integradas para a sua defesa, como organismos sésseis, acumulando indicações da existência de redes de sinalização de escala livre (Batistic & Kudla, 2004). Isso é visível quando plantas são feridas ou atacadas por herbívoros e patógenos, e elas induzem uma cicatrização específica e respostas de defesa, onde um conjunto de moléculas sinalizadoras é sintetizado a fim de coordenar e especificar essas respostas. O equilíbrio entre o ácido jasmônico (JA), etileno (ET), ácido salicílico (SA) e ácido abscísico (ABA), existente

nessas interações de resposta de defesa, possuem um papel crítico nessa regulação e nas vias de sinalização desses hormônios. Dessa forma, uma extensiva ativação cruzada (*crosstalk*) durante a defesa e resposta a ferimentos, ocorre entre esses hormônios, agindo em sinergia, aditivamente ou antagonicamente para ativar a expressão de genes de defesa (Manavella, et al., 2008).

Como demonstrado, muitos são os fatores que podem influenciar alterações nos mecanismos de defesa da planta, ocasionando o desencadeamento de reações adversas para sua proteção e manutenção. Como os experimentos foram desenvolvidos em casa de vegetação, fatores como temperatura e umidade, não puderam ser muito bem controlados, o que pode ter interferido na expressão gênica dessas plantas de *A. stenosperma*. Para tanto esses dados serão novamente avaliados, em bioensaios conduzidos em câmara de crescimento, com o método de folhas destacadas (Moraes & Salgado, 1982) de *A. stenosperma* desafiadas e não desafiadas com o fungo *P. personata*, com temperatura, umidade e fotoperíodos controlados, tentando verificar se haverá alteração de expressão.

7. CONCLUSÃO

Neste estudo, um número substancial de genes transcritos (362.631) foi gerado a partir do pirossequenciamento massal de folhas de *A. stenosperma* inoculadas e não inoculadas com *P. personata*, o agente causal da mancha preta do amendoim. A partir da análise deste banco de dados, foram gerados 17.912 unigenes, dos quais, 7.723 constituíam *contigs* com dois ou mais genes transcritos. A análise de similaridade de sequência com banco de dados públicos (BLASTx) identificou 52% desta sequências as quais classificadas em famílias protéicas baseadas em funções putativas (GO – ontologia gênica).

Dentre as anotações obtidas, 50 fatores de transcrição foram identificados, sendo três deles, bZIP, MYB e AP2/EREBP os mais representados, 48 RGAs e vários genes envolvidos em processos de defesa da planta. Dentre eles, 16 genes-candidatos, sendo cinco RGAs, foram avaliados através de RT-qPCR, quanto ao seu nível de expressão mediante a presença do patógeno.

A variação de expressão encontrada entre os genes-candidatos avaliados e os descritos em literatura mostra que, os membros dessas diversas famílias ou subfamílias dos genes estudados, apesar de possuírem domínios conservados em suas estruturas, podem exercer funções distintas das conhecidas. Isso por que o processo de co-evolução entre plantas e patógenos é um processo contínuo e de longo prazo, fazendo com que pressões seletivas do meio sobre os organismos envolvidos possam, aos poucos, suprimir velhas funções, mutar ou deletar algumas regiões do DNA, conduzindo a aparição de novas funções, porém sem, em alguns casos, apagar a informação evolutiva comum compartilhada entre os organismos ou grupos de organismos existentes.

Esta é a primeira análise massal do transcrito de espécie silvestre de *Arachis* (*A. stenosperma*) inoculada com *P. personata*, o agente causal da mancha preta do amendoim, realizada até o momento. Os dados gerados neste trabalho promoverão o enriquecimento do banco de dados de genes transcritos de *A. stenosperma*, os quais poderão ser utilizados para descoberta de genes-candidatos a resistência e a outras características de interesse agrônômico e novos marcadores moleculares (SNPs e SSRs), que uma vez mapeados, poderão auxiliar na identificação de QTLs no melhoramento da cultura do amendoim.

O desenvolvimento de novos marcadores moleculares contribuirá no enriquecimento dos mapas genéticos existente, gerando ferramentas genéticas e genômicas mais informativas, permitindo assim, a identificação de genes ortólogos através da análise de sintenia (Bertioli, et al., 2009), o que contribuirá no desenvolvimento dos escassos recursos genômicos entre as espécies de leguminosas.

Além disso, uma melhor caracterização e compreensão dos genes envolvidos nos mecanismos de defesa oriundos da interação *A. stenosperma* e *P. personata*, possuem importância prática e agrônômica, visando contribuir não somente para a produção de uma cultivar de amendoim mais resistente com maior capacidade de sobrevivência ao estresse, reduzindo o estresse ocasionado e aumentando da produtividade da cultura, com maior rendimento econômico.

8. REFERÊNCIAS BIBLIOGRÁFICAS

- Acharya, B. R., Raina, S., Magbool, S. B., Jagadeeswaran, G., Mosher, S. L., Appel, H. M., et al. (2007). Overexpression of CRK13, an *Arabidopsis* cysteine-rich receptor-like kinase, results in enhanced resistance to *Pseudomonas syringae*. *The Plant Journal: for cell and molecular biology*, 50 (3), pp. 488-499.
- Afzal, A. J., Wood, A. J., Lightfoot, D. A. (2008). Plant receptor-like serine threonine kinases: roles in signaling and plant defense. *Molecular Plant-Microbe Interactions*, 21 (5), pp. 507-517.
- Alagna, F., D'Agostino, N., Torchia, L., Servili, M., Rao, R., Pietrella, M., et al. (2009). Comparative 454 pyrosequencing of transcriptome from two olive genotypes during fruit development. *BMC Genomics*, 10 (399), p. 15.
- Alderman, S. C., & Nutter Jr, F. W. (1994). Effect of temperature and relative humidity on development of *Cercosporidium personatum* on peanut in Georgia. *Plant Disease*, 78, pp. 690-694.
- Almeida, A. M. (2001). Principais doenças fúngicas do amendoim e controle. In: J. R. Scarpellini, Z. A. Ramiro, A. S. Santos, G. A. Paula e Silva, M. Bergamashi, *Anais da IV Reunião Itinerante de Fitossanidade do Instituto Biológico e do V encontro sobre pragas e doenças do cafeeiro* (pp. 83-93). Ribeirão Preto - SP: Instituto Biológico.
- Almeida, P. S., & Saddi, V. A. (2007). Monitoramento de doença residual mínima em leucemia mielóide crônica por PCR em tempo real. *Revista Brasileira de Hematologia e Hemoterapia*, 29 (4), pp. 382-386.
- Alves, D. M., Pereira, R. W., Leal-Bertioli, S. C., Moretzsohn, M. C., Guimarães, P. M., Bertioli, D. J., et al. (2008). Development and use of single nucleotide polymorphism markers for candidate resistance genes in wild peanuts (*Arachis* spp). *Genetics and Molecular Research*, 7 (3), pp. 631-642.
- Amtmann, A., & Armengaud, P. (2007). The role of calcium sensor-interacting protein kinases in plant adaptation to potassium-deficiency: a new answers to old questions. *Cell Research*, 17, pp. 483-485.
- Andreote, F. D. (2011). Análises genômica e transcriptômica de *Methylobacterium mesophilicum* SR1.6/6 em interação com a planta hospedeira. *Dissertação apresentada ao Departamento de Genética e Melhoramento de plantas da Universidade de São Paulo, Escola Superior de Agricultura "Luiz de Queiroz" - ESALQ, Piracicaba, Brasil*, p. 82.
- Applied Biosystem, L. T. (s/d). *7300 Real-Time PCR System*. Acesso em Setembro de 2011, disponível em https://products.appliedbiosystems.com/ab/en/US/htdocs/productMgr/images/7300_Real_Time_big.jpg
- Applied Biosystem, L. T. (s/d). *Apostila de PCR em tempo real*. Acesso em Novembro de 2010, disponível em <http://www.appliedbiosystems.com.br/site/material/d1jg4lmm.pdf>

- Aquino, V. M., Shokes, F. M., Gorbet, D. W., Nutter Jr., F. W. (1995). Late leaf spot progression on peanut as affected by components of partial resistance. *Plant Disease*, *79*, pp. 74-78.
- Arce, A. L., Raineri, J., Capella, M., Cabello, J. V., Chan, R. L. (2011). Uncharacterized conserved motifs outside the HD-Zip domain in HD-Zip subfamily I transcription factors; a potential source of functional diversity. *BMC Plant Biology*, *11* (42), p. 19.
- Ariel, F. D., Manavella, P. A., Dezar, C. A., Chan, R. L. (2007). The true story of the HD-Zip family. *Trends in Plant Science*, *12* (9), pp. 419-426.
- Artico, S., Nardeli, S. M., Brilhante, O., Grossi-de-Sa, M. F., Alves-Ferreira, M. (2010). Identification and evaluation of new reference genes in *Gossypium hirsutum* for accurate normalization of real-time quantitative RT-PCR data. *BMC Plant Biology*, *10* (49), p. 12.
- Audic, S., & Claverie, J. (1997). The significance of digital gene expression profiles. *Genome Research*, *7*, pp. 986-995.
- Barakat, A., DiLoreto, D. S., Zhang, Y., Smith, C., Baier, K., Powell, W. A., et al. (2009). Comparison of the transcriptome of American chestnut (*Castanea dentata*) and Chinese chestnut (*Castanea mollissima*) in response to the chestnut blight infection. *BMC Plant Biology*, *9* (51), p. 11.
- Barreto, M. (1997). Doenças do amendoim. In: H. Kimati, L. Amorim, A. Bergamim Filho, L. E. Camargo, & J. A. Rezende (Eds.), *Manual de Fitopatologia - Doenças das Plantas Cultivadas* (3ª ed., Vol. 2, pp. 70-81). São Paulo: Ceres.
- Batistic, O., & Kudla, J. (2004). Integration and channeling of calcium signaling through the CBL calcium sensor/CIPK protein kinase network. *Planta*, *219* (6), pp. 915-924.
- Bays, N. W., Wilhovsky, S. K., Goradia, A., Hodgkiss-Harlow, K., Hampton, R. Y. (2001). HRD4/NPL4 is required for the proteasomal processing of ubiquitinated ER proteins. *Molecular Biology of the Cell*, *12*, pp. 4114-4128.
- Benito, E. P., Prins, T., Van Kan, J. A. (1996). Application of differential display RT-PCR to the analysis of gene expression in a plant-fungus interaction. *Plant Molecular Biology*, *32*, pp. 937-947.
- Bent, A. F., & Mackey, D. (2007). Elicitors, effectors, and R genes: The new paradigm and a lifetime supply of questions. *Annual Review of Phytopathology*, *45*, pp. 399-436.
- Bertioli, D. J., Moretzsohn, M. C., Madsen, L. H., Sandal, N., Leal-Bertioli, S. C., Guimarães, P. M., et al. (2009). An analysis of synteny of *Arachis* with *Lotus* and *Medicago* sheds new light on the structure, stability and evolution of legume genomes. *BMC Genomics*, *10* (45), p. 11.
- Bertioli, D. J., Seijo, G., Freitas, F. O., Valls, J. F., Leal-Bertioli, S. C., Moretzsohn, M. C., et al. (2011). An overview of peanut and its wild relatives. *Plant Genetic Resources: Characterization and Utilization*, *9* (1), pp. 134-149.

- Bossie, M. A., DeHoratius, C., Barcelo, G., Silver, P. (1992). A mutant nuclear protein with similarity to RNA binding proteins interferes with nuclear import in yeast. *Molecular Biology of the Cell*, 3, pp. 875-893.
- Boutrot, F., Segonzac, C., Chang, K. N., Qiao, H., Ecker, J. R., Zipfel, C., et al. (2010). Direct transcriptional control of the *Arabidopsis* immune receptor FLS2 by the ethylene-dependent transcription factors EIN3 and EIL1. *PNAS*, 107 (32), pp. 14502-14507.
- Brasileiro, A. C., Santos, C. M., Morgante, C. V., Martins, A. C., Silva, F. R., Araújo, A. C., et al. (2008). Análise in silico da expressão gênica diferencial de plantas de *Arachis magna* submetidas a estresse hídrico. *Boletim de Pesquisa e Desenvolvimento* 229, p. 22.
- Braun, U. (1995). Miscellaneous notes on phytopathogenic hyphomycetes (II). *Mycotaxon*, 55, pp. 223-241.
- Bräutigam, A., Mullick, T., Schliesky, S., Weber, A. P. (2011). Critical assessment of assembly strategies for non-model species mRNA-Seq data and application of next-generation sequencing to the comparison of C3 and C4 species. *Journal of Experimental Botany*, 62 (9), pp. 3093-3102.
- Buggs, R. J., Chamala, S., Wu, W., Gao, L., May, G. D., Schnable, P. S., et al. (2010). Characterization of duplicate gene evolution in the recent natural allopolyploid *Tragopogon miscellus* by next-generation sequencing and Sequenom iPLEX MassARRAY genotyping. *Molecular Ecology*, 19 (1), pp. 132-146.
- Bustin, S. A., Benes, V., Garson, J. A., Hellems, J., Huggett, J., Kubista, M., et al. (2009). The MIQE guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry*, 55 (4), pp. 1-12.
- Butler, D. R., Wadia, K. D., Jadhav, D. R. (1994). Effects of leaf wetness and temperature on late leaf-spot infection on groundnut. *Plant Pathology*, 43, 112-120.
- Cantonwine, E. G., Culbreath, A. K., Holbrook, C. C., Gorbett, D. W. (2008). Disease Progress of Early Leaf Spot and Components of Resistance to *Cercospora arachidicola* and *Cercosporidium personatum* in Runner-Type Peanut Cultivars. *Peanut Science*, 35, p. 10.
- Cao, A., Xing, L., Wang, X., Yang, X., Wang, W., Sun, Y., et al. (2011). Serine/threonine kinase gene Stpk-V, a key member of powdery mildew resistance gene Pm21, confers powdery mildew resistance in wheat. *Proceedings of the National Academy of Sciences of the United States of America*, 108 (19), pp. 7727-7732.
- Carvalho, M. C., & Silva, D. C. (2010). Sequenciamento de DNA de nova geração e suas aplicações na genômica de plantas. *Ciência Rural*, 40 (3), pp. 735-744.
- Caser, D. V., & Anefalos, L. C. (1994). Evolução regional das principais atividades agrícolas do estado de São Paulo, 1969 a 1992. Série Informações estatísticas da agricultura, (2ª ed., Vol. 5, p. 60). São Paulo: Instituto de Economia Agrícola.

- Chen, K., Du, L., Chen, Z. (2003). Sensitization of defense responses and activation of programmed cell death by a pathogen-induced receptor-like protein kinase in *Arabidopsis*. *Plant Molecular Biology*, 53, pp. 61-74.
- Chen, K., Fan, B., Du, L., Chen, Z. (2004). Activation of hypersensitive cell death by pathogen-induced receptor-like protein kinases from *Arabidopsis*. *Plant Molecular Biology*, 56, pp. 271-283.
- Chen, Z. (2001). A superfamily of proteins with novel cystein-rich repeats. *Plant Physiology*, 126, pp. 473-476.
- Cheng, D. W., Lin, H., Takahashi, Y., Walker, M. A., Civerolo, E. L., Stenger, D. C., et al. (2010). Transcriptional regulation of the grape cytochrome P450 monooxygenase gene CYP736B expression in response to *Xylella fastidiosa* infection. *BMC Plant Biology*, 10, p. 14.
- Cheung, F., Haas, B. J., Goldberg, M. D., May, G. D., Xiao, Y., Town, C. D., et al. (2006). Sequencing *Medicago truncatula* expressed sequenced tags using 454 Life Sciences technology. *BMC Genomics*, 7 (272), p. 10.
- Chirinos, F. V. (2011). *Breeding for Early Maturity in Peanuts (Arachis hypogaea L.) using Traditional Methods and Marker Assisted Selection (MAS)*. North Carolina State University.
- Christiansen, K. M., Gu, Y., Rodibaugh, N., Innes, R. W. (2011). Negative regulation of defence signalling pathways by the EDR1 protein kinase. *Molecular Plant Pathology*, 12 (8), pp. 746-758.
- Conab. (2011b). *Conab - Acompanhamento da Safra Brasileira - Grãos - Safra 2010/2011 - 11º levantamento*. Acesso em Agosto de 2011, disponível em Conab: http://www.conab.gov.br/OlalaCMS/uploads/arquivos/11_08_09_11_44_03_boletim_agosto-2011..pdf
- Conab. (2011a). *Conab - Indicadores da Agropecuária - Ano XX nº 5, 6 e 7*. Acesso em Agosto de 2011, disponível em Conab: <http://www.conab.gov.br/conab/Main.php?MagID=3&MagNo=125>
- Cordeiro, M. C., & Sá, M. F. (1999). Biotecnologia e resistência a patógenos - Interação planta-patógeno e o uso da biotecnologia na obtenção de plantas resistentes. *Biotecnologia Ciência e Desenvolvimento*, Ano 2 (10), pp. 34-39.
- Corrêa, L. G., Riaño-Pachón, D. M., Schrago, C. G., Santos, R. V., Mueller-Roeber, B., Vincentz, M., et al. (2008). The role of bZIP transcription factors in green plant evolution: Adaptive features emerging from four founder genes. *PLoS One*, 3 (8), p. 16.
- Corrent, A. R., Schwarz, L. L., Oster, A. H., Moraes, M. G., Bender, R. J. (2009). Caracterização de genes diferencialmente expressos na interação entre maçã 'Fuji' e *Botryosphaeria dothidea*. *Revista Brasileira de Fruticultura*, 31 (1), pp. 268-272.
- Coutinho, W. M., & Suassuna, N. D. (Dezembro de 2008). Principais doenças do Amendoimzeiro no Nordeste do Brasil. *Documentos* 206, p. 26.

- Crous, P. W., Kang, J. C., & Braun, U. (2001). A phylogenetic redefinition of anamorph genera in *Mycosphaerella* based on ITS rDNA sequence and morphology. *Mycologia*, 93 (6), pp. 1081-1101.
- Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K., Scheible, W. (2005). Genome-wide identification and testing of superior references genes for transcript normalization in *Arabidopsis*. *Plant Physiology*, 139, pp. 5-17.
- Czernic, P., Visser, B., Sun, W., Savouré, A., Deslandes, L., Marco, Y., et al. (1999). Characterization of an *Arabidopsis thaliana* receptor-like protein kinase gene activated by oxidative stress and pathogen attack. *The Plant Journal: for cell and molecular biology*, 18 (3), pp. 321-327.
- Damicone, J. P., & Melouk, H. A. (s/d). *Foliar Diseases of Peanut*. Acesso em Novembro de 2010, disponível em Oklahoma Cooperative Extension Service: <http://www.osuextra.com>
- Daniel, X., Lacomme, C., Morel, J., Roby, D. (1999). A novel MYB oncogene homologue in *Arabidopsis thaliana* related to hypersensitive cell death. *The Plant Journal*, 20 (1), pp. 57-66.
- Dassanayake, M., Haas, J. S., Bohnerr, H. J., Cheeseman, J. M. (2009). Shedding light on an extremophile lifestyle through transcriptome. *New Phytologist*, 183 (3), pp. 764-775.
- De Rybel, B., Audenaert, D., Vert, G., Rozhon, W., Mayerhofer, J., Peelman, F., et al. (2009). Chemical inhibition of a subset of *Arabidopsis thaliana* GSK3-like kinases activates brassinosteroid signaling. *Chemistry & Biology*, 16, pp. 594-604.
- DeHoratius, C., & Silver, P. A. (1996). Nuclear transport defects and nuclear envelope alterations are associated with mutation of the *Saccharomyces cerevisiae* NPL4 gene. *Molecular Biology of the Cell*, 7, pp. 1835-1855.
- Derveaux, S., Vandesomepele, J., Hellemans, J. (2010). How to do successful gene expression analysis using real-time PCR. *Methods*, 50 (4), pp. 227-230.
- Dezar, C. A., Giacomelli, J. I., Manavella, P. A., Ré, D. A., Alves-Ferreira, M., Baldwin, I. T., et al. (2011). HAHB10, a sunflower HD-ZIP II transcription factor, participates in the induction of flowering and in the control of phytohormone-mediated responses to biotic stress. *Journal of Experimental Botany*, 62 (3), pp. 1061-1076.
- Diener, A. C., & Ausubel, F. M. (2005). Resistance to *Fusarium oxysporum* 1, a dominant *Arabidopsis* disease-resistance gene, is not race specific. *Genetics*, 171, pp. 305-321.
- Ding, X., Hou, X., Xie, K., Xiong, L. (2009). Genome-wide identification of BURP domain-containing genes in rice reveals a gene family with diverse structures and responses to abiotic stress. *Planta*, 230, pp. 149-163.
- Dubey, A., Farmer, A., Schlueter, J., Cannon, S. B., Abernathy, B., Tuteja, R., et al. (2011). Defining the transcriptome assembly and its use for genome dynamics and transcriptome profiling studies in Pigeonpea (*Cajanus cajan* L.). *DNA Research*, pp. 1-12.

Dwivedi, S. L., Crouch, J. H., Nigam, S. N., Ferguson, M. E., Paterson, A. H. (2003). Molecular breeding of groundnut for enhanced productivity and food security in the semi-arid tropics: Opportunities and Challenges. *Advances in Agronomy*, 80, pp. 153-221.

DWSPL, & Cattlin, N. (s/d). *Late leaf spot (Phaeoisariopsis personata) infected peanut crop*. Acesso em Outubro de 2010, disponível em DWSPL: <http://www.dwpicture.com.au/picture.asp?picture=109293&cat=Plant%20Diseases&cat2=Plant%20Diseases&cat3=Leaf%20Spot>

Elhiti, M., & Stasolla, C. (2009). Structure and function of homodomain-leucine zipper (HD-Zip) proteins. *Plant Signaling & Behavior*, 4 (2), pp. 86-88.

Ellegren, H. (2008). Sequencing goes 454 and takes large-scale genomics into the wild. *Molecular Ecology*, p. 3.

Ellis, J. G., & Jones, D. A. (2003). Plant disease resistance genes. In: A. B. Ezekowitz, & J. A. Hoffmann (Eds.), *Innate Immunity*, pp. 27-45: Humana Press.

Ellis, J., & Jones, D. (1998). Structure and function of proteins controlling strain-specific pathogen resistance in plants. *Current Opinion in Plant Biology*, 1, pp. 288-293.

Elvira, M. I., Galdeano, M. M., Gilardi, P., García-Luque, I., Serra, M. T. (2008). Proteomic analysis of pathogenesis-related proteins (PRs) induced by compatible and incompatible interactions of pepper mild mottle virus (PMMoV) in *Capsicum chinense* L³ plants. *Journal of Experimental Botany*, 59 (6), pp. 1253-1265.

Emrich, S. J., Barbazuk, W. B., Li, L., Schnable, P. S. (2006). Gene discovery and annotation using LCM-454 transcriptome sequencing. *Genome Research*, 16, p. 5.

Evrogen. (s/d). *SMART cDNA synthesis*. Acesso em Novembro de 2011, disponível em <http://www.evrogen.com/technologies/SMART.shtml#>

Fávero, A. P., & Suassuna, T. F. (2007). *Obtenção de três eventos-elite através da hibridação entre A. hypogaea e anfidiplóides sintéticos e sua inserção no programa de melhoramento da Embrapa*. Brasília: Embrapa Recursos Genéticos e Biotecnologia.

Fávero, A. P., Moraes, S. A., Garcia, A. A., Valls, J. F., Vello, N. A. (2009). Characterization of rust, early and late leaf spot resistance in wild and cultivated peanut germplasm. *Scientia Agricola*, 66 (1), pp. 110-117.

Fernández, A., & Krapovickas, A. (1994). Cromosomas y evolución en *Arachis* (Leguminosae). *Bonplandia*, 8 (1/4), pp. 186-220.

Fleige, S., & Pfaffl, M. W. (2006). RNA integrity and the effect on the real-time qRT-PCR performance. *Molecular Aspects of Medicine*, 27, pp. 126-139.

Flor, H. H. (1971). Current status of the gene-for-gene concept. *Annual Review of Phytopathology*, 9, pp. 275-296.

Frank, M. R., Deyneka, J. M., Schuler, M. A. (1996). Cloning of wound-induced cytochrome P450 monooxygenases expressed in pea. *Plant Physiology*, 110 (3), pp. 1035-1046.

- Franssen, S. U., Shrestha, R. P., Bräutigam, A., Bornberg-Bauer, E., Weber, A. P. (2011). Comprehensive transcriptome analysis of the highly complex *Pisum sativum* genome using next generation sequencing. *BMC Genomics*, 12 (227), p. 16.
- Freire, F. C., & Braun, U. (2009). Hifomicetos Cercosporoides Associados a Plantas do Estado do Ceará. *Revista Ciência Agronômica*, 40 (1), pp. 150-156.
- Freitas, S. M., & Margarido, M. A. (2003). Fatores que influenciam o cultivo do amendoim das águas no estado de São Paulo: uma análise econométrica. *Revista de Economia Agrícola*, 50 (2), pp. 29-40.
- Frye, C. A., Tang, D., Innes, R. W. (2001). Negative regulation of defense response in plants by a conserved MAPKK kinase. *PNAS*, 98 (1), pp. 373-378.
- Gachomo, E. W., Shonukan, O. O., Kotchoni, S. O. (2003). The molecular initiation and subsequent acquisition of disease resistance in plants. *African Journal of Biotechnology*, 2 (2), pp. 26-32.
- Gan, D., Jiang, H., Zhang, J., Zhao, Y., Zhu, S., Cheng, B., et al. (2011). Genome-wide analysis of BURP domain-containing genes in maize and sorghum. *Molecular Biology Reports*, 38 (7), pp. 4553-4563.
- Gan, Y., Zhou, Z., An, L., Bao, S., Forde, B. G. (March de 2011). A comparison between Northern Blotting and quantitative real-time PCR as a means of detecting the nutritional regulation of genes expressed in roots of *Arabidopsis thaliana*. *Agricultural Sciences in China*, 10 (3), pp. 335-342.
- Garg, R., Patel, R. K., Jhanwar, S., Priya, P., Bhattacharjee, A., Yadav, G., et al. (2011). Gene discovery and tissue-specific transcriptome analysis in chickpea with massively parallel pyrosequencing and web resource development. *Plant Physiology*, 156 (4), pp. 1661-1678.
- Gilles, A., Megléc, E., Pech, N., Ferreira, S., Malausa, T., Martin, J., et al. (2011). Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC Genomics*, 12 (245), p. 11.
- Glazebrook, J. (2005). Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual Reviews Phytopathology*, 43, pp. 205-227.
- Godoy, I. J., Moraes, S. A., Zanotto, M. D., Santos, R. C. (1999). Melhoramento do amendoim. In: A. Borem (Ed.), *Melhoramento de espécies cultivadas* (pp. 51-94). Viçosa: UFV.
- Göhre, V., & Robatzek, S. (2008). Breaking the barriers: microbial effector molecules subvert plant immunity. *Annual Reviews Phytopathology*, 46, pp. 189-215.
- Gómez-Gómez, L., & Boller, T. (2002). Flagellin perception: a paradigm for innate immunity. *Trends in Plant Science*, 7 (6), pp. 251-256.
- Gómez-Gómez, L., & Boller, T. (2000). FLS2: An LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. *Molecular Cell*, 5, pp. 1003-1011.

- González-Mendoza, D. (2007). El complejo enzimático citocromo P450 en las plantas. *Revista internacional de contaminación ambiental*, 23 (4), pp. 177-183.
- Govind, G., Thamme Gowda, H. V., Kalaiarasi, P. J., Iyer, D. R., Muthappa, S. K., Nese, S., et al. (2009). Identification and functional validation of a unique set of drought induced genes preferentially expressed in response to gradual water stress in peanut. *Molecular Genetics and Genomics*, 281 (6), pp. 591-605.
- Graham, T. L., & Graham, M. Y. (1991). Cellular coordination of molecular responses in plant defense. *Molecular plant-microbe interactions*, 4 (5), pp. 415-422.
- Granger, C., Coryell, V., Khanna, A., Keim, P., Vodkin, L., Shoemaker, R. C., et al. (2002). Identification, structure, and differential expression of members of a BURP domain containing protein family in soybean. *Genome*, 45 (4), pp. 693-701.
- Greenshields, D. L., & Jones, J. D. (2008). Plant pathogen effectors: getting mixed messages. *Current Biology: CB*, 18 (3), pp. R128-130.
- Gu, Y., & Innes, R. W. (2011). The keep on going (KEG) protein of *Arabidopsis* recruits the enhanced disease resistance 1 protein to trans-golgi network/early endosome vesicles. *Plant Physiology*, 155 (4), pp. 1827-1838.
- Guimarães, P. M., Gasmeur, O., Proite, K., Leal-Bertioli, S. C., Seijo, G., Chaine, C., et al. (2008). BAC libraries construction from the ancestral diploid genomes of the allotetraploid cultivated peanut. *BMC Plant Biology*, 8 (14), p. 8.
- Guo, B., Chen, X., Dang, P., Scully, B. T., Liang, X., Holbrook, C. C., et al. (2008). Peanut gene expression profiling in developing seeds at different reproduction stages during *Aspergillus parasiticus* infection. *BMC Developmental Biology*, 8 (12), p. 16.
- Guo, B., Chen, X., Hong, Y., Liang, X., Dang, P., Brenneman, T., et al. (2009). Analysis of gene expression profiles in leaf tissues of cultivated peanuts and development of EST-SSR markers and gene discovery. *International Journal of Plant Genomics*, 2009, p. 14.
- Gutierrez, L., Mauriat, M., Pelloux, J., Bellini, C., Wuytswinkel, O. V. (2008). Towards a systematic validation of references in real-time RT-PCR. *The Plant Cell*, 20, pp. 1734-1735.
- Hagan, A. (1998). Foliar diseases of peanuts. *Alabama Cooperative Extension System - ANR-369*, p. 4.
- Hammond-Kosack, K. E., & Jones, J. D. (1997). Plant disease resistance genes. *Annual Review of Plant Physiology and Plant Molecular Biology*, 48, pp. 575-607.
- Harris, J. C., Hrmova, M., Lopato, S., Langridge, P. (2011). Modulation of plant growth by HD-Zip class I and II transcription factors in response to environmental stimuli. *New Phytologist*, 190, pp. 823-837.
- Hattori, J., Boutilier, K. A., Campagne, M. M., Miki, B. L. (1998). A conserved BURP domain defines a novel group of plant proteins with unusual primary structures. *Molecular & General Genetics*, 259 (4), pp. 424-428.

- Heath, M. C. (2000). Nonhost resistance and nonspecific plant defenses. *Current Opinion in Plant Biology*, 3, pp. 315-319.
- Henry, M., Borland, C. Z., Bossie, M., Silver, P. A. (1996). Potential RNA binding proteins in *Saccharomyces cerevisiae* identified as suppressors of temperature-sensitive mutations in NPL3. *Genetics*, 142 (1), pp. 103-105.
- Hitchcock, A. L., Krebber, H., Fietze, S., Lin, A., Latterich, M., Silver, P. A., et al. (2001). The conserved Npl4 protein complex mediates proteasome-dependent membrane-bound transcription factor activation. *Molecular Biology of the Cell*, 12, pp. 3226-3241.
- Holbrook, C. C., Stephenson, M. G., Johnson, A. W. (2000). Level and geographical distribution of resistance to *Meloidogyne arenaria* in the US peanut germplasm collection. *Crop Science* (40), pp. 1168-1171.
- Hougaard, B. K., Madsen, L. H., Sandal, N., Moretzsohn, M. C., Fredslund, J., Schauser, L., et al. (2008). Legume Anchor Markers Link Syntenic Regions Between *Phaseolus vulgaris*, *Lotus japonicus*, *Medicago truncatula* and *Arachis*. *Genetics Society of America*, 179, pp. 2299-2312.
- Hrabak, E. M., Chan, C. W., Gribskov, M., Harper, J. F., Choi, J. H., Halford, N., et al. (2003). The *Arabidopsis* CDPK-SnRK superfamily of protein kinases. *Plant Physiology*, 132, pp. 666-680.
- Hu, H., Wang, Y., Tsay, Y. (2009). AtCIPK8, a CBL-interacting protein kinase, regulates the low-affinity phase of the primary nitrate response. *The Plant Journal*, 57, pp. 264-278.
- Hughes, K., Nikolakaki, E., Plyte, S. E., Totty, B. F., Woodgett, J. R. (1993). Modulation of the glycogen synthase kinase-3 family by tyrosine phosphorylation. *The EMBO Journal*, 12 (2), pp. 803-808.
- Hutcheson, S. W. (1998). Current concepts of active defense in plants. *Annual Review Phytopathology*, 36, pp. 59-90.
- Index Fungorum, O. (2011). *Species Fungorum*. Acesso em Agosto de 2011, disponível em <http://www.speciesfungorum.org/Names/SynSpecies.asp?RecordID=258363>
- Isin, E. M., & Guengerich, F. P. (2007). Complex reactions catalyzed by cytochrome P450 enzymes. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1770 (3), pp. 314-329.
- Izge, A. U., Mohammed, Z. H., Goni, A. (2007). Levels of variability in groundnut (*Arachis hypogaea* L.) to cercospora leaf spot disease - implication for selection. *African Journal of Agricultural Research*, 2 (4), pp. 182-186.
- Javelle, M., Klein-Cosson, C., Vernoud, V., Boltz, V., Maher, C., Timmermans, M., et al. (2011). Genome-wide characterization of the HD-Zip IV transcription factor family in maize: preferential expression in the epidermis. *Plant Physiology*, 157 (2), pp. 790-803.
- Jenkis, W. A. (1938). Two fungi causing leaf spot of peanut. *Journal of Agricultural Research*, 56, pp. 317-322.

- Jonak, C., & Hirt, H. (2002). Glycogen synthase kinase 3/SHAGGY-like kinases in plants: an emerging family with novel functions. *Trends in Plant Science*, 7 (10), pp. 457-461.
- Jonak, C., Beisteiner, D., Beyerly, J., Hirt, H. (2000). Wound-induced expression and activation of WIN, a novel glycogen synthase kinase 3. *The Plant Cell*, 12, pp. 1467-1475.
- Jones, J. D., & Dangl, J. L. (2006). The plant immune system. *Nature*, 444 (7117), pp. 323-329.
- Jung, J., Won, S. Y., Suh, S. C., Kim, H., Wing, R., Jeong, Y., et al. (2007). The barley ERF-type transcription factor HvRAF confers enhanced pathogen resistance and salt tolerance in *Arabidopsis*. *Planta*, 225 (3), pp. 575-588.
- Jyosthna, M. K., Reddy, N. P., Chalam, T. V., Reddy, G. L. (2004). Morphological and biochemical characterization of *Phaeoisariopsis personata* resistant and susceptible cultivars of groundnut (*Arachis hypogaea*). *Plant Pathology Bulletin*, 13, pp. 243-250.
- Kabir, M. H., & Wang, M. (2010). Response of calcineurin B-like protein kinase gene in tomato to various abiotic stress. *Journal of the Korean Society for Applied Biological Chemistry*, 53 (1), pp. 15-21.
- Kaur, S., Cogan, N. O., Pembleton, L. W., Shinozuka, M., Savin, K. W., Materne, M., et al. (2011). Transcriptome sequencing of lentil based on second-generation technology permits large-scale unigene assembly and SSR marker discovery. *BMC Genomics*, 12 (265), p. 11.
- Khan, M. A. (2009). Importance and use of resistance gene analogs. *The Open Plant Science Journal*, 3, pp. 54-60.
- Khedikar, Y. P., Gowda, M. V., Sarvamangala, C., Patgar, K. V., Upadhyaya, H. D., Varshney, R. K., et al. (2010). A QTL study on late leaf spot and rust revealed one major QTL for molecular breeding for rust resistance in groundnut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics*, 121 (5), pp. 971-984.
- Kim, G., & Tsukaya, H. (2002). Regulation of the biosynthesis of plant hormones by cytochrome P450s. *Journal of Plant Research*, 115, pp. 169-177.
- Kim, Y., Juang, E., Hwang, S., Go, S., Hwang, D. (2004). AP2/EREB transcription factors in rice. *The Plant Pathology Journal*, 20 (1), pp. 41-45.
- Kishore, G. K., Pande, S., Rao, J. N. (2001). Control of Late Leaf Spot of groundnut (*Arachis hypogaea*) by extracts from non-host plant species. *Plant Pathology Journal*, 17 (5), pp. 264-270.
- Kliebenstein, D. J., Rowe, H. E., Denby, K. J. (2005). Secondary metabolites influence *Arabidopsis/Botrytis* interactions: variation in host production and pathogen sensitivity. *The Plant Journal*, 44, pp. 25-36.
- Kochert, G., Halward, T., Branch, W. D., Simpson, C. E. (1991). RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. *Theoretical and Applied Genetics*, 81 (5), pp. 565-570.

- Kochert, G., Stalker, H. T., Gimenes, M., Galgaro, L., Lopes, C. M., Moore, K., et al. (1996). RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). *American Journal of Botany*, 83 (10), pp. 1282-1291.
- Krapovickas, A., & Gregory, W. C. (1994). Taxonomía del género *Arachis* (Leguminosae). 8, 186p.
- Kumar, K. R., & Kirti, P. B. (2011). Differential gene expression in *Arachis diogeni* upon interaction with peanut late leaf spot pathogen, *Phaeoisariopsis personata* and characterization of a pathogen induced cyclophilin. *Plant Molecular Biology*, 75, pp. 497-513.
- Kumar, P. V. (2007). Guide to agricultural meteorological practices (WMO Nº 134), Draft, 3rd Ed., GAMP, Chapter 13B. Disponível em: <http://www.agrometeorology.org>.
- Kumar, S., & Blaxter, M. (2010). Comparing de novo assemblers for 454 transcriptome data. *BMC Genomics*, 11 (571), p. 1-12.
- Leal-Bertioli, S. C., Farias, M. P., Silva, P. I., Guimarães, P. M., Brasileiro, A. C., Bertioli, D. J., et al. (2010). Ultrastructure of initial interaction of *Puccinia arachidis* and *Cercosporidium personatum* with leaves of *Arachis hypogaea* and *Arachis stenosperma*. *Journal of Phytopathology*, 158, pp. 792-796.
- Leal-Bertioli, S. C., José, A. C., Alves-Freitas, D. M., Moretzsohn, M. C., Guimarães, P. M., Nielen, S., et al. (2009). Identification of candidate genome regions controlling disease resistance in *Arachis*. *BMC Plant Biology*, 9 (112), p. 12.
- Leal-Bertioli, S. C., Guedes, L., Guimarães, P. M., Fávero, A. P., Bertioli, D. J. (2005). Isolamento de sequências de *Arachis stenosperma* em resposta à infecção de *Puccinia arachidis* e *Cercosporidium personatum* por cDNA-AFLP. *Boletim de Pesquisa e Desenvolvimento* 94, p. 16.
- Leal-Bertioli, S. C., Guimarães, P. M., Fávero, A. P., Moretzsohn, M. C., Proite, K., Bertioli, D. J., et al. (2003). Amendoim Selvagem - Uma fonte de resistência a pragas. *Revista Biotecnologia Ciência & Desenvolvimento* (31), 116-119.
- Lee, J. M., Roche, J. R., Donaghy, D. J., Thrush, A., Sathish, P. (2010). Validation of reference genes for quantitative RT-PCR studies of gene expression in perennial ryegrass (*Lolium perenne* L.). *BMC Molecular Biology*, 11 (8), p. 14.
- Lee, S., & Choi, D. (2002). Toward functional genomics of plant-pathogen interactions: Isolation and analysis of defense-related genes of hot pepper expressed during resistance against pathogen. *Plant Pathology Journal*, 18 (2), pp. 63-67.
- Li, C., Su, R., Cheng, C., Sanjaya, You, S., Hsieh, T., et al. (2011). Tomato RAV transcription factor is a pivotal modulator involved in the AP2/EREBP-mediated defense pathway. *Plant Physiology*, 156, pp. 213-227.
- Li, Y., L, H., Sun, C., Song, J., Sun, Y., Wu, Q., et al. (2010). EST analysis reveals putative genes involved in glycyrrhizin biosynthesis. *BMC Genomics*, 11 (268), p. 11.

- Liang, C., Zhang, X., Chi, X., Guan, X., Li, Y., Qin, S., et al. (2011). Serine/threonine protein kinase SpkG is a candidate for high salt resistance in the unicellular cyanobacterium *Synechocystis* sp. PCC 6803. *PLoS One*, 6 (5), p. 6.
- Liang, X., Chen, X., Hong, Y., Liu, H., Zhou, G., Li, S., et al. (2009). Utility of EST-derived SSR in cultivated peanut (*Arachis hypogaea* L.) and *Arachis* wild species. *BMC Plant Biology*, 9 (35), p. 9.
- Liberato, J. R., & Shivas, R. G. (Dezembro de 2006). *Late leaf spot of peanut (Passalora personata)*. Acesso em Agosto de 2011, disponível em <http://www.padil.gov.au>
- Life Technologies, A. B. (s/d). *Real-Time Polymerase Chain Reaction*. Acesso em Setembro de 2011, disponível em <http://media.invitrogen.com.edgesuite.net/ab/applications-technologies/real-time-pcr/real-time-polymerase-chain-reaction/index.html?ICID=EDI-Lrn1>
- Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*, 25, pp. 402-408.
- Lopez-Gomez, M., Sandal, N., Stougaard, J., Boller, T. (2011). Interplay of flg22-induced defence responses and nodulation in *Lotus japonicus*. *Journal of Experimental Botany*, First published online September 20, p. 9.
- Lu, Z., Gaudet, D. A., Frick, M., Puchalski, B., Genswein, B., Laroche, A., et al. (2005). Identification and characterization of genes differentially expressed in the resistance reaction in wheat infected with *Tilletia tritici*, the common bunt pathogen. *Journal of Biochemistry and Molecular Biology*, 38 (4), pp. 420-431.
- Luan, S., Kudla, J., Rodriguez-Concepcion, M., Yalovsky, S., Gruissem, W. (2002). Calmodulins and calcineurin B-loke proteins: calcium sensors for specific signal response coupling in plants. *The Plant Cell*, 14, pp. S389-S400.
- Lucas, J. A. (2011). Advances in plant disease and pest management. *The Journal of Agricultural Science*, 149, pp. 91-114.
- Luo, M., Dang, P., Bausher, M. G., Holbrook, C. C., Lee, R. D., Lynch, R. E., et al. (2005b). Identification of transcriptis involved in resistance responses to leaf spot disease caused by *Cercosporidium personatum* in peanut (*Arachis hypogaea*). *Phytopathology*, 95, pp. 381-385.
- Luo, M., Dang, P., Guo, B. Z., He, G., Holbrook, C. C., Bausher, M. G., et al. (2005a). Generation of Expressed Sequence Tags (ESTs) for gene discovery and marker development in cultivated peanut. *Crop Science*, 45, pp. 346-353.
- Manavella, P. A., Dezar, C. A., Bonaventure, G., Baldwin, I. T., Chan, R. C. (2008). HAHB4, a sunflower HD-Zip protein, integrates signals from the jasmonic acid and ethylene pathways during wounding and biotic stress responses. *The Plant Journal*, 56, pp. 376-388.
- Mardis, E. R. (2008a). Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics*, 9, pp. 387-402.

- Mardis, E. R. (2008b). The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, 24 (3), pp. 133-141.
- Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437 (7057), pp. 376-380.
- Martin, C., & Paz-Ares, J. (1997). MYB transcription factor in plants. *Trends in Genetics*, 13 (2), pp. 67-73.
- Martin, G. B., Brommonschenkel, S. H., Chunwongse, J., Frary, A., Ganai, M. W., Spivey, R., et al. (1993). Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science*, 262 (5138), pp. 1432-1436.
- Martins, A. C. (2008). *Isolamento de RNA total de boa qualidade visando a prospecção de genes de tolerância ao estresse hídrico em espécies silvestres de Arachis*. Brasília: Monografia apresentada ao Curso de bacharelada em Ciências Biológicas do Centro de Ciência da Educação e Humanidades da Universidade Católica de Brasília, Brasil.
- Matz, M., Shagin, D., Bogdanova, E., Britanova, O., Lukyanov, S., Diatchenko, L., et al. (1999). Amplification of cDNA ends based on template-switching effect and step-out PCR. *Nucleic Acids Research*, 27 (6), pp. 1558-1560.
- Mazzara, M., & James, D. J. (2000). The influence of photoperiodic growth condition on isolation of RNA from strawberry (*Fragaria x ananassa* Duch.) tissue. *Molecular Biotechnology*, 15 (3), pp. 237-241.
- McDonald, D., Subrahmanyam, P., Gibbons, R. W., Smith, D. H. (1985). *Early and Late Leaf Spots of Groundnut - Information Bulletin nº 21*. Andhra Pradesh, Patancheru P.O., Índia: ICRISAT.
- McDowell, J. M., & Woffenden, B. J. (2003). Plant disease resistance genes: recent insights and potential applications. *Trends in biotechnology*, 21 (4), pp. 178-183.
- Mekhedov, S., Ilárduya, O. M., Ohlrogge, J. (2000). Toward a functional catalog of the plant genome - A survey of genes for lipid biosynthesis. *Plant Physiology*, 122, pp. 389-401.
- Mersmann, S., Bourdais, G., Rietz, S., Robatzek, S. (2010). Ethylene signaling regulates accumulation of the FLS2 receptor and is required for the oxidative burst contributing to plant immunity. *Plant Physiology*, 154, pp. 391-400.
- Meyer, J. D., Silva, D. C., Yang, C., Pedley, K. F., Zhang, C., van de Mortel, M., et al. (2009). Identification and analyses of candidate genes for Rpp4-mediated resistance to asian soybean rust in soybean. *Plant Physiology*, 150, pp. 295-307.
- Moe, K. T., Chung, J., Cho, Y., Moon, J., Ku, J., Jung, J., et al. (2011). Sequence information on simple sequence repeats and single nucleotide polymorphisms through transcriptome analysis of mungbean. *Journal of Integrative Plant Biology*, 53 (1), pp. 63-73.

- Mondego, J. M., Vidal, R. O., Carazzolle, M. F., Tokuda, E. K., Parizzi, L. P., Costa, G. G., et al. (2011). An EST-based analysis identifies new genes and reveals distinctive gene expression features of *Coffea arabica* and *Coffea canephora*. *BMC Plant Biology*, 11 (30), p. 22.
- Moraes, S. A., & Salgado, C. L. (1982). Utilização da técnica de folhas destacadas de amendoim (*Arachis hypogaea* L.) para inoculações com *Cercospora arachidicola* Hori e *Cercosporidium personatum* (Bert. & Curt.) Ell & Ev. *Summa Phytopathologica*, 8, pp. 39-55.
- Moraes, S. A., Godoy, I. J., Martins, A. L., Pereira, J. C., Pedro Júnior, M. J. (1994). Epidemiologia da mancha preta (*Cercosporidium personatum*) em amendoim: resistência, controle químico e progresso da doença. *Fitopatologia Brasileira*, 19, pp. 532-540.
- Moretzsohn, M. C., Leoi, L., Proite, K., Guimarães, P. M., Leal-Bertioli, S. C., Gimenes, M. A., et al. (2005). A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (Fabaceae). *Theoretical and applied genetics*, 111, pp. 1060-1071.
- Morgante, C. V., Guimarães, P. M., Martins, A., Araujo, A. G., Leal-Bertioli, S. M., Bertioli, D. J., et al. (2011). Reference Genes for Quantitative Reverse Transcription-Polymerase Chain Reaction Expression Studies in Wild and Cultivated Peanut. *BMC Plant Biology*, p. 31.
- Morgante, C. V., Martins, A. C., Araujo, A. C., Leal-Bertioli, S. C., Bertioli, D. J., Guimarães, P. M., et al. (2009). Seleção de genes de referência para estudos de expressão gênica no gênero *Arachis*. *Boletim de Pesquisa e Desenvolvimento*, 249, p. 26.
- Morozova, O., Hirst, M., Marra, M. A. (2009). Applications of new sequencing technologies for transcriptome analysis. *Annual Review of Genomics and Human Genetics*, 10, pp. 135-151.
- Mukherjee, K., & Bürglin, T. R. (2006). MEKHLA, a novel domain with similarity to PAS domains, is fused to plant homeodomain-leucine zipper III proteins. *Plant Physiology*, 140 (4), pp. 1142-1150.
- Mukherjee, K., Brocchieri, L., Bürglin, T. R. (2009). A comprehensive classification and evolutionary analysis of plant homeobox genes. *Molecular Biology and Evolution*, 26 (12), pp. 2775-2794.
- Mysore, K. S., & Ryu, C. (2004). Nonhost resistance: how much do we know? *Trends in Plant Science*, 9 (2), pp. 97-104.
- Natarajan, P., & Parani, M. (2011). De novo assembly and transcriptome analysis of five major tissues of *Jatropha curcas* L. using GS FLX titanium platform of 454 pyrosequencing. *BMC Genomics*, 12 (191), p. 12.
- Niggeweg, R., Thurow, C., Kegler, C., Gatz, C. (2000). Tobacco transcription factor TGA2.2 is the main component of as-1-binding factor ASF-1 and is involved in salicylic acid- and auxin-inducible expression of as-1-containing target promoters. *The Journal of Biological Chemistry*, 275 (26), pp. 19897-19905.

- Niu, J., Zhang, L., Hong, D., Wang, Y. (2005). Cloning, characterization and expression of wheat EDR1 (enhanced disease resistance) gene. *Journal of Plant Physiology and Molecular Biology*, 31 (5), pp. 477-484.
- Nobile, P. M., Lopes, C. R., Barsalobres-Cavallari, C., Quecini, V., Coutinho, L. L., Hoshino, A. A., et al. (2008). Peanut genes identified during initial phase of *Cercosporidium personatum* infection. *Plant Science*, 174, pp. 78-87.
- Nogueira Júnior, S. (1976). Evolução da produção e comercialização do amendoim no Brasil. In: *Anais do Simpósio Nacional y Latino Americano de Oleaginosas*, 8, Buenos Aires, Argentina.
- Novaes, E., Drost, D. R., Farmerie, W. G., Pappas Jr, G. J., Grattapaglia, D., Sederoff, R. R., et al. (2008). High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. *BMC Genomics*, 9 (312), p. 14.
- Novais, C. M., Pires-Alves, M., Silva, F. F. (2004). PCR em tempo real. *Bioteχνologia Ciência & Desenvolvimento*, VII (33), pp. 10-13.
- O'Brian, M. R., & Vance, C. P. (2007). Legume biology: sequence to seeds. *Plant Physiology*, 144, p. 537.
- Ogihara, Y., Mochida, K., Nemoto, Y., Murai, K., Yamazaki, Y., Shin-I, T., et al. (2003). Correlated clustering and virtual display of gene expression patterns in the wheat life cycle by large-scale statistical analyses of expressed sequence tags. *The Plant Journal*, 33, pp. 1001-1011.
- Oh, B., Ko, M. K., Kim, Y. S., Kim, K. S., Kostenyuk, I., Kee, H. K., et al. (1999). A cytochrome P450 gene is differentially expressed in compatible and incompatible interactions between pepper (*Capsicum annuum*) and the anthracnose fungus, *Colletotrichum gloeosporioides*. *Molecular Plant-Microbe Interactions*, 12 (12), pp. 1044-1052.
- Ohnishi, T., Yokota, T., Mizutani, M. (2009). Insights into the function and evolution of P450s in plant steroids metabolism. *Phytochemistry*, 70, pp. 1918-1929.
- Osborne, M. A., & Silver, P. A. (1993). Nucleocytoplasmic transport in the yeast *Saccharomyces cerevisiae*. *Annual Review of Biochemistry*, 62, pp. 219-254.
- Pacheco, G., Gagliardi, R. F., Carneiro, L. A., Valls, J. F., Mansur, E. (2008). Plant regeneration in *Arachis stenosperma* Krapov. and W. C. Gregory from roots and calluses derived from leaflets of in vitro plants. *In Vitro Celular & Development Biology - Plant*, 44, pp. 14-17.
- Pande, S., & Rao, J. N. (2001). Resistance of Wild *Arachis* Species to Late Leaf Spot and Rust in Greenhouse Trials. *Plant Disease*, 85, pp. 851-855.
- Parchman, T. L., Geist, K. S., Grahnen, J. A., Benkman, C. W., Buerkle, A. (2010). Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. *BMC Genomics*, 11 (180), p. 16.

- Park, J., Lee, H., Cheon, C., Kim, S., Hur, Y., Auh, C., et al. (2011). The *Arabidopsis thaliana* homeobox gene ATHB12 is involved in symptom development caused by geminivirus infection. *PLoS ONE*, 6 (5), p. 7.
- Pascholati, S. F., & Leite, B. (1995). Hospedeiro: mecanismos de resistência. In: A. Bergamin Filho, H. Kimati, & L. Amorin (Eds.), *Manual de Fitopatologia - Princípios e conceitos* (pp. 417-452). São Paulo: Agronômica Ceres.
- Passos, G. A., Nguyen, C., Jordan, B. (2000). Projeto Transcriptoma - Análise da expressão gênica em larga escala usando DNA - Arrays. *Biotecnologia*, 12 (2), pp. 34-37.
- Pattee, H. E. (2005). Peanut Oil. In: F. Shahidi, *Barley's Industrial Oil and Fat Product* (6th ed., Vol. VI, pp. 431-463). Raleigh, Carolina do Norte: Wiley-Interscience.
- Payton, P., Kottapalli, K. R., Rowland, D., Faircloth, W., Guo, B., Burrow, M., et al. (2009). Gene expression profiling in peanut using high density oligonucleotide microarrays. *BMC Genomics*, 10 (265), p. 11.
- Peirson, S. N., & Butler, J. N. (2007). Quantitative Polymerase Chain Reaction. *Springer Protocols*, 362, pp. 349-362.
- Pensuk, V., Patanothai, A., Jogloy, S., Wongkaew, S., Akkasaeng, C., Vorasoot, N., et al. (2003). Reaction of peanut cultivars to late leafspot and rust. *Songklanakarin Journal of Science and Technology*, 25 (3), pp. 289-295.
- Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, 29 (9), pp. 2002-2007.
- Pfaffl, M. W., Horgan, G. W., Dempfle, L. (2002). Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Research*, 30 (9), p. 10.
- PGP. (2011). *The Peanut Genome Project*. Acesso em Agosto de 2011, disponível em http://www.peanutbioscience.com/images/PGP_StratPlan_Goal_2_May11.pdf
- Piao, H., Xuan, Y., Park, S. H., Je, B. I., Park, S. J., Park, S. H., et al. (2010). OsCIPK31, a CBL-interacting protein kinase is involved in germination and seedling growth under abiotic stress conditions in rice plants. *Molecular and Cells*, 30, pp. 19-27.
- Pinot, F., & Beisson, F. (2011). Cytochrome P450 metabolizing fatty acids in plants: characterization and physiological roles. *The FEBS Journal*, 278 (2), pp. 195-205.
- Ponchel, F., Toomes, C., Bransfield, K., Leong, F. T., Douglas, S. H., Field, S. L., et al. (2003). Real-time PCR based SYBR-Green I fluorescence: An alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. *BMC Biotechnology*, 3 (18), p. 13.
- Ponciano, G., Yoshikawa, M., Lee, J. L., Ronald, P. C., Whalen, M. C. (2006). Pathogenesis-related gene expression in rice is correlated with developmentally controlled Xa21-mediated resistance against *Xanthomonas oryzae* pv. *oryzae*. *Physiological and Molecular Plant Pathology*, 69, pp. 131-139.

- Ponting, C. P., & Aravind, L. (1999). START: a lipid-binding domain in StAR, HD-Zip and signalling proteins. *Trends in Biochemical Sciences*, 24 (4), pp. 130-132.
- Pretorius, A. E. (2006). Evaluation of groundnut (*Arachis hypogaea* L.) germoplasm for resistance to leaf diseases and related cytoplasmic factors, testa colour and cup leaf. *Dissertação apresentada ao Departamento de Ciências Vegetais (Melhoramento Vegetal), na Faculdade de Ciências Naturais e Agronômicas da Universidade de Free State, Bloemfontein, África do Sul*, p. 216.
- Proite, K., Leal-Bertioli, S. C., Bertioli, D. J., Moretzsohn, M. C., Silva, F. R., Martins, N. F., et al. (2007). ESTs from wild *Arachis* species for gene discovery and marker development. *BMC Plant Biology*, 7 (7), p. 10.
- Rajeevan, M. S., Vernon, S. D., Taysavang, N., Unger, E. R. (2001). Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR. *Journal of Molecular Diagnostics*, 3 (1), pp. 26-31.
- Rajkarnikar, A., Kwon, H., Ryu, Y., Suh, J. (2006). Catalytic domain of AfsKav modulates both secondary metabolism and morphologic differentiation in *Streptomyces avermitilis* ATCC 31272. *Current Microbiology*, 53, pp. 204-208.
- Ramos, V. R. (2007). Caracterização da resistência às cercosporioses, lagarta do cartucho e lagarta da soja em espécies silvestres do gênero *Arachis*, para uso no melhoramento genético do amendoim. *Tese apresentada ao Instituto de Biociências da Universidade Estadual Paulista "Júlio de Mesquita Filho", Botucatu, SP, Brasil*, p. 109.
- Rao, N. K., Reddy, L. J., Bramel, P. J. (2003). Potential of wild species for genetic enhancement of some semi-arid food crops. *Genetic Resources and Crop Evolution*, 50, pp. 707-721.
- Roche Applied Science, D. (2004). Assay formats for use in Real-Time PCR. *Technical Note - Nº LC 18/2004*, p. 14.
- Roche. (s/d). *GS FLX System*. Acesso em Setembro de 2011, disponível em <http://www.my454.com/products/gs-flx-system/index.asp>
- Roche. (2009). *GS FLX Titanium - General library preparation method manual*. Acesso em Novembro de 2011, disponível em <http://dna.uga.edu/docs/GS-FLX-Titanium-General-Library-Preparation-Method-Manual%20%28Roche%29.pdf>
- Roche. (2011). *Now delivering sequencing reads up to 1,000 bp in length!* Acesso em Novembro de 2011, disponível em <http://my454.com/products/gs-flx-system/index.asp>
- Roche, A. S. (2006). *Genome Sequencer FLX*. Acesso em Setembro de 2011, disponível em http://www.roche-applied-science.com/publications/multimedia/genome_sequencer/flx_multimedia/wbt.htm
- Rothberg, J. M., & Leamon, J. H. (2008). The development and impact of 454 sequencing. *Nature Biotechnology*, 26 (10), 1117-1124.

- Rozhon, W., Mayerhofer, J., Petutschnig, E., Fujioka, S., Jonak, C. (2010). ASK0, a group-III *Arabidopsis* GSK3, functions in the brassinosteroid signalling pathway. *The Plant Journal*, 62, pp. 215-223.
- Rusnak, F., & Mertz, P. (2000). Calcineurin: form and function. *Physiological Reviews*, 80 (4), pp. 1483-1521.
- Russel, J. R., Bayer, M., Booth, C., Cardle, L., Hackett, P. E., Hedley, P. E., et al. (2011). Identification, utilisation and mapping of novel transcriptome-based markers from blackcurrant (*Ribes nigrum*). *BMC Plant Biology*, 11 (147), p. 36.
- Sadler, I., Kurihara, T., Rothblatt, J., Way, J., Silver, P. (1989). A yeast gene important for protein assembly into the endoplasmic reticulum and the nucleus has homology to DnaJ, an *Escherichia coli* heat shock protein. *The Journal of Cell Biology*, 109 (6), pp. 2665-2675.
- Sangha, J. S., Gu, K., Kaur, J., Yin, Z. (2010). An improved method for RNA isolation and cDNA library construction from immature seeds of *Jatropha curcas* L. *BMC Research Notes*, 3 (126), p. 6.
- Sanine, P. R. (2006). *Caracteres estruturais e dimensão fractal da folha de espécies de Arachis e sua relação com a cercosporiose*. Botucatu - SP: Universidade Estadual Paulista - UNESP.
- Santos, R. C. (1999). *Utilização de recursos genéticos e melhoramento de Arachis hypogaea L. no Nordeste Brasileiro*. Acesso em Agosto de 2011, disponível em Embrapa Semi Árido - CPATSA: <http://www.cpatosa.embrapa.br:8080/catalogo/livro/>
- Schmidt, W. M., & Mueller, M. W. (1999). CapSelect: A highly sensitive method for 5' CAP-dependent enrichment of full-length cDNA in PCR-mediated analysis of mRNAs. *Nucleic Acids Research*, 27 (21), p. 4.
- Schuler, M. A., & Werck-Reichhart, D. (2003). Functional genomics of P450s. *Annual Reviews Plant Biology*, 54, pp. 629-667.
- Seijo, G. J., Lavia, G. I., Fernandez, A., Krapovickas, A., Ducasse, E., Moscone, D. E., et al. (2004). Physical mapping of the 5S and 18S-25S rRNA genes by FISH as evidence that *Arachis duranensis* and *A. ipaensis* are the wild diploid progenitors of *A. hypogaea* (Leguminosae). *American Journal of Botany*, 91, pp. 1294-1303.
- Sharoni, A. M., Nuruzzaman, M., Satoh, K., Shimizu, T., Kondoh, H., Sasaya, T., et al. (2010). Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant and Cell Physiology*, 52 (2), pp. 344-360.
- Shi, C., Chaudhary, S., Yu, K., Park, S. J., Navabi, A., McClean, P. E., et al. (2011). Identification of candidate genes associated with CBB resistance in common bean HR45 (*Phaseolus vulgaris* L.) using cDNA-AFLP. *Molecular Biology Reports*, 38 (1), pp. 75-81.
- Shi, J., Kim, K., Ritz, O., Albrecht, V., Gupta, R., Harter, K., et al. (1999). Novel protein kinases associated with calcineurin B-like calcium sensors in *Arabidopsis*. *The Plant Cell*, 11, pp. 2393-2405.

- Shokes, F. M., & Culbreath, A. K. (1997). Early and late leaf spots. In: N. Kokalis-Burele, D. M. Porter, R. Rodriguez-Kábana, D. H. Smith, & P. Subrahmanyam (Eds.), *Compendium of Peanut Diseases* (2nd ed., pp. 17-20). St. Paul: APS Press.
- Shunwu, Y., Zhang, L., Zuo, K., Li, Z., Tang, K. (2004). Isolation and characterization of a BURP domain-containing gene BnBDC1 from *Brassica napus* involved in abiotic and biotic stress. *Physiologia Plantarum*, 122, pp. 210-218.
- Silva, G. S. (2008). *Contribuição à taxonomia do gênero Arachis - secção Arachis à luz do estudo de espécies e híbridos interespecíficos*. Dissertação. Universidade de Brasília, Instituto de Ciências Biológicas, Departamento de Botânica, p. 114.
- Silver, P. A. (1991). How proteins enter the nucleus. *Cell*, 64 (3), pp. 489-487.
- Simpson, C. E. (1991). Pathways for introgression of pest resistance into *Arachis hypogaea* L. *Peanut Science*, 18 (1), pp. 22-26.
- Simpson, C. E., Krapovickas, A., Valls, J. F. (2001). History of *Arachis* including evidence of *A. hypogaea* L. progenitors. *Plant Science*, 28, pp. 78-80.
- Singh, A. K., & Moss, J. P. (1982). Utilization of wild relatives in genetics improvement of *Arachis hypogaea* L. Part. 2. Chromosome complements of species in section *Arachis*. *Theoretical and Applied Genetics*, 61 (4), pp. 305-314.
- Singh, B., & Singh, U. (1991). Peanut as a source of protein for human foods. *Plant foods for human nutrition*, 41, pp. 165-177.
- Smartt, J. (1964). Cross-compatibility relationships between the cultivated peanut *Arachis hypogaea* L. and other species of the genus *Arachis*. *Tese de Ph. D - Universidade do Estado da Carolina do Norte*, Nº. 65, 8968.
- Smartt, J., Gregory, W. C., Gregory, M. P. (1978). The genome of *Arachis hypogaea* L. Cytogenetic studies of putative genome donors. *Euphytica*, 27, pp. 665-675.
- Sommartya, T., & Beute, M. K. (1986). Temperature effects on germination and comparative morphology of conidia for Thai and USA isolates of *Cercosporidium personatum*. *Peanut Science*, 13, pp. 67-70.
- Somssich, I. E., & Hahlbrock, K. (1998). Pathogen defence in plants - a paradigm of biological complexity. *Trends in Plant Science*, 3 (3), pp. 86-90.
- Son, L. V., Tiedemann, J., Rutten, T., Hillmer, S., Hinz, G., Zank, T., et al. (2009). The BURP domain protein AtUSPL1 of *Arabidopsis thaliana* is destined to the protein storage vacuoles and overexpression of the cognate gene distorts seed development. *Plant Molecular Biology*, 71, pp. 319-329.
- Song, G. Q., Li, M. J., Xiao, H., Wang, X. J., Tang, R. H., Xia, H., et al. (2010). EST sequencing and SSR marker development from cultivated peanut (*Arachis hypogaea* L.). *Electronic Journal of Biotechnology*, 13 (3), p. 9.

- Song, W., Wang, G., Chen, L., Kim, H., Pi, L., Holsten, T., et al. (1995). A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. *Science*, 270 (5243), pp. 1804-1806.
- Stalker, H. T. (1991). A new species in section *Arachis* of peanuts with a D genome. *American Journal of Botany*, 78 (5), pp. 630-637.
- Stalker, H. T. (1992). Utilizing *Arachis* germplasm resources. In: S. N. Nigam, *Groundnut - A Global Perspective* (pp. 281-295). Patancheru: ICRISAT.
- Stalker, H. T., & Moss, J. P. (1987). Speciation, cytogenetics, and utilization of *Arachis* species. *Advances in Agronomy*, 41, p. 40.
- Stalker, H. T., Weissinger, A. K., Mill-Lewis, S., Holbrook, C. C. (2009). Genomics: An evolving science in peanut. *Peanut Science*, 36 (1), pp. 2-10.
- Staskawicz, B. J. (2001). Genetics of plant-pathogen interactions specifying plant disease resistance. *Plant Physiology*, 125, pp. 73-76.
- Stewart, E. L., Liu, Z., Crous, P. W., Szabo, L. J. (1999). Phylogenetic relationships among some cercosporoid anamorphs of *Mycosphaerella* based on rDNA sequence analysis. *Mycological Research*, 103 (11), pp. 1491-1499.
- Subrahmanyam, P., Williams, J. H., McDonald, D., Gibbons, R. W. (1984). The influence of foliar diseases and their control by selective fungicides on a range of groundnut (*Arachis hypogaea* L.) genotypes. *Annals of Applied Biology*, 104 (3), pp. 467-476.
- Suzuki, K. (2002). MAP kinase cascades in elicitor signal transduction. *Journal of Plant Research*, 115 (3), pp. 237-244.
- Suzuki, Y., Mae, T., Makino, A. (2008). RNA extraction from various recalcitrant plant tissues with a Cethyltrimethylammonium Bromide-containing buffer followed by an acid guanidium thiocyanate-phenol-chloroform treatment. *Bioscience, Biotechnology, and Biochemistry*, 72 (7), pp. 1951-1953.
- Takemoto, D., Hayashi, M., Doke, N., Nishimura, M., Kawakita, K. (1999). Molecular cloning of a defense-response-related Cytochrome P450 gene from Tobacco. *Plant & Cell Physiology*, 40 (12), pp. 1232-1242.
- Tang, D., Christiansen, M., Innes, R. W. (2005). Regulation of plant disease resistance, stress responses, cell death, and ethylene signaling in *Arabidopsis* by the EDR1 protein kinase. *Plant Physiology*, 138, pp. 1018-1026.
- Tangphatsornruang, S., Sangsrakru, D., Chanprasert, J., Uthaipaisanwong, P., Yoocha, T., Jomchai, N., et al. (2010). The chloroplast genome sequence of mungbean (*Vigna radiata*) determined by high-throughput pyrosequencing: Structural organization and phylogenetic relationships. *DNA Research*, 17, pp. 11-22.
- Tavares, R., Vidal, J., van Lammeren, A., Kreis, M. (2002). AtSKtheta, a plant homologue of SGG/GSK-3 marks developing tissues in *Arabidopsis thaliana*. *Plant Molecular Biology*, 50 (2), pp. 261-271.

Temsch, E. M., & Greilhuber, J. (2001). Genome size in *Arachis duranensis*: a critical study. *Genome*, 44 (5), pp. 828-830.

Terefe-Ayana, D., Yasmin, A., Le, T. L., Kaufmann, H., Biber, A., Kuhr, A., et al. (2011). Mining disease-resistance genes in roses: functional and molecular characterization of the Rdr1 locus. *Frontiers in Plant Science - Crop Science and Horticulture*, 2, pp. 1-12.

Torres, T. T., Metta, M., Ottenwalder, B., Schlotterer, C. (2008). Gene expression profiling by massively parallel sequencing. *Genome Research*, 18, pp. 172-177.

Tripathi, V., Parasuraman, B., Laxmi, A., Chattopadhyay, D. (2009). CIPK6, a CBL-interacting protein kinase is required for development and salt tolerance in plants. *The Plant Journal*, 58, pp. 778-790.

Udvardi, M. K., Kakar, K., Wandrey, M., Montanari, O., Murray, J., Andriankaja, A., et al. (2007). Legume transcription factors - global regulators of plant development and response to the environment. *Plant Physiology*, 144, pp. 538-549.

Uehara, T., Sugiyama, S., Masuta, C. (2007). Comparative serial analysis of gene expression of transcript profiles of tomato roots infected with cyst nematode. *Plant Molecular Biology*, 63 (2), pp. 185-194.

USDA-FAS. (2011a). *Table 01: Major Oilseeds: World Supply and Distribution (Commodity View)*. Acesso em Agosto de 2011, disponvel em USDA - United States Department of Agriculture/FAS - Foreign Agricultural Service: <http://www.fas.usda.gov/psdonline/psdreport.aspx?hidReportRetrievalName=BVS&hidReportRetrievalID=700&hidReportRetrievalTemplateID=5>

USDA-FAS. (2011b). *Table 13: Peanut Area, Yield, and Production*. Acesso em Agosto de 2011, disponvel em USDA - United States Department of Agriculture/FAS - Foreign Agricultural Service: <http://www.fas.usda.gov/psdonline/psdreport.aspx?hidReportRetrievalName=BVS&hidReportRetrievalID=918&hidReportRetrievalTemplateID=1>

Uzarowska, A., Dionisio, G., Sarholz, B., Piepho, H., Xu, M., Ingvaridsen, C. R., et al. (2009). Validation of candidate genes putatively associated with resistance to SCMV and MDMV in maize (*Zea mays* L.) by expression profiling. *BMC Plant Biology*, 9 (15), p. 15.

Valentine, H. (2011). Feeding the hungry world with peanuts. In: *5th International Conference of the Peanut Research Community on Advances in Arachis through genomics and biotechnology* (p. 9). Braslia: Embrapa Recursos Genticos e Biotecnologia.

Valls, J. F., & Simpson, C. E. (2005). New species of *Arachis* from Brazil, Paraguay, and Bolivia. *Bonplandia*, 14, pp. 35-64.

van Vliet, A. H. (2009). Next generation sequencing of microbial transcriptomes: challenges and opportunities. *FEMS Microbiology Letters*, 302, p. 1-7.

Varshney, R. K., Close, T. J., Singh, N. K., Hoisington, D. A., Cook, D. R. (2009). Orphan legume crops enter the genomics era! *Current Opinion in Plant Biology*, 12 (2), pp. 202-210.

- Veiga, R. F., Queiroz-Voltan, R. B., Valls, J. F., Fávero, A. P., Barbosa, W. (2001). Caracterização morfológica de acessos de germoplasma de quatro espécies brasileiras de amendoim-silvestre. *Bragantia*, 60 (3), pp. 167-176.
- Vleeshouwers, V. G., Martens, A., van Dooijeweert, W., Colon, L. T., Govers, F., Kamoun, S., et al. (2001). Ancient diversification of the Pto kinase family preceded speciation in *Solanum*. *Molecular Plant-Microbe interactions*, 14 (8), pp. 996-1005.
- Vrebalov, J., Pan, I. L., Arroyo, A. J., McQuinn, R., Chung, M., Poole, M., et al. (2009). Fleshy fruit expansion and ripening are regulated by the tomato SHATTERPROOF gene TAGL1. *The Plant Cell*, 21, pp. 3041-3062.
- Wagner, C. (2008). Quantitative resistance of barley against scald - A candidate gene approach. In: J. L. Molina Cano, P. Christou, A. Graner, K. Hammer, N. Jouve, B. Keller, et al. (Eds.), *Cereal science and technology for feeding ten billion people: genomics era and beyond* (pp. 175-177). Zaragoza: CIHEAM-IAMZ/IRTA.
- Wang, W., Wang, Y., Zhang, Q., Qi, Y., Guo, D. (2009). Global characterization of *Artemisia annua* glandular trichome transcriptome using 454 pyrosequencing. *BMC Genomics*, 10 (465), p. 10.
- Wang, X., Wu, W., Jin, G., Zhu, J. (2005). Genome-wide identification of R genes and exploitation of candidate RGA markers in rice. *Chinese Science Bulletin*, 50 (11), pp. 1120-1125.
- Wang, Y., Qun, L., Zuhua, H. E. (2004). Blast fungus-induction and development and tissue-specific expression of a rice P450 CYP72A gene cluster. *Chinese Science Bulletin*, 49 (2), pp. 131-135.
- Wang, Z., Gerstein, M., Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10, pp. 57-63.
- Wang, Z., Libault, M., Joshi, T., Valliyodan, B., Nguyen, H. T., Xu, D., et al. (2010). SoyDB: a knowledge database of soybean transcription factors. *BMC Plant Biology*, 10 (14), p. 12.
- Wawrzynska, A., Christiansen, K. M., Lan, Y., Rodibaugh, N. L., Innes, R. W. (2008). Powdery mildew resistance conferred by loss of the enhanced disease resistance 1 protein kinase is suppressed by a missense mutation in keep on going, a regulator of abscisic acid signaling. *Plant Physiology*, 148, pp. 1510-1522.
- Weber, A. P., Weber, K. L., Carr, K., Wilkerson, C., Ohlrogge, J. B. (2007). Sampling the *Arabidopsis* transcriptome with massively parallel pyrosequencing. *Plant Physiology*, 144, pp. 32-42.
- Werck-Reichhart, D., Hehn, A., Didierjean, L. (2000). Cytochrome P450 for engineering herbicide tolerance. *Trends in plant science*, 5 (3), pp. 116-123.
- Wicker, T., Schlagenhauf, E., Graner, A., Close, T. J., Keller, B., Stein, N., et al. (2006). 454 sequencing put to the test using the complex genome of barley. *BMC Genomics*, 7 (275), pp. 1-11.

- Wrzaczek, M., Brosché, M., Salojärvi, J., Kangasjärvi, S., Idänheimo, N., Mersmann, S., et al. (2010). Transcription regulation of the CRK/DUF26 group of receptor-like protein kinases by ozone and plant hormones in *Arabidopsis*. *BMC Plant Biology*, 10 (95), p. 19.
- Wrzaczek, M., Rozhon, W., Jonak, C. (2007). A proteasome-regulated glycogen synthase kinase-3 modulates disease response in plants. *The Journal of Biological Chemistry*, 282 (8), pp. 5249-5255.
- Wynne, J. C., Beute, M. K., Nigam, S. N. (1991). Breeding for disease resistance in peanut (*Arachis hypogaea* L.). *Annual Review of Phytopathology*, 29, pp. 279-303.
- Xian-Guang, Y., Jin-Yu, F., Chuan-Liang, D. (2011). Molecular cloning of a novel GSK3/shaggy-like gene from *Triticum monococcum* L. and its expression in response to salt, drought and other abiotic stresses. *African Journal of Biotechnology*, 10 (20), pp. 4065-4071.
- Xiao, S., Wang, W., Yang, X. (2008). Evolution of resistance genes in plants. In: H. Heine (Ed.), *Innate immunity of plants, animals and humans* (Vols. 21 - Nucleic Acids and Molecular Biology, p. 20). Springer.
- Xu, H., Li, Y., Yan, Y., Wang, K., Gao, Y., Hu, Y., et al. (2010). Genome-scale identification of Soybean BURP domain-containing genes and their expression under stress treatments. *BMC Plant Biology*, 10 (197), p. 16.
- Xu, Q., & Deng, X. (2010). Cloning and phylogenetic analyses of serine/threonine kinase class defense-related genes in a wild fruit crop 'chestnut rose'. *BMC Research Notes*, 3 (202), p. 8.
- Xu, Q., Wen, X., Deng, X. (2008). Genomic organization, rapid evolution and meiotic instability of NBS-encoding genes in a new fruit crop "chestnut rose". *Genetics*, 178, pp. 2081-2091.
- Yoo, M., Albert, V. A., Soltis, P. S., Soltis, D. E. (2006). Phylogenetic diversification of glycogen synthase kinase 3/SHAGGY-like kinase genes in plants. *BMC Plant Biology*, 6 (3), p. 14.
- Young, N. D., Weeden, N. F., Kochert, G. (1996). Genome mapping in legumes (Family Fabaceae). In: A. H. Paterson (Ed.), *Genome mapping in plants* (pp. 212-227). RG Landes: Austins, Texas.
- You-Xiong, Q., Jian-Wei, L., Xian-Xian, S., Li-Ping, X., Ru-Kai, C. (2011). Differential gene expression in sugarcane in response to challenge by fungal pathogen *Ustilago scitaminea* revealed by cDNA-AFLP. *Journal of Biomedicine and Biotechnology*, 2011, p. 10.
- Zeng, Y., & Yang, T. (2002). RNA isolation from highly viscous samples rich in polyphenols and polysaccharides. *Plant Molecular Biology Reporter*, 20 (4), pp. 417a-417e.
- Zhang, L., Tian, L., Zhao, J., Song, Y., Zhang, C., Guo, Y., et al. (2009). Identification of an apoplastic protein involved in the initial phase of salt stress response in rice root by two-dimensional electrophoresis. *Plant Physiology*, 149, pp. 916-928.

Zhang, Y., Zhang, G., Xia, N., Wang, X., Huang, L., Kang, Z., et al. (2008). Cloning and characterization of a bZIP transcription factor gene in wheat and its expression in response to strip rust pathogen infection and abiotic stress. *Physiological and Molecular Plant Pathology*, 73 (4-5), pp. 88-94.

Zhao, Q., & Guo, H. (2011). Paradigms and paradox in the ethylene signaling pathway and interaction network. *Molecular Plant*, First published online June 20, p. 9.

Zhu, Y. Y., Machleder, E. M., Chenchik, A., Li, R., Siebert, P. D. (2001). Reverse transcriptase template switching: A SMART™ approach for full-length cDNA library construction. *BioTechniques*, 30, pp. 892-897.

Zorina, A., Stepanchenko, N., Novikova, G. V., Sinetova, M., Panichkin, V. B., Moshkov, I. E., et al. (2011). Eukaryotic-like Ser/Thr protein kinases SpkC/F/K are involved in phosphorylation of GroES in the cyanobacterium *Synechocystis*. *DNA Research*, First published online, p. 15.

9. APÊNDICES

Tabela 4 – Cem primeiros unigenes (formados por dois ou mais *contigs*), diferencialmente expressos indicados pela subtração *in silico*, entre as bibliotecas de folhas de *A. stenosperma* controle (SFC) e inoculadas com *P. personata* (SFI). ID = identidade do *contig*; SFI/SFC = razão entre amostras inoculadas e controle, em vezes. Regulados negativamente (em azul); Regulados positivamente (em vermelho); Exclusivos nas amostras controle (SFC – em verde) e exclusivo nas amostras inoculadas (SFI – em roxo).

ID	Tamanho (pb)	SFI/SFC	Hit	Melhor pontuação	E-value	Identidade (%)
Astenosperma_large_rep_c136	425	377,00	<i>Vitis vinifera</i>	<i>Hypothetical protein isoform 2</i>	4,76e ⁻²⁵	85,71
Astenosperma_large_rep_c77	1450	287,50	<i>Glycine max</i>	<i>Unknown</i>	2,76e ⁻⁹¹	82,65
Astenosperma_small_rep_c128	280	272,00	<i>No hit</i>	-	-	-
Astenosperma_large_rep_c83	672	217,00	<i>Glycine max</i>	<i>Unknown</i>	1,48e ⁻⁴⁷	54,19
Astenosperma_large_rep_c47	1230	210,50	<i>Vitis vinifera</i>	<i>Unnamed protein product</i>	3,69e ⁻¹⁰²	75,93
Astenosperma_large_rep_c78	538	160,00	<i>Arabidopsis lyrata</i>	<i>Predicted protein</i>	2,63e ⁻³⁷	60,17
Astenosperma_large_rep_c207	703	131,00	<i>Glycine max</i>	<i>Unknown</i>	2,96e ⁻⁶⁸	88,73
Astenosperma_large_rep_c66	441	122,00	<i>No hit</i>	-	-	-
Astenosperma_small_rep_c168	267	117,00	<i>No hit</i>	-	-	-
Astenosperma_large_rep_c144	477	116,50	<i>Glycine max</i>	<i>Unknown</i>	4,09e ⁻²²	72,15
Astenosperma_large_rep_c91	740	105,50	<i>Populus trichocarpa</i>	<i>Unknown</i>	1,13e ⁻⁴⁴	88,78
Astenosperma_small_rep_c70	1795	103,00	<i>Ricinus communis</i>	<i>Tubulin alpha chain, putative</i>	0,00e ⁺⁰⁰	94,44
Astenosperma_large_rep_c201	588	101,00	<i>Glycine max</i>	<i>Homeodomain-leucine zipper protein 56</i>	7,96e ⁻²²	72,73
Astenosperma_small_rep_c26	427	100,00	<i>No hit</i>	-	-	-
Astenosperma_large_rep_c227	219	86,50	<i>No hit</i>	-	-	-
Astenosperma_large_rep_c265	371	78,00	<i>No hit</i>	-	-	-
Astenosperma_large_rep_c349	387	75,00	<i>No hit</i>	-	-	-
Astenosperma_small_rep_c122	1751	68,00	<i>Glycine max</i>	<i>Glycogen synthase kinase</i>	0,00e ⁺⁰⁰	97,80
Astenosperma_large_rep_c11	423	61,45	<i>Populus trichocarpa</i>	<i>Predicted protein</i>	3,66e ⁻⁰⁹	76,92
Astenosperma_small_rep_c114	1146	61,00	<i>Nicotiana tabacum</i>	<i>Putative beta6 proteasome subunit</i>	2,10e ⁻⁸⁹	92,25

ID	Tamanho (pb)	SFI/SFC	Hit	Melhor pontuação	E-value	Identidade (%)
Astenosperma_small_rep_c118	319	56,67	<i>Glycine max</i>	<i>Unknown</i>	6,93e ⁻¹⁶	72,73
Astenosperma_large_rep_c211	367	53,50	<i>No hit</i>	-	-	-
Astenosperma_large_rep_c187	469	53,00	<i>Sorghum bicolor</i>	<i>Hypothetical protein SORBIDRAFT_01g019370</i>	3,21e ⁻¹³	89,47
Astenosperma_small_rep_c100	1148	52,00	<i>Glycine max</i>	<i>Unknown</i>	5,99e ⁻¹⁰⁴	93,85
Astenosperma_large_rep_c270	911	52,00	<i>Glycyrrhiza uralensis</i>	<i>Actin</i>	1,17e ⁻⁷⁰	100,00
Astenosperma_large_rep_c131	902	50,50	<i>Vitis vinifera</i>	<i>Hypothetical protein</i>	3,36e ⁻⁶⁶	74,42
Astenosperma_large_rep_c85	1042	49,20	<i>Glycine max</i>	<i>Unknown</i>	4,86e ⁻⁷⁸	71,98
Astenosperma_small_rep_c390	2405	47,00	<i>Populus trichocarpa</i>	<i>Predicted protein</i>	0,00e ⁺⁰⁰	61,67
Astenosperma_large_rep_c212	503	47,00	<i>No hit</i>	-	-	-
Astenosperma_small_rep_c62	326	46,00	<i>No hit</i>	-	-	-
Astenosperma_large_rep_c165	671	45,31	<i>Bruguiera gymnorhiza</i>	<i>BURP domain-containing protein</i>	4,74e ⁻⁰⁹	57,45
Astenosperma_small_rep_c13	654	44,00	<i>Populus trichocarpa</i>	<i>Predicted protein</i>	3,33e ⁻²⁹	59,48
Astenosperma_small_rep_c503	269	44,00	<i>No hit</i>	-	-	-
Astenosperma_large_rep_c206	1891	44,00	<i>Phaseolus acutifolius</i>	<i>Heat shock transcription factor</i>	5,53e ⁻¹⁵⁴	67,49
Astenosperma_small_rep_c310	354	42,00	<i>No hit</i>	-	-	-
Astenosperma_large_c335	876	41,00	<i>Medicago truncatula</i>	<i>Unknown</i>	1,70e ⁻⁷⁰	89,04
Astenosperma_small_rep_c17	368	39,67	<i>No hit</i>	-	-	-
Astenosperma_large_rep_c145	1116	39,33	<i>Glycine max</i>	<i>Unknown</i>	2,05e ⁻⁷⁷	66,29
Astenosperma_large_c395	1587	39,00	<i>Vitis vinifera</i>	<i>Hypothetical protein</i>	0,00e ⁺⁰⁰	88,10
Astenosperma_large_rep_c364	301	38,00	<i>No hit</i>	-	-	-
Astenosperma_large_rep_c93	543	37,60	<i>No hit</i>	-	-	-
Astenosperma_small_rep_c179	1035	37,00	<i>Glycine max</i>	<i>Epsilon1-COP</i>	8,10e ⁻¹²⁶	90,36
Astenosperma_large_rep_c437	393	37,00	<i>No hit</i>	-	-	-
Astenosperma_large_rep_c87	406	36,83	<i>No hit</i>	-	-	-

ID	Tamanho (pb)	SFI/SFC	Hit	Melhor pontuação	E-value	Identidade (%)
Astenosperma_large_rep_c92	932	35,55	No hit	-	-	-
Astenosperma_large_rep_c67	869	35,40	<i>Glycine max</i>	Unknown	1,33e ⁻⁷²	98,56
Astenosperma_large_rep_c456	194	35,00	No hit	-	-	-
Astenosperma_large_rep_c160	791	35,00	<i>Arachis hypogaea</i>	Putative L24 ribosomal protein	2,05e ⁻⁵⁰	73,97
Astenosperma_small_rep_c142	771	34,00	<i>Arabidopsis thaliana</i>	Histone H3.2	3,27e ⁻⁶⁴	100,00
Astenosperma_large_rep_c513	298	33,00	<i>Glycine max</i>	Unknown	4,39e ⁻¹⁵	81,25
Astenosperma_small_rep_c271	1594	33,00	<i>Vitis vinifera</i>	Hypothetical protein	1,13e ⁻⁴⁸	33,79
Astenosperma_large_rep_c445	207	31,00	No hit	-	-	-
Astenosperma_large_rep_c497	187	31,00	<i>Vitis vinifera</i>	Hypothetical protein	6,97e ⁻¹⁶	95,12
Astenosperma_small_rep_c380	2263	30,00	<i>Vitis vinifera</i>	Hypothetical protein	3,67e ⁻¹⁷⁸	65,40
Astenosperma_large_rep_c53	709	29,94	No hit	-	-	-
Astenosperma_large_rep_c150	854	29,67	<i>Arachis hypogaea</i>	Galactose-binding lectin precursor	7,32e ⁻⁹⁶	89,16
Astenosperma_large_c236	1367	29,67	<i>Glycine max</i>	Unknown	0,00e ⁺⁰⁰	87,76
Astenosperma_large_rep_c118	412	29,00	No hit	-	-	-
Astenosperma_small_rep_c373	833	29,00	<i>Glycine max</i>	Unknown	1,76e ⁻⁸⁷	71,97
Astenosperma_small_rep_c180	678	29,00	<i>Glycine max</i>	Ferredoxin thioredoxin reductase precursor	2,22e ⁻⁵⁴	91,67
Astenosperma_large_c281	1520	29,00	<i>Vitis vinifera</i>	Hypothetical protein	6,04e ⁻¹⁵³	68,14
Astenosperma_large_c384	1017	29,00	<i>Ricinus communis</i>	Conserved hypothetical protein	3,94e ⁻⁹³	70,00
Astenosperma_large_rep_c104	567	28,90	No hit	-	-	-
Astenosperma_large_rep_c342	312	28,50	<i>Glycine max</i>	Unknown	1,46e ⁻¹⁰	70,45
Astenosperma_small_rep_c10	444	28,14	No hit	-	-	-
Astenosperma_large_rep_c129	1648	27,83	<i>Glycine max</i>	Calreticulin-1	0,00e ⁺⁰⁰	84,51
Astenosperma_large_rep_c323	364	27,50	No hit	-	-	-
Astenosperma_small_rep_c368	324	26,00	No hit	-	-	-
Astenosperma_large_rep_c50	1670	25,57	<i>Medicago sativa</i>	Fructose bisphosphate aldolase	0,00e ⁺⁰⁰	91,58
Astenosperma_small_rep_c347	1486	25,50	<i>Glycine max</i>	Unknown	1,54e ⁻¹⁴⁵	76,66

ID	Tamanho (pb)	SFI/SFC	Hit	Melhor pontuação	E-value	Identidade (%)
Astenosperma_large_rep_c74	507	25,21	<i>Glycine max</i>	<i>Unknown</i>	1,03e ⁻¹⁴	90,24
Astenosperma_large_c430	327	25,00	No hit	-	-	-
Astenosperma_large_rep_c116	430	24,50	<i>Oryza sativa Japonica Group</i>	<i>Os09g0361200</i>	5,15e ⁻³³	91,14
Astenosperma_small_rep_c278	1051	24,33	<i>Phaseolus vulgaris</i>	<i>Gag-pol polyprotein</i>	8,04e ⁻²⁰	30,00
Astenosperma_small_rep_c560	594	24,00	No hit	-	-	-
Astenosperma_small_rep_c1042	221	24,00	No hit	-	-	-
Astenosperma_large_rep_c322	994	24,00	<i>Glycine max</i>	<i>Unknown</i>	2,86e ⁻³²	81,18
Astenosperma_large_c346	527	24,00	<i>Vitis vinifera</i>	<i>Hypothetical protein</i>	2,20e ⁻²⁹	68,60
Astenosperma_large_rep_c111	828	23,40	<i>Vitis vinifera</i>	<i>Unnamed protein product</i>	6,06e ⁻⁵⁶	99,10
Astenosperma_small_rep_c547	1340	23,00	<i>Vitis vinifera</i>	<i>Hypothetical protein</i>	1,78e ⁻¹³¹	69,83
Astenosperma_large_c324	718	23,00	<i>Glycine max</i>	<i>Unknown</i>	2,14e ⁻³⁷	68,29
Astenosperma_large_rep_c178	1407	22,60	<i>Medicago truncatula</i>	<i>Unknown</i>	1,10e ⁻¹⁶²	87,90
Astenosperma_large_rep_c12	1479	22,32	<i>Vitis vinifera</i>	<i>Hypothetical protein isoform 2</i>	4,01e ⁻⁵³	52,36
Astenosperma_small_rep_c268	2022	22,00	<i>Vitis vinifera</i>	<i>Unnamed protein product</i>	0,00e ⁺⁰⁰	93,32
Astenosperma_large_rep_c61	446	21,86	<i>Glycine max</i>	<i>Unknown</i>	3,00e ⁻¹¹	80,00
Astenosperma_small_rep_c716	192	21,67	No hit	-	-	-
Astenosperma_small_rep_c836	333	21,50	No hit	-	-	-
Astenosperma_large_rep_c38	531	21,40	No hit	-	-	-
Astenosperma_large_rep_c101	958	21,13	<i>Physcomitrella patens</i>	<i>Predicted protein</i>	1,27e ⁻⁷⁰	70,23
Astenosperma_small_rep_c354	691	21,00	<i>Vitis vinifera</i>	<i>Unnamed protein product</i>	3,42e ⁻¹¹	73,91
Astenosperma_large_rep_c240	504	21,00	No hit	-	-	-
Astenosperma_large_rep_c199	1039	20,50	<i>Ipomoea batatas</i>	<i>Cysteine proteinase precursor</i>	5,20e ⁻⁷²	84,35
Astenosperma_large_rep_c357	885	20,50	No hit	-	-	-
Astenosperma_small_rep_c101	1613	20,50	<i>Glycine max</i>	<i>Unknown</i>	1,94e ⁻¹⁵⁵	67,81
Astenosperma_small_rep_c403	2363	20,50	<i>Vitis vinifera</i>	<i>Unnamed protein product</i>	0,00e ⁺⁰⁰	74,36

ID	Tamanho (pb)	SFI/SFC	Hit	Melhor pontuação	E-value	Identidade (%)
Astenosperma_small_rep_c323	604	20,00	<i>Glycine max</i>	<i>Unknown</i>	1,06e ⁻⁷²	77,78
Astenosperma_small_rep_c659	252	19,50	<i>No hit</i>	-	-	-
Astenosperma_small_rep_c269	783	19,00	<i>Glycine max</i>	<i>Magnesium chelatase subunit</i>	1,33e ⁻⁷⁸	96,67
Astenosperma_small_rep_c794	1022	19,00	<i>Glycine max</i>	<i>Unknown</i>	4,77e ⁻¹⁰⁴	73,42
Astenosperma_large_c375	1114	19,00	<i>Glycine max</i>	<i>Unknown</i>	1,01e ⁻⁹²	76,89