



UNIVERSIDADE  
ESTADUAL de LONDRINA

---

EUZIANE JOANA LINO

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE TRANSCRITOS  
RELACIONADOS A MECANISMOS DE RESISTÊNCIA À FERRUGEM  
ASIÁTICA DA SOJA**



Universidade Estadual de Londrina

Instituto Agronômico do Paraná

Empresa Brasileira de Pesquisa Agropecuária

---

**EUZIANE JOANA LINO**

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE TRANSCRITOS  
RELACIONADOS A MECANISMOS DE RESISTÊNCIA À FERRUGEM  
ASIÁTICA DA SOJA**

EUZIANE JOANA LINO

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE TRANSCRITOS  
RELACIONADOS A MECANISMOS DE RESISTÊNCIA À FERRUGEM  
ASIÁTICA DA SOJA**

Dissertação apresentada ao Programa de Pós-graduação em Genética e Biologia Molecular, da Universidade Estadual de Londrina, como requisito parcial para a obtenção do título de Mestre.

Orientador: *Dr. Ricardo Vilela Abdelnoor*  
Co-orientadora: *Dra. Francismar Correa  
Marcelino Guimarães*  
Co-orientadora: *Dra. Renata Stolf Moreira*

Londrina  
2011

**Catálogo elaborado pela Divisão de Processos Técnicos da Biblioteca Central da  
Universidade Estadual de Londrina.**

**Dados Internacionais de Catalogação-na-Publicação (CIP)**

L758i Lino, Euziane Joana.

Identificação e caracterização de transcritos relacionados a mecanismos de resistência à ferrugem asiática da soja / Euziane Joana Lino. – Londrina, 2011.  
93 f. : il.

Orientador: Ricardo Vilela Abdelnoor.

Co-orientador: Francismar Correa Marcelino Guimarães.

Co-orientador: Renata Stolf Moreira.

Dissertação (Mestrado em Genética e Biologia Molecular) – Universidade Estadual de Londrina, Centro de Ciências Biológicas, Programa de Pós-Graduação em Genética e Biologia Molecular, 2011.

Inclui bibliografia.

1. Soja – Melhoramento genético – Teses. 2. Soja – Resistência a doenças e pragas – Teses. 3. Soja – Ferrugem asiática – Teses. 4. Ferrugem asiática – Teses. I. Abdelnoor, Ricardo Vilela. II. Guimarães, Francismar Correa Marcelino. III. Moreira, Renata Stolf. IV. Universidade Estadual de Londrina. Centro de Ciências Biológicas. Programa de Pós-Graduação em Genética e Biologia Molecular. V. Instituto Agronômico do Paraná. VI. EMBRAPA. VII. Título.

CDU 631.52:633.34

EUZIANE JOANA LINO

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE TRANSCRITOS  
RELACIONADOS A MECANISMOS DE RESISTÊNCIA À FERRUGEM  
ASIÁTICA DA SOJA**

Dissertação apresentada ao Programa de Pós-graduação em Genética e Biologia Molecular, da Universidade Estadual de Londrina, como requisito parcial para a obtenção do título de Mestre.

**BANCA EXAMINADORA**

---

Prof. Dr. Ricardo Vilela Abdelnoor  
EMBRAPA/UEL – Londrina – PR

---

Prof. Dr. Laurival Antônio Vilas Bôas  
UEL – Londrina – PR

---

Prof. Dra. Danielle Cristina Gregorio da Silva  
UENP – Bandeirantes – PR

Londrina, 31 de março de 2011.

*Dedico aos meus pais Marlúcia e Euzebio, às minhas avós “Carme” e “Alda”, às minhas tias Nereide e Euza e à minha prima “Guga”, todos os meus exemplos de “batalha” e vida, que contribuíram para a construção dos meus conceitos de mundo e do meu caráter!*

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus pela minha existência, por trilhar o meu caminho e me guiar em todas as etapas de minha vida, sempre me dando forças para superar os obstáculos impostos.

Agradeço ao meu orientador Dr. Ricardo Vilela Abdelnoor, pela confiança ao me aceitar como sua orientanda, por todo auxílio e disponibilizado para a realização do trabalho e pela orientação.

Agradeço a minha co-orientadora Dra. Francismar Corrêa Marcelino Guimarães, por todos os conselhos e, que mesmo estando longe por um período, se fez presente em cada momento da realização do trabalho, sempre contribuindo com sugestões e orientações para que esse trabalho pudesse ser concretizado. Muito obrigada Fran por sempre ter me amparado!

Agradeço a minha co-orientadora Dra. Renata Stolf Moreira, por todos os ensinamentos compartilhados, por toda ajuda e apoio durante esses dois anos de mestrado, por todas as dicas, por toda a paciência e por sempre se preocupar comigo. Com certeza sem ela nem um terço desse trabalho teria sido concretizado. “Rê esse trabalho é tão meu quanto seu”!

Agradeço o Dr. Alexandre Lima Nepomuceno e ao Dr. Eliseu Binnek pelas correções e dicas sobre o meu trabalho que com certeza contribuíram para uma melhor estruturação deste.

Agradeço a Dra. Joice Felipes Barbosa que foi a primeira pessoa que acompanhei no laboratório e que me ensinou os primeiros protocolos. Agradeço também ao Dr. Rodrigo Matheus Pereira que participou imensamente deste trabalho também e que me auxiliou na parte de bioinformática.

Agradeço as pós-doc, Beatriz de Almeida Barros, Fabiana Rodrigues e Mayra Gallo de Carvalho, pelas conversas, pelos momentos de descontração, pelas dicas e discussões sobre o trabalho.

Agradeço aos meus queridos amigos André Passianoto e Paula Camargo, que com certeza tornaram esses anos de mestrado bem mais fáceis, me presenteando com suas convivências, escutando os meus desabafos e vivenciando momentos memoráveis de risadas!

Agradeço a minha querida Cynara Romero, amiga e companheira de mestrado, a quem sempre me ajudou quando eu precisei, seja nas disciplinas, seja na bancada, por todos os momentos peculiares vivenciados.

Agradeço aos doutorandos Salvador Lima Brito Júnior e a Adriana Polizel, por toda a ajuda que me deram durante os experimentos em casa de vegetação, por toda ajuda em bancada e paciência com que me ensinaram muitas das coisas que aprendi.

Agradeço ao Lucas Soares por estar sempre disposto a me ajudar com as análises de bioinformática, e por compartilhar comigo um pouco dos seus conhecimentos.

Agradeço aos técnicos e analistas do laboratório, Verinha, Silvana, Márcia, César e Vieira, que sempre estiveram por perto dispostos a ajudar e explicar o que fosse necessário.

Agradeço ao pessoal do laboratório Águida Morales, Alan Pereira, Amanda Paiva, Cibelle Engels, Edgar Mitani, Elton Gargioni, Gislaine Vasquez, Idenize Orsini, João Vitor Maldonado, Juliane Marinho, Juliana Leite, Juliana Marcolino, Kleber Miranda, Larissa Giroto, Lívia Nogueira, Lizandra Catelli, Maísa Pinheiro, Maria Cecília, Mayla Molinari, Michelle Rincão, Paulo Camargo, Paulo Silla, Renata Fuganti, Selma Pereira (*in memoriam*), Tânia Hitomi (*in memoriam*), Tatiana Fraga, Thiago Nakayama e Valéria Lopes que fizeram parte do meu cotidiano durante esses anos de mestrado e pela boa convivência.

Agradeço aos meus pais Marlúcia e Euzebio, por todo o apoio e esforços que fizeram durante os meus estudos.

Agradeço à Embrapa Soja e ao CNPq pela concessão de bolsas durante o mestrado, a Embrapa Soja mais uma vez pela disponibilização do laboratório e equipamentos para o desenvolvimento do meu trabalho e a Universidade Estadual de Londrina por me receber como aluna e a todos os professores que contribuíram para o meu crescimento cognitivo. Agradeço também a secretaria e coordenação da pós-graduação em genética em biologia molecular pela dedicação e carinho dispendido aos alunos.

MUITO OBRIGADA!

"Aprender é a única coisa de que a mente  
nunca se cansa, nunca tem medo e nunca  
se arrepende."

(Leonardo da Vinci)

LINO, Euziane Joana. **Identificação e Caracterização de Transcritos Relacionados a Mecanismos de Resistência à Ferrugem Asiática da Soja**. 2011. Dissertação (Mestrado em Genética e Biologia Molecular) – Universidade Estadual de Londrina, Londrina 2011.

## RESUMO

O Brasil é o segundo maior exportador de soja no mundo, portanto esta cultura é de grande importância para o país, porém devido ao clima ameno que pode contribuir para o surgimento e estabilização de patógenos, a cultura da soja fica sujeita a desenvolver doenças como é o caso da ferrugem asiática da soja, que no Brasil tem causado grandes prejuízos, sendo uma das doenças que mais contribui para a perda na produção do grão nas últimas safras. A melhor e menos dispendiosa solução para evitar essas perdas causadas pelo fungo biotrófico *Phakopsora pachyrhizi* seria o desenvolvimento de cultivares que apresentem resistência efetiva contra a doença. Algumas cultivares já foram lançadas no país visando solucionar esse problema contendo um ou poucos genes de resistência, porém sabe-se que esta resistência não é duradoura, uma vez que patógeno e hospedeiro co-evoluem e o fungo apresenta elevada variabilidade genética. Estudos das diferentes formas de resistência da planta, com objetivo de contribuir para o desenvolvimento de estratégias na obtenção de cultivares que possam vencer o patógeno, fazem-se necessários. Uma das formas de compreender o que ocorre na planta (órgão infectado) durante o ataque do patógeno, seria o estudo dos transcritos induzidos nesse período. Alguns dos métodos que poderiam contribuir para o estudo seria a construção de bibliotecas subtrativas supressivas (SSH) associadas com tecnologias de sequenciamento tradicionais e/ou novas tecnologias. Neste trabalho, visando compreender diferentes formas de resistência contra a ferrugem asiática em soja, foram sequenciadas 6 SSH em dois genótipos de soja, sendo um com resistência qualitativa (PI561356) e o outro resistência quantitativa (BRS231). As SSH foram transformadas em bulks correspondendo aos horários iniciais (12, 24, 48h), médios (72 e 96h) e tardios (192h) de infecção. Muitos genes correspondentes a diferentes processos biológicos como resposta a estresse e estímulos, tradução e transcrição, óxido-redução e transporte foram induzidos após a infecção. Sete transcritos pertencentes a diferentes processos biológicos, como transdução de sinais, defesa e transcrição, foram selecionados para a validação das bibliotecas por RT-qPCR.

**Palavras chave:** Hibridação subtrativa supressiva. Solexa. *Phakopsora pachyrhizi*.

LINO, Euziane Joana. **Identification and Characterization of Transcripts Related to Mechanisms of Resistance to Asian Soybean Rust.** 2011. Dissertation (Masters in Genetics and Molecular Biology) – State University of Londrina, Londrina 2011.

## ABSTRACT

Brazil is the second largest exporter of soybeans in the world, so this culture is very important for the country, but due to the mild climate which can contribute to the emergence and stabilization of pathogens, the soybean crop is subject to develop diseases as the case of Asian soybean rust, which causes great losses in Brazil, one of the diseases that most contributes to the loss in grain production in the last seasons. The best and least expensive solution to avoid the losses caused by the biotrophic fungus *Phakopsora pachyrhizi* is the development of cultivars that have effective resistance against the disease. Some cultivars have been released in the country seeking to solve this problem containing one or few resistance genes, but this resistance will not be durable once the pathogen and the host co-evolve and the fungus has a high genetic variability. Studies of different forms of plant resistance, in order to contribute to the development of strategies to obtain cultivars that can win the pathogen, become necessary. One way to understand what happens in the plant body (infected) during the attack of the pathogen is the study of transcripts induced during this period. Some of the methods that could contribute to the study are the construction of suppressive subtractive libraries (SSH) associated with traditional sequencing technologies and / or new technologies. In this work, seeking to understand different forms of resistance against soybean rust, 6 SSH were sequenced in two soybean genotypes being a qualitative resistant (PI561356) and other resistant quantitative (BRS231). The SSH were transformed in bulks that corresponding start times (12, 24, 48h), medium (72 and 96h) and late (192h) of infection. Many genes corresponding to different biological processes like response to stress and stimuli, translation and transcription, redox and transport, were induced. Seven transcripts belonging to different biological processes such as signal transduction, defense and transcription, were selected for the validation of the libraries by RT-qPCR.

**Keywords:** Suppressive subtractive hybridization. Solexa. *Phakopsora pachyrhizi*.

## LISTA DE FIGURAS

### CAPÍTULO 1 - CONSIDERAÇÕES GERAIS

- Figura 1 -** Ciclo da infecção do fungo *p. pachyrhizi* em plantas hospedeiras (fonte: consórcio antiferrugem 2011a).....21
- Figura 2 -** Mapa de dispersão da Ferrugem Asiática da Soja no território brasileiro até 05 de fevereiro de 2011. Os pontos vermelhos representam os focos de FAS. ....22
- Figura 3 -** Diagrama da SSH. Linhas representam os cDNAs digeridos com *Rsa* I. Retângulos pretos representam a parte externa do adaptador 1 e 2R, iguais e complementares ao iniciador da primeira PCR. Retângulos brancos representam a parte interna do adaptador 1 e são complementares ao iniciador interno 1 da segunda PCR. Retângulos cinza representam a parte interna do adaptador 2R e são complementares ao iniciador interno 2R da segunda PCR. Adaptado de Clontech, 2008 .....30

### CAPÍTULO 2 - CARACTERIZAÇÃO DO PERFIL TRANSCRICIONAL EM RESPOSTA A INFECÇÃO POR PHAKOPSORA PACHYRHIZI EM GENÓTIPOS TOLERANTE E RESISTENTE À FERRUGEM ASIÁTICA

- Figura 1 -** Representação do delineamento experimental em casa de vegetação. ....47
- Figura 2 -** Experimento em casa de vegetação. A: plantas antes da inoculação distribuídas na casa de vegetação, B: plantas após a inoculação dos uredósporos, cobertas com sacos plásticos, para manter a umidade alta, facilitando a germinação dos esporos da FAS. ....48

<b>Figura 3 -</b>	Esquema de construção das bibliotecas subtrativas para os genótipos PI561356 e BRS231. Nas caixas verdes estão os horários após a inoculação da doença correspondentes aos “bulks” montados com os cDNA inoculado (“tester”) e falso inoculado (“driver”), em seguida o enriquecimento das sequências que se dá ao longo das duas hibridizações por PCRs.....	49
<b>Figura 4 -</b>	Esquema utilizado para análises das sequências Solexa obtidas para PI561356.....	56
<b>Figura 5 -</b>	Esquema utilizado para análises das sequências Sanger obtidas para BRS231.....	57
<b>Figura 6 -</b>	A: Diagrama de Venn obtido através da comparação das sequências da PI561356 em cada um dos bulks obtidas através do alinhamento dos “contigs” do Solexa nos ESTs e Modelos de genes de soja, B: unissequências da PI561356 após análise pelo AutoFact e C: unissequências da BRS231 após análise pelo AutoFACT. Diagramas obtidos a partir do programa Venny (Oliveros, 2007).....	58
<b>Figura 7 -</b>	Processos biológicos compartilhados e exclusivos entre os “Bulks” obtidos em ambos os genótipos. A: BRS231 e B: PI561356.....	59
<b>Figura 8 -</b>	Gráficos da categoria Processos biológicos mais comuns encontrados no genótipo PI561356. Em A: Bulk1, em B: Bulk 2 e em C: Bulk3 .....	60
<b>Figura 9 -</b>	Gráficos da categoria Processos biológicos mais comuns encontrados no genótipo BRS231. Em A: Bulk1, em B: Bulk 2 e em C: Bulk3 .....	60
<b>Figura 10 -</b>	Principais cinases envolvidas na transdução de sinais em resposta a estresse, encontradas em PI561356 (colunas em azul) e em BRS231 (colunas em cinza).....	70

- Figura 11** - Gráficos obtidos através do programa geNorm apontando os genes mais estáveis com apresentando menores valores de M. Em A: Ambos os genótipos quando analisados conjuntamente, B: em PI561356 quando analisada em separado, e em C: BRS231 quando analisada separadamente..... 78
- Figura 12** - Níveis de expressão gênica nos dois genótipos de soja (BRS231 e PI561356) em seis horários após inoculação com FAS. A: Ankyrin, B: Disease, C: KR4, D:LRR, E: MPKK, F: Serina/Treonina e G: WRKY23. .... 80

## LISTA DE TABELAS

<b>Tabela 1</b> -Quantidade de “reads”, “contigs” e “singlets” e tamanhos dos ESTs das bibliotecas subtrativas dos genótipos PI561356 e BRS231.....	53
<b>Tabela 2</b> -Quantidade de sequências conhecidas e “no hits” nos genótipos PI561356 e BRS231.....	54
<b>Tabela 3</b> -Processos Biológicos relacionado a defesa e resposta a estress expressos em reposta a FAS nos genótipos PI561356 e BRS 231.....	62
<b>Tabela 4</b> -Fatores de Transcrição mais representativos em PI561356 e BRS231.....	73
<b>Tabela 5</b> -Valores de estabilidade obtidos através do programa NormFinder, aonde os melhores genes normalizadores são os que apresentam os valores mais próximos a 0. ....	76
<b>Tabela 6</b> -Transcritos escolhidos para a validação das bibliotecas. A identificação corresponde as sequências das bibliotecas SSH de BRS231 depositadas no banco de dados da Embrapa Soja e ao Gene model do Phytozome.....	79

## LISTA DE ABREVIATURAS E SIGLAS

<b>ABA</b>	Ácido Abscísico
<b>ABC transporter</b>	Atp binding cassette transporter
<b>AJ</b>	Ácido Jasmônico
<b>ARF</b>	Auxin response factor
<b>AS</b>	Ácido Salicílico
<b>AutoFACT</b>	Automatic Annotation and Classification Tool
<b>avr</b>	Gene de avirulência
<b>CC</b>	Coiled coil
<b>CDPK</b>	Proteína cinase dependente de cálcio
<b>DRP</b>	Disease Resistance Protein
<b>EDS1</b>	Enhanced Disease Susceptibility1
<b>ELF1b</b>	Fator de alongação 1b
<b>EST</b>	Expressed Sequence Tags ou Sequências de Etiquetas Expressas
<b>FAS</b>	Ferrugem asiática da soja
<b>GAPDH</b>	Gliceraldeído-3-fosfato desidrogenase
<b>hai</b>	Horas após a inoculação
<b>HR</b>	Resposta de hipersensibilidade
<b>LRR</b>	Leucine rich repeat
<b>MAMPs/PAMPs</b>	Padrões moleculares associados ao microrganismo ou ao patógeno
<b>MAPK</b>	Mitogen Activated Protein Kinase
<b>MPKK</b>	Mitogen Activated Protein Kinase Kinase
<b>MPSS</b>	Massively Parallel Signature Sequencing
<b>NBS-LRR</b>	Nucleotide Binding Site – Leucine Rich Repeat
<b>NPR1</b>	Non-Expressor of PR1
<b>PAD4</b>	Phytoalexin deficient4
<b>PAL</b>	Fenilalanina Amonia-Liase
<b>pb</b>	Pares de base
<b>PCR</b>	Polymerase Chain Reaction ou Reação em Cadeia da Polimerase
<b>PR-1</b>	Pathogen Related 1

<b>R</b>	Gene de resistência
<b>RB</b>	Reddish-Brown ou castanho-avermelhada
<b>RLK</b>	Receptor-like kinase
<b>ROS</b>	Espécies reativas de oxigênio
<b>RPKM</b>	Reads por kilobase por milhão
<b>Rpp</b>	Resistance to <i>Phakopsora pachyrhizi</i> ou Resistência a
<b>Phakopsora</b>	<i>pachyrhizi</i>
<b>RT-qPCR</b>	Real time quantitative Polymerase Chain Reaction ou Reação em Cadeia de Polimerase quantitativo em tempo real
<b>SAGE-</b>	Serial Analysis for Gene Expression
<b>Ser/Ter</b>	Serina/Treonina
<b>SNPs</b>	Single Nucleotide Polymorphism
<b>SOAP</b>	Short Oligonucleotide Alignment Program
<b>SSH</b>	Hibridização Subtrativa Supressiva
<b>TAN</b>	Castanho-clara
<b>TIR</b>	Domínio “ <i>Drosophila Toll</i> ” e receptor de Interleucina-1 humana

## SUMÁRIO

<b>CAPÍTULO 1 – CONSIDERAÇÕES GERAIS</b> .....	17
<b>1 INTRODUÇÃO</b> .....	17
<b>2 FUNDAMENTAÇÃO TEÓRICA</b> .....	19
2.1 SOJA: ASPECTOS CULTURAIS E ECONÔMICOS .....	19
2.2 FERRUGEM ASIÁTICA DA SOJA .....	20
2.3 RESPOSTA DE RESISTÊNCIA A DOENÇAS EM PLANTAS.....	23
2.4 ESTUDOS MOLECULARES: EXPRESSÃO GÊNICA.....	27
2.4.1 Estudos de Expressão Gênica em Soja Infectada por Ferrugem Asiática .....	31
2.5 AVANÇOS NAS TECNOLOGIAS DE SEQUENCIAMENTO .....	33
<b>3 OBJETIVO GERAL</b> .....	37
3.1 OBJETIVOS ESPECÍFICOS.....	37
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	38
<b>CAPÍTULO 2 – CARACTERIZAÇÃO DO PERFIL TRANSCRICIONAL EM RESPOSTA A INFECÇÃO POR PHAKOPSORA PACHYRHIZI EM GENÓTIPOS TOLERANTE E RESISTENTE À FERRUGEM ASIÁTICA</b> .....	44
<b>RESUMO</b> .....	44
<b>ABSTRACT</b> .....	45
<b>1 INTRODUÇÃO</b> .....	45
<b>2 MATERIAL E MÉTODOS</b> .....	46
2.1 MATERIAL BIOLÓGICO E DESENHO EXPERIMENTAL.....	46
2.2 PREPARO DOS INÓCULOS E INOCULAÇÃO .....	47

2.3	CONSTRUÇÃO DAS BIBLIOTECAS SUBTRATIVAS.....	48
2.4	SEQUENCIAMENTO E BIOINFORMÁTICA.....	49
2.5	VALIDAÇÃO DOS GENES POR RT-qPCR .....	51
<b>3</b>	<b>RESULTADOS E DISCUSSÃO .....</b>	<b>53</b>
3.1	ANÁLISE COMPARATIVA ENTRE OS SEQUENCIAMENTOS POR SOLEXA E SANGER .....	53
3.2	CATEGORIZAÇÃO FUNCIONAL .....	57
3.2.1	Categoria Funcional: Defesa e Resposta a Estresse Bióticos e Abióticos .....	63
3.2.2	Categoria Funcional: Cinases.....	68
3.2.3	Categoria Funcional: Regulação da Expressão Gênica .....	70
3.2.4	Categoria Funcional: Modificação/Catabolismo de Proteínas .....	73
3.3	SEQUÊNCIAS DESCONHECIDAS “UNKNOW” E “NO HITS” .....	75
3.4	RT-QPCR: IDENTIFICAÇÃO DE GENES NORMALIZADORES PARA ESTUDOS DE EXPRESSÃO GÊNICA EM RESPOSTA A FAS POR RT-QPCR .....	76
3.6	RT-QPCR: VALIDAÇÃO DOS GENES IDENTIFICADOS NAS BIBLIOTECAS VIA HIBRIDAÇÃO SUBTRATIVA .....	77
	<b>CONCLUSÃO .....</b>	<b>85</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>93</b>

## CAPÍTULO 1 CONSIDERAÇÕES GERAIS

### 1 INTRODUÇÃO

A soja [*Glycine max* (L.) Merrill] é considerada uma das culturas mais importantes do mundo, devido a ampla utilização de seus grãos na alimentação humana e animal, na produção de óleos e biocombustíveis, como uma alternativa aos combustíveis não renováveis.

A soja é o principal produto que movimenta o mercado de exportações de grãos no Brasil, sendo este o segundo maior exportador de soja e a China um de seus principal importadores. Porém, existem fatores que influenciam negativamente a produção de soja, causando grandes perdas como, por exemplo, o estresse biótico causado pela Ferrugem Asiática da Soja (FAS).

A FAS é causada pelo fungo *Phakopsora pachyrhizi* (Sydow & Sydow), um dos principais patógenos da soja atualmente, pois pode atacar a planta em qualquer estágio de seu desenvolvimento causando perdas de até 80% em sua produção. O controle desse fungo é basicamente feito através de aplicações de fungicidas, que oneram a manutenção da cultura. Uma estratégia de controle do patógeno, que apresenta menor custo e menores danos ao ambiente, seria a obtenção de cultivares capazes de resistir ao ataque da doença e garantir a produção de grãos de qualidade para o mercado.

A compreensão dos mecanismos fisiológicos bem como moleculares envolvidos na resistência de uma planta aos seus patógenos pode auxiliar na definição de estratégias para o desenvolvimento de cultivares melhoradas, visando a eliminar os prejuízos financeiros e colapso na produção de alimentos.

Sabe-se que a resistência qualitativa ou raça-específica é mediada por um ou poucos genes de efeito maior, portanto é mais passível de ser quebrada por diferentes raças fisiológicas do patógeno, já a resistência quantitativa ou raça inespecífica é mediada por mais de um gene, genes de menor efeito, que quando somados contribuem aditivamente para o fenótipo de resistência. A união desses dois tipos de resistência, em uma mesma planta, é almejado no manejo da

resistência a doenças em muitos patossistemas, o que garantiria uma resistência mais duradoura.

A análise de transcriptomas através da geração das sequências de etiquetas expressas (“Expressed Sequence Tags” ou ESTs) por Hibridização Subtrativa Supressiva (“Suppressive Subtractive Hybridization” ou SSH) associada a novas tecnologias de sequenciamento tem permitido grandes avanços e rapidez na identificação de genes induzidos na resposta de resistência, o que tem auxiliado nos estudos da interação molecular entre patógenos e hospedeiros.

O presente trabalho teve como objetivo a identificação e caracterização de sequências induzidas em soja ao longo do ciclo infectivo a FAS, nos genótipos BRS231 e PI561356, que apresentam resistência à infecção com *P. pachyrhizi*, respectivamente (RIBEIRO *et al.*, 2007; CAMARGO, 2010). A subtração diferencial unidirecional entre os tratamentos inoculado e falso inoculado permitiu a identificação de 3.277 e 2.018 sequências induzidas em soja nos genótipos PI561356 e BRS231, respectivamente. As ESTs foram categorizadas em pelo menos 326 processos biológicos para PI561356 e 558 para BRS231, contendo pelo menos um transcrito em cada categoria. Os processos biológicos contendo o maior número de sequências estão relacionados a resposta a stress, produção de metabólitos secundários, transporte e tradução.

A identificação dos genes altamente expressos, baseados nos valores de RPKM (Reads por Kilobase por Milhão), bem como as principais rotas biológicas ativadas após a infecção, permitirão que análises futuras possam definir os principais atores que medeiam a resposta de resistência e tolerância a FAS nos genótipos PI5651356 e BRS131, bem como o emprego futuro de novas alternativas no desenvolvimento de variedades resistentes ao fungo.

## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 SOJA: ASPECTOS CULTURAIS E ECONÔMICOS

Por apresentar grande destaque mundial, a soja é amplamente estudada no intuito de compreender e quantificar as relações ecofisiológicas e moleculares que interferem na produção da cultura, devido aos desafios encontrados em sua adaptação nos diferentes ambientes (POPP, DILLON e KEISLING, 2003).

A soja, conhecida atualmente, provém do cruzamento de dois ancestrais com porte rasteiro, oriundos das margens do rio Yangtse na China, que foram domesticados e melhorados por cientistas da antiga China. (EMBRAPA, 2009a).

A soja foi introduzida no Brasil, no estado do Rio Grande do Sul, em 1950, para cultivo massivo pelos imigrantes alemães e italianos que ali se instalaram, devido à necessidade de povoar os vazios demográficos para que o Paraguai, Argentina e Uruguai não tomassem as terras brasileiras. Ali se encontravam ótimas condições para o cultivo desse grão, que era cultivado em rotação com o trigo (MUELLER e BUSTAMANTE, 2002), pois o solo, composto por terra roxa, apresentava alta fertilidade. Além disso, políticas públicas nacionais de desenvolvimento agrícola estavam favorecendo o cultivo nas regiões sul e sudeste (SIEBEN e MACHADO, 2007).

Atualmente, a cultura de soja ocupa um lugar de destaque no mercado de “commodities” do Brasil, sendo este o segundo maior produtor de soja do mundo, com uma produção nacional de 68,69 milhões de toneladas na safra 2009/10, e estimativa de 68,55 milhões de toneladas para a safra 2010/11, aproximadamente, sendo 0,2% ou em torno de 134,9 mil hectares inferior a safra anterior (CONAB, 2011). Além disso, a soja lidera os produtos agropecuários na pauta de exportação do Brasil.

A expansão da soja para áreas agrícolas brasileiras com diferentes condições edafo-climáticas acarretou numa maior demanda por tecnologias de apoio à produção dos grãos em locais onde essa cultura é negativamente afetada por fatores bióticos e abióticos. Esforços biotecnológicos tem sido feito para desenvolver novas cultivares com características que contribuam para minimizar perdas na sua

produção, contribuindo para o esclarecimento genético de respostas adaptativas perante estresses que pode ser utilizada para melhorar, por exemplo, a resistência a doenças e a tolerância à seca em plantas (STOLF, 2007).

Com a expansão da cultura para novas áreas o número de doenças causadas por bactérias, fungos, nematóides e vírus na soja aumentaram, sendo que, até o momento, cerca de 50 doenças já foram identificadas na cultura da soja. Atualmente, a ferrugem asiática da soja (FAS) é considerada uma das maiores causadoras de danos para a cultura da soja no mundo, ocasionando perdas acumuladas estimadas em sete bilhões de dólares (EMBRAPA, 2009b).

## 2.2 FERRUGEM ASIÁTICA DA SOJA

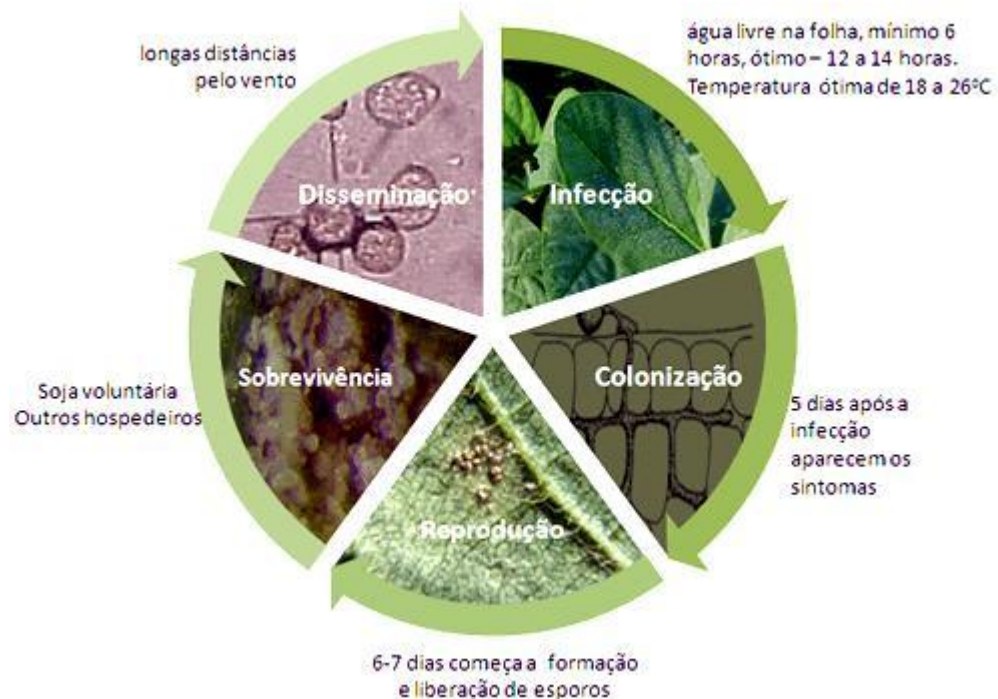
A ferrugem asiática tem como agente causador o fungo *Phakopsora pachyrhizi*, que ataca não somente a soja, mas vários hospedeiros de aproximadamente 95 espécies, distribuídas em 42 gêneros da família dos feijões (Fabaceae). Os primeiros relatos de ferrugem asiática datam de 1902 no Japão. Na África do Sul, a doença foi identificada em 2001, tendo recorrência em 2002, causando perdas na produção de grãos de soja entre 60% a 80%, e até 100% em locais onde não havia rotação de culturas (CALDWELL e LAING, 2002).

Os danos causados pelo fungo são pequenas lesões foliares e, à medida que ocorre a esporulação, o tecido da folha adquire coloração castanho-clara (lesão tipo TAN, característica de genótipos suscetíveis) a castanho-avermelhada (lesão tipo RB – “reddish-brown”, característica de genótipos resistentes) (BROMFIELD, 1982). O acúmulo das lesões causa rápido amarelecimento, bronzeamento ou crestamento (queimaduras) e queda prematura das folhas, impedindo a plena formação dos grãos que não amadurecem por completo e apresentam baixo peso (YORINORI *et al.*, 2004).

*P. pachyrhizi* é um patógeno biotrófico, pertencente à Classe Pucciniomycetes, Ordem Pucciniales, e Família Phakopsoraceae (INDEX FUNGORUM, 2011). Ao contrário de outras ferrugens como *Uromyces appendiculatus* (ferrugem do feijoeiro), que necessita identificar uma superfície específica para penetrar na planta, no caso, os estômatos (WYNN, 1976), a FAS pode penetrar diretamente pela cutícula das plantas hospedeiras (ZAMBOLIN, 2006).

O ciclo da doença (Figura 1) se inicia com o transporte dos uredósporos (esporos assexuados da ferrugem) pelo vento e posteriormente sua germinação na superfície abaxial da folha aproximadamente seis horas após o contato com esta (ZAMBOLIN, 2006). Com a germinação forma-se o tubo germinativo que termina em uma estrutura dilatada (com praticamente o mesmo tamanho do uredósporo) denominada apressório. Dessa estrutura é lançada uma hifa de infecção rumo ao interior da célula, que se ramifica formando o haustório, cujo contato com a membrana da célula hospedeira, permite a aquisição de nutrientes para o fungo (PANSTRUGA, 2003). A formação do primeiro haustório se dá de 24 a 48 horas após o contato do uredósporo com a célula hospedeira de uma planta suscetível. O ciclo da doença está completo quando uma nova urédia (estrutura de reprodução e acomodação dos esporos) é originada da primeira urédia a surgir no hospedeiro. (KOCH *et al.*, 1983). Cada urédia permanece produzindo esporos por aproximadamente 21 dias (CONSÓRCIO ANTIFERRUGEM, 2011a).

**Figura 1** - Ciclo da infecção do fungo *P. pachyrhizi* em plantas hospedeiras (Fonte: Consórcio Antiferrugem 2011a).



No Brasil, a FAS acometeu primeiramente o norte e o noroeste do Paraná. Posteriormente, a doença se espalhou para muitos campos brasileiros principalmente nas regiões sul, sudeste e centro-oeste, com um potencial de

redução de produtividade maior que 75% (SILVA, MAIA e MEYER, 2004; YORINORI, NUNES JÚNIOR e LAZZAROTTO, 2004). Em 2002, as perdas brasileiras de soja pela FAS superaram 31 milhões de toneladas de grãos e os investimentos para tentar combater a doença ultrapassaram US\$ 11 bilhões. O número de focos detectados, até 05 de fevereiro de 2011 referentes à safra 2010/2011, já chega a 430 relatos (Figura 2) (CONSÓRCIO ANTIFERRUGEM, 2011b).

**Figura 2** - Mapa de dispersão da Ferrugem Asiática da Soja no território brasileiro até 05 de fevereiro de 2011. Os pontos vermelhos representam os focos de FAS



**Fonte:** Consórcio Antiferrugem 2011c.

Algumas medidas preventivas vêm sendo tomadas visando à diminuição das perdas no cultivo da soja por *P. pachyrhizi*, como o vazio sanitário vegetal nas regiões produtoras de soja, que estabelece a entressafra sem plantas de soja, para evitar a ponte verde, o que significa reduzir inóculos do fungo no ambiente, para retardar a entrada da doença nos próximos cultivos (BARROS, 2008). Além disso, o estudo e desenvolvimento de novas variedades de soja que apresentam resistência à ferrugem asiática é uma estratégia para o manejo da doença no Brasil, uma vez que o controle dessa doença é feito basicamente por

fungicidas que oneram a produção e podem prejudicá-la caso as aplicações não sejam corretamente elaboradas (ARIAS *et al.*, 2008).

Alguns genes de resistência à ferrugem asiática são conhecidos há bastante tempo: *Rpp1*, *Rpp2*, *Rpp3* e *Rpp4* (BROMFIELD & HARTWIG, 1980; HARTWIG, 1986). Recentemente o gene *Rpp5* foi descrito e mapeado no grupo de ligação N da soja (GARCIA *et al.*, 2008). Interações incompatíveis (resistência) mediadas pelo alelo *Rpp1* apresentam um fenótipo de imunidade contra alguns isolados (MILES *et al.*, 2006), enquanto resistência conferida pelos outros quatro genes é caracterizada por crescimento e esporulação limitada do fungo e a formação de lesões RB (BONDE *et al.*, 2006). A resistência conferida pelos genes *Rpp1* e *Rpp3* foi quebrada pela raça fisiológica do fungo encontrada atualmente no Brasil, enquanto *Rpp2* e *Rpp4* continuam efetivos no controle da doença. No momento, muito pouco se sabe sobre como os mecanismos de resistência são mediados por esses genes.

O entendimento dos mecanismos moleculares envolvidos na interação planta-patógeno é de extrema importância no planejamento de estratégias alternativas para o controle dessa doença (MATSUMURA *et al.*, 2003). A aplicação de técnicas de análise e identificação de transcritos expressos durante a interação patógeno-hospedeiro tem levado a importantes avanços no entendimento dos mecanismos de patogenicidade, e está em atual expansão à medida que sequências gênicas do hospedeiro e patógeno vêm sendo disponibilizadas (MATSUMURA *et al.*, 2005).

### 2.3 RESPOSTA DE RESISTÊNCIA A DOENÇAS EM PLANTAS

A resposta de resistência pode ser entendida fisiologicamente como o potencial da planta em evitar ou retardar a entrada e subsequente infecção de um patógeno em seus tecidos, visto que na natureza, devido à exposição da planta a um grande número de microorganismos, a resistência se faz regra, enquanto a suscetibilidade aos patógenos mostra-se como exceção (PASCHOLATI & LEITE, 1995).

Existem diferentes concepções sobre resistência que devem ser destacados: o primeiro é a resistência monogênica ou de gene de efeito maior, que é mediada por um único gene capaz de conferir a resposta de resistência. Ela

também é denominada resistência qualitativa, pois o fenótipo apresentado pela planta resistente é bem característico e distinto da planta suscetível - é a resistência do tipo “tudo ou nada”, pois a planta apresenta ou não a doença. Já a resistência poligênica, também conhecida como quantitativa, é mediada por mais de um gene sendo que a soma dos seus efeitos confere o caráter de resistência. A característica da resistência quantitativa é a presença da variação contínua nas populações vegetais referente ao grau da doença, com plantas que apresentam desde a extrema suscetibilidade até a resistência, por isso, é um caráter que pode ser quantificado e não facilmente qualificado. A resistência qualitativa pode ser mais facilmente vencida por se tratar de um único gene que medeia a resposta, o que não acontece com a resistência quantitativa, uma vez que o patógeno teria que burlar vários genes para a ocorrência da doença (CAMARGO, 1995).

Existe ainda o conceito epidemiológico de resistência, que define como resistência horizontal ou raça inespecífica, aquela eficaz para uma ampla gama de raças do patógeno e a resistência vertical como aquela eficaz para uma única, ou poucas raças do patógeno. Comumente a resistência vertical é monogênica e a horizontal é poligênica, o que tem levado a confusões quanto aos conceitos. No entanto, a resistência horizontal pode ser conferida por um único gene e a vertical por vários genes (CAMARGO & BERGAMIN FILHO, 1995).

Entre resistência e suscetibilidade encontra-se um termo muito utilizado quanto ao caráter agrônomico da planta em relação a uma doença. Esse termo é a tolerância. O que caracteriza uma planta tolerante é a capacidade de suportar o patógeno mesmo desenvolvendo a doença, mas sem que ocorram perdas significativas na produção. Tolerância é muitas vezes confundida com resistência horizontal ou resistência quantitativa, porém, a fisiologia da tolerância não é bem compreendida e pode ser mediada por mono ou poligenes (CAMARGO, 1995).

A interação planta-patógeno se dá, segundo Flor (1971), através do reconhecimento do produto de um gene de avirulência (*avr*), o elicitor, que está presente no patógeno, pelo produto de um gene de resistência (*R*) que está presente na planta hospedeira, ou seja, ocorre o reconhecimento direto do elicitor pela proteína de resistência. A interação do produto gênico *avr* com o produto gênico *R* acarreta em uma interação incompatível, que culmina em uma resposta de hipersensibilidade (HR), que através da transdução de sinais mediada pelo reconhecimento do elicitor do patógeno (gene *avr*) desencadeia a morte celular

programada, bloqueando a reprodução e dispersão do patógeno para as demais células. Acontece um “altruísmo celular” na planta infectada, onde algumas células morrem para que as outras sobrevivam. Ao contrário da interação incompatível onde não acontece a doença, na interação compatível o gene *avr* não é reconhecido e, portanto, não é elicitada resposta específica de defesa caracterizando a planta como suscetível à doença.

O reconhecimento do elicitor pela planta hospedeira leva a uma cascata de eventos de transdução de sinais que irá ativar genes da célula vegetal hospedeira, resultando em uma rápida resposta de defesa. Em alguns casos, o patógeno pode não ser reconhecido pelo hospedeiro e, portanto, não ocorrerá indução dos mecanismos de defesa, ou este pode ocorrer lentamente, possibilitando o crescimento do patógeno e, conseqüentemente, o desenvolvimento da doença (TAIZ e ZEIGER, 2009).

A reação de hipersensibilidade é uma das respostas de defesa gerada principalmente pela resistência qualitativa e, muitas moléculas e fitormônios estão envolvidas na transdução de sinais que a desencadeia, tais como: o ácido salicílico (AS), etileno, ácido jasmônico (AJ), ácido abscísico (ABA), óxido nítrico e as espécies reativas de oxigênio (ROS) derivadas da combustão oxidativa nas células (TORRES *et al.*, 2006; ADIE *et al.*, 2007; KOORNNEEF & PIETERSE, 2008). A partir da teoria de Flor, muitas outras hipóteses foram propostas e com base em diferentes patossistemas, algumas levando em consideração essa teoria inicial, outras, pressupondo que em muitos casos o que ocorre não é a interação direta, mas sim, uma interação indireta entre proteína R e elicitor.

No modelo guarda ou protetor os efetores do patógeno se ligam a proteínas alvos do hospedeiro, essas proteínas podem ser tanto proteínas que compõe a defesa da planta [proteínas com domínio serina/treonina (Ser/Tre), por exemplo] como outras proteínas que podem ter a função modificada para nutrir o patógeno. As proteínas guardas ou protetoras da planta são as Nucleotide Binding Site – Leucine Rich Repeat (NBS-LRR), cuja função seria ligar-se indiretamente aos elicitores do patógeno e desencadear os mecanismos de defesa, uma vez que NBS-LRR se liga as proteínas alvos da planta (proteínas guardadas ou protegidas), pois só faz o reconhecimento do elicitor através do conjunto proteína alvo - elicitor. Nesse cenário, as proteínas R monitoram se as proteínas celulares estão sob ataque do patógeno. Portanto o elicitor só é reconhecido intracelularmente e indiretamente pela

ligação a proteína hospedeira alvo do elicitor e esta com a proteína R (DANGL E JONES, 2001).

O modelo do chamariz (van der HOORN e KAMOUN, 2008) é similar ao modelo guarda de interação indireta, e explica que a proteína alvo do efector (proteína protegida), de um ponto de vista evolutivo, encontra-se em uma situação instável e está sujeita a duas forças naturais de seleção opostas, na população de plantas com genes R polimórficos: uma na ausência do gene R, onde a proteína protegida sofre seleção para diminuir a afinidade de ligação ao efector evitando a sua detecção e modificação pelo efector. A segunda seleção seria na presença do gene R, onde a proteína protegida tem a capacidade de interação aumentada com o efector, acentuando a percepção do patógeno pela planta. Essas duas pressões de seleção conflitantes, no mesmo plano de interação entre efector e protegido, resulta numa situação evolucionariamente instável (trade off) que pode ser amenizada sob a evolução de uma outra proteína hospedeira denominada “chamariz”. Chamariz tem a função de ser reconhecido pelo elicitor, mimetizando a proteína protegida e também ser reconhecida pela proteína R quando ligada ao elicitor para ativação dos mecanismos de resposta imune, porém nenhum benefício é trazido ao patógeno quando o elicitor se liga ao chamariz, o que promove a diminuição do “fitness” do patógeno em relação a planta, quando na ausência dos genes R.

Na interação patógeno-hospedeiro a maioria dos eventos que leva a suscetibilidade ou à resistência da planta hospedeira ao patógeno ocorre em nível celular. Existem duas amplas categorias de defesa: a estrutural, morfológica ou anatômica e a bioquímica ou fisiológica, ambas podem ser pré-formadas, quando já estão presentes na planta, ou pós-formadas, que serão produzidas mediante ativação de genes (INGHAM, 1973). Os fatores estruturais da planta atuam como barreiras físicas, impedindo a entrada do patógeno e a colonização dos tecidos, enquanto as reações químicas produzem substâncias tóxicas ao patógeno ou criam condições adversas ao crescimento deste no interior do hospedeiro. Como exemplo de defesa estrutural pode-se citar cutícula, tricomas e a própria parede celular, bem como lignificação da parede celular. Em relação aos mecanismos bioquímicos de defesa têm-se como exemplo os compostos fenólicos, saponinas, inibidores de proteinase (mecanismos pré-formados), e espécies reativas de oxigênio e fitoalexinas (mecanismo pós-formados) e proteínas antimicrobianas (PASCHOLATI & LEITE, 1995).

O principal hormônio envolvido em respostas de defesa e transdução de sinais contra patógenos biotróficos é o ácido salicílico (SCHWEIGHOFER *et al.*, 2007). Alguns genes desenvolvem papéis chave na indução e acumulação desse hormônio na planta, mediante infecção por patógenos como é o caso do produto gênico codificado por *NPR1* (Non-Expressor of PR1), *PAD4* (phytoalexin deficient4) e *EDS1* (Enhanced Disease Susceptibility1) entre outros genes. *PAD4* codifica uma lipase envolvida na acumulação de AS e camalexina (fitoalexina) (JIRAGE *et al.*, 1999). *EDS1* também codifica uma lipase que está envolvida na acumulação de AS, e mutação nesse gene causa suscetibilidade em plantas inoculadas com *Peronospora parasitica* (FALK, *et al.*, 1999).

Recentemente um estudo de silenciamento gênico em plantas de soja contendo *Rpp2*, contendo *NPR1*, *PAD4* e *EDS1*, genes já sabidos estarem envolvidos na rota de sinalização de defesa contra a ferrugem asiática. Quando esses genes foram silenciados separadamente, as plantas resistentes ao isolado utilizado tornaram-se suscetíveis a ferrugem. Houve também, a “down” regulação de vários fatores de transcrição com domínios WRKY e MYB que são importantes na regulação de vários processos biológicos da planta, incluindo a defesa (PANDEY *et al.*, 2011).

O maior desafio é selecionar genes candidatos para serem estudados em detalhes entre os milhares de genes presentes no genoma que apresenta alternativas promissoras para o desenvolvimento de estratégias de controle do patógeno. A análise de transcriptomas é uma abordagem muito útil para este propósito.

## 2.4 ESTUDOS MOLECULARES: EXPRESSÃO GÊNICA

Devido à grande quantidade de estudos com RNA (pela facilidade das técnicas empregadas ao RNA e por ser a primeira etapa na expressão gênica) comparados com os estudos de proteínas, a primeira aplicação da genômica em muitos organismos é catalogar e medir a atividade transcricional. Com a disponibilidade das sequências gênicas, há o interesse em conseguir recursos e tecnologias capazes de analisar a atividade e a interação transcricional em larga escala em diferentes espécies de plantas (MEYERS *et al.*, 2004). Desse modo, nos

últimos anos várias técnicas vêm sendo desenvolvidas para permitir a análise global e/ou pontual da expressão gênica nos mais diversos organismos.

Tecnologias como SAGE (“Serial Analysis for Gene Expression”) (VELCULESCU *et al.*, 2005), MPSS (“Massively Parallel Signature Sequencing”) (BRENNER *et al.*, 2000) ou SSH (“Supressive Subtractive Hibridization”) (DIATCHENKO *et al.*, 1996) não requerem conhecimento prévio das sequências dos transcritos e, dessa forma, é possível a identificação de transcritos desconhecidos (NIELSEN *et al.*, 2003; TALLA *et al.*, 2003). Essa característica define uma arquitetura do tipo aberta para análise da expressão gênica, ou seja, o organismo não precisa ter genoma ou transcritos sequenciados. Em contrapartida, arquiteturas fechadas, como a maioria dos microarranjos de DNA e PCR em tempo real são baseadas na existência do conhecimento prévio (sequências) dos genes, com sondas desenhadas para se ligarem a transcritos conhecidos (KUO *et al.*, 2002). Os dados derivados a partir de tecnologias abertas são usados para anotação de sequências genômicas, enquanto dados de tecnologias fechadas são mais facilmente usados para experimentos focalizados (YUEN *et al.*, 2002).

Acompanhando os avanços realizados para o sequenciamento de DNA, já existem técnicas que permitem sequenciar RNA diretamente. Esta tecnologia conhecida como RNAseq permite o sequenciamento com ou sem a amplificação do RNA tem permitido uma análise global de transcriptoma sem a necessidade de hibridização a partir de pequenas quantidades de RNA. Outro fator importante é a distinção das diferentes isoformas de genes que esse método tem possibilitado analisar. Porém, é um método novo que ainda precisa de melhorias em alguns aspectos, como nas análises de bioinformática (WANG *et al.*, 2009; HAAS & ZODY, 2010).

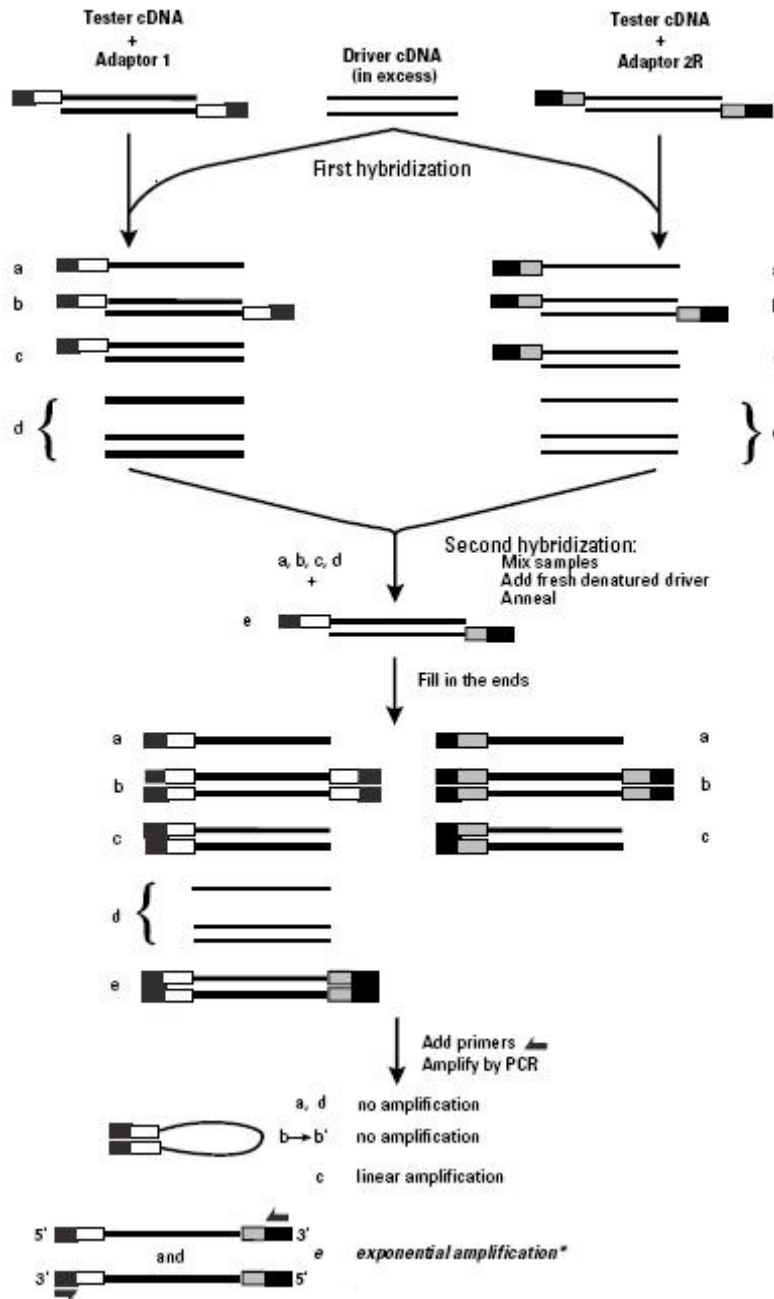
Entre os métodos apresentados acima, a hibridização subtrativa supressiva (SSH), que é o foco de estudo deste trabalho, é capaz de enriquecer uma população de cDNAs com transcritos diferencialmente expressos, pois é um método baseado em PCR, o que permite enriquecer em 1000 a quantidade de transcritos diferencialmente expressos, permitindo assim a identificação de transcritos raros (JI & CAI, 2004).

Na SSH o cDNA é sintetizado a partir de RNAm obtido das duas amostras que se pretende comparar: a testadora (“tester”) e a referência (“driver”). O mecanismo de subtração ocorre a partir de duas hibridizações por PCR: a amostra

testadora é subdividida em duas partes, cada uma receberá um adaptador diferente (a sequência do adaptador é complementar aos iniciadores que serão utilizados posteriormente para amplificação por PCR). Na primeira hibridização o cDNA controle (“driver”) é adicionado em excesso a cada amostra do testador, que são então desnaturadas e se complementam, gerando as moléculas ‘a’, ‘b’, ‘c’ e ‘d’ em cada amostra (Figura 3). As moléculas que hibridizarem com o cDNA controle são as não diferencialmente expressas, denominadas “housekeeping”, pois tem um nível de expressão constitutivo na célula. Outras moléculas que se hibridizam são genes que estão sendo induzidos tanto na condição controle como na condição testada. Após o enriquecimento e equalização dos transcritos ocorre a segunda hibridização que é subtrativa; as duas amostras são misturadas, as moléculas diferencialmente expressas que são complementares e contendo adaptadores diferentes formarão um novo tipo de híbrido: as moléculas ‘e’ (Figura 3), e terão o produto exponencialmente amplificado uma vez que os dois iniciadores irão maximizar a amplificação dessas moléculas e as que apresentarem apenas um iniciador, moléculas ‘c,’ serão amplificadas de forma linear (CLONTECH, 2008).

As moléculas ‘a’ e ‘d’ não podem ser amplificadas, por não possuírem o sítio de complementação dos iniciadores (Figura 3). Devido ao efeito de supressão da PCR, a maioria das moléculas ‘b’ forma uma estrutura em grampo, que não permite sua amplificação. As moléculas ‘e’ amplificadas podem ser diretamente inseridas em um vetor T/A, o qual é adaptado para produtos de PCR, para a clonagem e obtenção de uma biblioteca subtrativa (DIATCHENKO *et al.*, 1996).

**Figura 3** - Diagrama da SSH. Linhas representam os cDNAs digeridos com *Rsa* I. Retângulos pretos representam a parte externa do adaptador 1 e 2R, iguais e complementares ao iniciador da primeira PCR.



Retângulos brancos representam a parte interna do adaptador 1 e são complementares ao iniciador interno 1 da segunda PCR. Retângulos cinza representam a parte interna do adaptador 2R e são complementares ao iniciador interno 2R da segunda PCR.

**Fonte:** Adaptado de Clontech, 2008 .

Outra vantagem da técnica de SSH é a capacidade de contemplar tanto os genes induzidos como os genes reprimidos com apenas algumas mudanças no protocolo. Portanto podem-se estudar as duas formas de genes diferencialmente

expressos com a inversão das amostras: o “driver” passa a ser o “tester”, portanto recebem adaptadores, e a amostra que antes era “tester” passa a ser “driver” sem adaptadores e colocados em excesso (CLONTECH, 2008).

Uma das estratégias mais utilizadas para quantificar e validar a expressão de genes é a PCR em Tempo Real (RT-qPCR), que permite o acompanhamento, ciclo a ciclo, da amplificação das amostras (PFAFFL, 2001), pois os dados são lidos através da fluorescência que aumenta de acordo com a síntese de DNA, essa fluorescência é captada por mecanismos ópticos presentes no próprio termociclador que está acoplado em um computador que recebe os dados em tempo real.

Existem várias estratégias para marcação fluorescente dos produtos da RT-qPCR, porém uma das mais utilizadas é o SYBR Green<sup>®</sup> devido ao menor custo em relação as outras estratégias em RT-qPCR. O SYBR Green é um agente intercalante de fita dupla de DNA semelhante ao brometo de etídeo, portanto quanto maior o número de moléculas de fitas duplas, maior será a intensidade de fluorescência captada, porém esta técnica apresenta uma desvantagem caso não esteja devidamente otimizada pois o SYBR Green se liga a qualquer fita dupla de DNA sendo ela a molécula alvo de estudo ou não, podendo se ligar ao DNA inespecífico, como DNA de outros organismos e dímeros de “primers” que esteja contaminando as amostras. Uma forma encontrada para evitar contaminação de corridas anteriores com SYBR Green é a amplificação do DNA com uracila ao invés de timina durante a PCR, isso para que em reações posteriores, o DNA de reações anteriores, que se torna um contaminante, seja degradado pela uracil DNA glicosilase, uma enzima capaz de degradar DNA com uracila (GIULIETTI et al., 2001).

#### 2.4.1 Estudos de Expressão Gênica em Soja Infectada por Ferrugem Asiática

A técnica de SSH tem sido utilizada para analisar a interação molecular de plantas com diferentes patógenos, incluindo soja, como na resposta a infecção pelo vírus do mosaico da soja (JEONG *et al.* 2005). Em relação a ferrugem, Silva (2008) utilizou a técnica SSH para a análise da interação patógeno-hospedeiro em plantas de soja do genótipo PI230970, que carrega o alelo de resistência *Rpp2*,

e de um genótipo suscetível, Embrapa 48. Em outro estudo, Choi et al (2008) investigou a expressão de genes de soja na PI200492, que carrega o alelo *Rpp1* identificando uma grande quantidade de lipoxigenases e peroxidases nessa interação.

Mais recentemente, a técnica de SSH também foi utilizada para estudo de resposta a FAS em outra espécie do gênero *Glycine*, a *G. tomentella* (SORIA-GUERRA *et al.*, 2010b), em que foram construídas quatro bibliotecas subtrativas de um genótipo suscetível (PI441101) e outro resistente (PI509501). Neste estudo foram encontrados, transcritos cujos produtos participam da síntese de ácido jasmônico, homônio este, importante na resposta contra patógenos também, como a aleno oxido sintase na planta suscetível e 4-coumarate-CoA ligase na planta resistente, o produto deste ultimo transcrito também está envolvido na rota dos fenilproanóides. Um estudo de microarranjos de DNA desenvolvido por van de Mortel *et al.* (2007) analisou a expressão diferencial de genes em dois genótipos de soja (PI230970 contendo *Rpp2*, resistente a FAS e Embrapa 48, suscetível) infectados por FAS utilizando o *GeneChip Soybean Genome Array*, que contém cerca de 37.500 grupos de sondas representando 35.611 transcritos de soja. Essa análise indicou uma expressão diferencial bifásica, com “up” regulação de genes, principalmente nos tempos iniciais e finais de inoculação, e “down” regulação nos tempos intermediários. Essa mesma plataforma para microarranjos foi utilizada também para avaliar a expressão de genes no genótipo de soja c.v. 5601T inoculado com FAS em um estágio inicial de desenvolvimento (V2). Esse Estudo demonstrou um número limitado de genes envolvidos na resposta à FAS em soja nos em estádios iniciais de desenvolvimento, talvez porque em condições naturais o patógeno infecte os estádios de V4 e R1, ou os cultivares norte americanos não são sabidos apresentarem genes de resistência á FAS. (PANTHEE *et al.*, 2007).

Panthee *et al.* (2009) analisou se o estágio de desenvolvimento em soja interfere na quantidade de genes envolvidos na resposta a FAS tanto suscetível (c.v. 5601T) como resistente (*G. tomentella* PI446998) através da técnica de microarranjos e RT-qPCR. A maior quantidade de genes diferencialmente expressos está presente no estágio V4 aonde 4.723 são exclusivos deste estágio, 333 são compartilhados entre os dois estádios e 246 são únicos de R1. Um grande número desses genes foi relacionado com defesa, estresse e metabolismo secundário, também genes envolvidos na rota do AS e de fenilpropanóides foram

encontrados. Os dados de RT-qPCR foram congruentes com os dados de microarranjos.

Outro estudo foi conduzido por Soria-Guerra *et al.* (2010b) analisando a expressão global de genes em resposta a FAS, utilizando somente microarranjos, nos mesmos genótipos de *G. tomentella* do trabalho anterior. Os horários analisados foram 12, 24, 48 e 72 hai. A validação dos genes foi feita por RT-qPCR. 1342 genes foram diferencialmente expressos em ambos os genótipos ao longo dos 4 tempos. Genes envolvidos na rota dos fenilpropanóides, como é o caso de chalcona sintase e PAL, foram encontrados, mostrando a importância desses compostos, também nesta espécie de soja, contra a ferrugem asiática da soja.

A técnica de microarranjo também tem sido utilizada em combinação com a técnica de Microdissecção e captura a laser para estudo de resposta de células da planta infectada com a ferrugem. Recentemente, essa técnica foi utilizada para identificar genes e rotas metabólicas envolvidas na resposta a infecção a ferrugem asiática (TREMBLAY *et al.*, 2010), bem como genes do fungo que possam estar envolvidos com a patogênese (TREMBLAY *et al.*, 2009), o que levou a identificação de 117 potenciais genes do fungo e 18 genes do hospedeiro, sendo as principais rotas ativadas durante a interação com o patógeno, as rotas dos fenilpropanóides.

## 2.5 AVANÇOS NAS TECNOLOGIAS DE SEQUENCIAMENTO

A técnica de sequenciamento de DNA mais utilizada nos últimos anos é baseada em inibidores de terminação de cadeia (didesoxinucleotídeos) e polimerização de DNA. Este método, descrito por Sanger, Nicklen e Colson (1977) apresenta a facilidade de não usar a clivagem por várias enzimas de restrição, empregada na técnica de degradação química desenvolvida por Maxam & Gilbert (1977). Outra vantagem que permitiu a ampla utilização do sequenciamento Sanger foi o seu aperfeiçoamento pela retirada dos radioisótopos utilizados para observar os fragmentos, por corantes fluorescentes e não radioativos desenvolvido por Prober *et al.* (1987), que permitiu a automação da técnica e visualização dos quatro didesoxinucleotídeos correspondentes aos quatro desoxinucleotídeos (dNTPs) presentes no DNA, em uma só amostra.

O ponto negativo da técnica é que a variação do comprimento de onda entre os fluoróforos é bem pequena, o que pode acarretar em erros de leitura das bases que podem trazer como consequências, por exemplo, deleções de bases e mudança do frame de leitura (no caso de transcritos) e, portanto interpretação errônea da proteína. Essas deleções também podem acarretar no não alinhamento e/ou agrupamento das sequências, ou alinhamento com baixo e-value. Há também a necessidade da formulação de iniciadores específicos para um conjunto de reação de sequenciamento que também restringe o emprego de muitos vetores de clonagem (PROBER *et al.*, 1987; VARSHNEY *et al.*, 2009).

A replicação da sequência através de clonagem em bactérias é outro fator limitante, já que demanda tempo e espaço para armazenamento de placas, além do fato de que algumas sequências são difíceis de serem clonadas como, por exemplo, sequências de fatores de transcrição que muitas vezes podem ser tóxicas para as bactérias. Além disso, como o sequenciamento é feito a partir de uma reação de PCR, tem-se a desvantagem da concentração de DNA atuar na eficiência da DNA polimerase que pode ficar sobrecarregada no início da reação devido à alta concentração de DNA e parar a polimerização sem estender a cadeia satisfatoriamente; se faz necessário também a separação dos fragmentos gerados, em gel de poliacrilamida nos capilares, um para cada amostra (PROBER *et al.*, 1987; VARSHNEY *et al.*, 2009). Contudo, a técnica de Sanger tem sido uma das principais ferramentas para a caracterização de ESTs devido ao baixo investimento em equipamentos, geração de sequências relativamente longas e por ser a única técnica disponível até o ano de 2005.

Entretanto, a tecnologia de Sanger ainda apresenta vantagens sobre as novas tecnologias de sequenciamento (next-generation sequencing - NGS), como o tamanho das sequências geradas que pode ultrapassar 1kilo base, o que facilita na montagem e análise pelas ferramentas de bioinformática (VARSHNEY *et al.*, 2009).

A necessidade de ressequenciar partes do genoma humano e também de outros genomas, além dos avanços na nanotecnologia, microfluídica e informática, propiciaram o surgimento das NGS que geram uma quantidade de dados muitas vezes maior do que as tecnologias tradicionais e num menor tempo, não utilizam clonagem em vetores para *E. coli*, pois partem do produto de PCR em fase sólida, ou PCR de emulsão, além de apresentar menor custo por reação do que

as tecnologias tradicionais. Desde que foram lançadas no mercado em 2005 as tecnologias de sequenciamento mais utilizadas são: Solexa (Illumina), Pirosequenciamento 454 (Roche) e Solid (Applied Biosystems) (ILLUMINA, 2010; ROCHE, 2010; APPLIED BIOSYSTEMS, 2011).

As novas tecnologias cada vez mais são aperfeiçoadas para aumentar o tamanho dos fragmentos gerados, assim como para aumentar o aproveitamento dos reagentes utilizados nas reações e diminuir ruídos que podem causar problemas de leitura. Atualmente os fragmentos de alguns sequenciadores da Illumina estão alcançando até 150pb (pares de base), os sequenciadores Solid da Applied geram fragmentos de 75pb enquanto que os maiores fragmentos ainda são gerados pelos sequenciadores da Roche que, em média, apresentam tamanho de 400 pb (ILLUMINA, 2011; ROCHE, 2011; APPLIED BIOSYSTEMS, 2011).

Muitos avanços e pesquisas estão sendo feitos para desenvolver tecnologias de sequenciamento cada vez mais eficientes e a um menor custo. Existem novas tecnologias de sequenciamento cuja grande característica é o sequenciamento de moléculas únicas, ou seja, sem a necessidade da amplificação clonal realizada através de PCR. Algumas dessas máquinas capazes de sequenciar moléculas únicas encontram-se disponíveis no mercado, como é o caso da “Heliscope True Single Molecule Sequencing” comercializada pela empresa Helicos (HELICOSBIO, 2010). Mas o grande anseio dos pesquisadores é pelo desenvolvimento do sequenciamento direto, sem que o DNA precise ser fragmentado e sem a necessidade de síntese de outra cadeia de DNA para a obtenção dos resultados. No Japão, um grupo de pesquisadores de Osaka está desenvolvendo uma forma de sequenciar fragmentos de DNA através de microscopia de força atômica e/ou microscopia de tunelamento, diferenciando o arranjo dos átomos nas bases nitrogenadas, das quais eles já conseguem distinguir guanina das demais (ZHOU *et al.*, 2010).

As novas tecnologias de sequenciamento estão sendo utilizadas, basicamente, para os mesmos fins que as tecnologias tradicionais, porém, principalmente trazendo maior rapidez nas análises. No âmbito do melhoramento vegetal, as NGS estão sendo amplamente utilizadas para análises de variação genômica e produção de marcadores moleculares para seleção assistida como a caracterização de SNPs (“Single Nucleotide Polymorphism”), ressequenciamento de espécies bem caracterizadas para encontrar polimorfismos, sequenciamento *de*

*novo*, bem como sequenciamento de ESTs para análises de transcriptomas (VARSHNEY *et al.*, 2009). Dentre as NGS, a tecnologia Solexa vem sendo bastante empregada para análises de transcriptoma e também de microRNA.

A tecnologia Solexa consiste primeiramente na clonagem das sequências que foram previamente ligadas a uma lâmina de vidro, onde estão presentes adaptadores para a amplificação clonal das amostras a serem sequenciadas. Nessa lâmina são acrescentados iniciadores e dNTPs convencionais (sem marcação) em fase sólida e uma alta densidade de amostras é amplificada gerando mais de mil cópias de cada molécula, num diâmetro de 1µm. Portanto, não exige clonagem em bactérias. Ainda, quatro dideoxynucleotídeos modificados (terminação reversível) são empregados de uma só vez na amostra, cada um possui uma fluorescência distinta, diferindo de outras técnicas que necessitam empregar os diferentes dideoxynucleotídeos de cada vez. Outro fator importante é que a tecnologia Solexa é capaz de determinar a sequência de homopolímeros mais precisamente do que alguns métodos (ILLUMINA, 2009).

### 3 OBJETIVO GERAL

Caracterizar o perfil de transcritos da soja em resposta à infecção por ferrugem asiática, visando disponibilizar informações referentes à funcionalidade de genes relacionados com esta interação, através da geração de etiquetas de sequências expressas (ESTs).

#### 3.1 OBJETIVOS ESPECÍFICOS

- Construir três bibliotecas subtrativas do genótipo PI561356, que apresenta resistência qualitativa à FAS, em diferentes fases do ciclo infeccioso, para obtenção de ESTs através de sequenciamento pelo método Solexa;
- Construir três bibliotecas subtrativas do genótipo BRS231, com resistência quantitativa à FAS, em diferentes fases do ciclo infeccioso, para obtenção de ESTs através de sequenciamento pelo método de Sanger;
- Sequenciar as bibliotecas obtidas, utilizando diferentes plataformas de sequenciamento, com a finalidade de disponibilizar sequências que complementem as ESTs já disponíveis nos bancos de dados;
- Identificar e analisar *in silico* as ESTs obtidas dos dois genótipos utilizados em resposta à FAS;
- Validar os resultados de expressão diferencial dos genes identificados nas bibliotecas.

## REFERÊNCIAS

ADIE, B. A. T. et al. ABA Is an Essential Signal for Plant Resistance to Pathogens Affecting JA Biosynthesis and the Activation of Defenses in *Arabidopsis*. **The Plant Cell**, v. 19: p. 665–1681, 2007.

APPLIED BIOSYSTEMS. Disponível em:  
<<http://www.appliedbiosystems.com/absite/us/en/home/applications-technologies/solid-next-generation-sequencing/next-generation-systems.html>>. Acesso em : 19 mar. 2011.

ARIAS, C. A. A. et al. Asian Rust in Brazil: Varietal Resistance. Facing the Challenge of Soybean Rust in South America. **JIRCAS Working Report**, n. 58, 2008.

BARROS, R. Doenças da cultura da soja. **Tecnologia e Produção - Soja**, 2007/2008. Disponível em <<http://www.fundacaoms.org.br/request.php?22>> Acesso 3 ago. 2009.

BONDE, M. R. et al. Evaluation of virulence of *Phakopsora pachyrhizi* and *P. meibomia* isolates. **Plant Disease**, v. 90, p. 708-716, 2006.

BROMFIELD, K. R.; HARTWIG, E. E. Resistance to soybean rust and mode of inheritance. **Crop Science**, v. 20, p. 254-255, 1980.

BROMFIELD, K. R. Soybean Rust. Monograph 11. **The American Phytopathological Society**. 1982.

CALDWELL, P; LAING, M. **Soybean rust – a new disease on the move**. 2002. Disponível em <[HTTP://www.saspp.org/achieved/tablesoybeanrust](http://www.saspp.org/achieved/tablesoybeanrust)>. Acesso em 17 ago. 2009.

CAMARGO, L. E. A. Análise genética da Resistência e da Patogenicidade. In: BERGAMIN FILHO, A; KIMATI, H.; AMORIN, L. (Ed.). **Manual de Fitopatologia: Princípios e Conceitos**. São Paulo. Agronômica CERES, ed. 3; v. 1, p. 470-492, 1995.

CAMARGO, L. E. A; BERGAMIN FILHO, A. Controle Genético. In: BERGAMIN FILHO, A; KIMATI, H.; AMORIN, L. (Ed.). **Manual de Fitopatologia: Princípios e Conceitos**. São Paulo. Agronômica CERES, ed. 3; v. 1, p. 729-758, 1995.

CAMARGO, P. O. **Estudo da Herança de Caracteres Quali-Quantitativos e Mapeamento Genético de Alelos Resistentes à Ferrugem Asiática da Soja Presentes nas Pi561356 e Pi594754**. Tese (Mestrado) - Universidade Estadual de Londrina, Londrina, Paraná, 2010.

CHOI, J. J. et al. Expression patterns in soybean resistant to *Phakopsora pachyrhizi* reveal the importance of peroxidases and lipoxygenases. **Functional & integrative genomics**, v. 8, p. 341-359, 2008.

CLONTECH. PCR -Select™ cDNA Subtraction Kit/ **Clontech**/ User manual. 24 set 2008.

CONAB. **Acompanhamento de safra brasileira: grãos 2010-2011, quarto levantamento, janeiro 2011** / Companhia Nacional de Abastecimento. – Brasília: Conab, 2011

CONSÓRSIO ANTIFERRUGEM. **Ciclo e Epidemiologia**. Disponível em: <[http://www.consorcioantiferrugem.net/portal/?page\\_id=67](http://www.consorcioantiferrugem.net/portal/?page_id=67)>. Acesso em 18/02/2011a. CONSÓRSIO Antiferrugem . **Ferrugem em números**. Disponível em <[http://www.consorcioantiferrugem.net/portal/?page\\_id=698](http://www.consorcioantiferrugem.net/portal/?page_id=698)>. Acesso em 5 fev. 2011.

CONSÓRSIO Antiferrugem . **Mapa de dispersão**. Disponível em <<http://www.consorcioantiferrugem.net/portal/>>. Acesso em 5 fev. 2011.

DANGL, J. L.; JONES, J. D. G. Plant pathogens and integrated defenses responses to infection. **Nature**, v. 411, p. 826-833, 2001.

DIATCHENKO, L. et al. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. **PNAS**, Washington, v. 93, n. 11, p. 6025-6030, June 1996.

EMBRAPA – **Empresa Brasileira de pesquisa agropecuária**. Disponível em <[http://www.cnpso.embrapa.br/index.php?op\\_page=112&cod\\_pai=33](http://www.cnpso.embrapa.br/index.php?op_page=112&cod_pai=33)>. Acesso em 15 ago. 2009.

EMBRAPA – **Empresa Brasileira de pesquisa agropecuária**. Disponível em <[http://www.cnpso.embrapa.br/alerta/ver\\_alerta.php?cod\\_pagina\\_sa=160](http://www.cnpso.embrapa.br/alerta/ver_alerta.php?cod_pagina_sa=160)>. Acesso em 15 ago. 2009b.

FALK, A. et al. EDS1, an essential component of R gene-mediated disease resistance in *Arabidopsis* has homology to eukaryotic lipases **PNAS** v. 96, p. 3292-3297, 1999.

FLOR, H. H. Current status of gene-a-gene concept. **Annual review of Phytopathology.**, v. 9, p. 275-296, 1971.

GARCIA, A. et al. Molecular mapping of soybean rust (*Phakopsora pachyrhizi*) resistance genes: discovery of a novel locus and alleles. **Theoretical and Applied Genetics**, v.117, p. 545-553, 2008.

GIULIETTI, A. et al. An Overview of Real-Time Quantitative PCR: Applications to Quantify Cytokine Gene Expression. **Methods**, v. 25, p. 386-401, 2001.

HAAS, B. J.; ZODY, M. C. Advancing RNA-Seq analysis. **Nature Biotechnology**. v. 28 n. 5, p. 421-423, 2010.

HARTWIG, E. E. Identification of a fourth major gene conferring resistance to soybean rust. **Crop Science**, v. 26, p. 1135-1136, 1986.

HELICOS Bioscience Corporation. Disponível em: <<http://www.helicosbio.com/Portals/0/Documents/Helicos%20tSMS%20Technology%20Primer.pdf>>. Acesso em 31 ago. 2011.

ILLUMINA. DNA Sequencing with Solexa® Technology. Disponível em: <<http://illumina.com/?sources=careers>>. Acesso em 5 jul. 2009.

ILLUMINA. Disponível em: <[http://www.illumina.com/systems/genome\\_analyzer\\_iix.ilmn](http://www.illumina.com/systems/genome_analyzer_iix.ilmn)>. Acesso em 31 ago. 2011.

ILLUMINA. Disponível em: <[http://www.illumina.com/systems/genome\\_analyzer\\_iix.ilmn#workflow\\_specs](http://www.illumina.com/systems/genome_analyzer_iix.ilmn#workflow_specs)>. Acesso em : 19 mar. 2011.

INDEX FUNGORUM. ***Phakopsora pachyrhizi* Syd. & P. Syd.** Disponível em: <<http://www.indexfungorum.org/names/NamesRecord.asp?RecordID=121037>>. Acesso em 21 fev. 2011.

INGHAM, J. L. Disease resistance in higher plants. The concept of pre-infectious and post-infectious resistance. **Journal of Phytopathology**, v. 78, p. 314-335, 1973.

JEONG, R. D. et al. Identification of *Glycine max* Genes Expressed in Response to Soybean mosaic virus Infection. **Plant Pathology**. J, n. 21, p. 47-54, 2005.

JI, W.; CAI, L. Suppression Subtractive Hybridization PCR. In: WEISSENSTEINER, T.; GRIFFIN, H. G.; GRIFFIN, A. (Ed.). **PCR Technology: Current Innovations**, 2. ed., p. 287-298, 2004.

JIRAGE, D. et al. *Arabidopsis thaliana* PAD4 encodes a lipase-like gene that is important for salicylic acid signaling. **PNAS**, v. 96, n. 23, p. 13583-13588, 1999.

KOCH, E.; EBRAHIM-NESBAT, F.; HOPPE, H. H. Light and electron microscopic studies on the development of soybean rust (*Phakopsora pachyrhizi* Syd.) in susceptible soybean leaves. **Phytopathology**, v. 106, n. 4, p. 302-320, 1983.

KOORNNEEF, A.; PIETERSE, C. M. J. Cross Talk in Defense Signaling. **Plant Physiology**, v. 146, p. 839–844, 2008.

KUO, W. P et al. Analysis of matched mRNA measurements from two different microarray technologies. **Bioinformatics**. v. 18, n. 3, p. 405-412, 2002.

MATSUMURA, H. et al. Gene expression analysis of plant host-pathogen interactions by SuperSAGE. **PNAS**, v. 100, p. 15718-15723, 2003.

MATSUMURA, H. et al. SuperSAGE. **Cellular Microbiology**, v. 7, p. 11–18, 2005.

MAXAM, A. M.; GILBERT, W. A new method for sequencing DNA. **PNAS** v. 74, n. 2, p. 560-564, 1977.

MEYERS, B. C. et al. Methods for Transcriptional Profiling in Plants. Be Fruitful and Replicate. **Plant Physiology**, v. 135, p. 637-652, 2004.

MILES, M. R.; FREDERICK, R. D.; HARTMAN, G. L. Evaluation of soybean germplasm for resistance to *Phakopsora pachyrhizi*. **Plant Health Progress**, 0104-01-RS, 2006.

MUELLER, C. C. & BUSTAMANTE, M. Análise da expansão da soja no Brasil. s/l, abr. 2002. Disponível em: <[www.worldbank.org/rfpp/news/debates/mueller.pdf](http://www.worldbank.org/rfpp/news/debates/mueller.pdf)>. Acesso em 7 jun. 2009.

NIELSEN, H. B.; WERNERSSON, R.; KNUDSEN, S. Design of oligonucleotides for microarrays and perspectives for design of multi-transcriptome arrays. **Nucleic Acids Research**, v. 31, p. 3491–3496, 2003.

PANDEY, A. K. et al. Functional Analysis of the Asian Soybean Rust Resistance Pathway Mediated by Rpp2. **Molecular Plant-Microbe Interactions**, v. 24, n. 2, p. 194–206, 2011.

PANSTRUGA, R. Establishing compatibility between plants and obligate biotrophic pathogens. **Current Opinion in Plant Biology**, v. 6, n. 4, p. 320-326, 2003.

PANTHEE, D. R. et al. Gene expression analysis in soybean in response to the causal agent of Asian soybean rust (*Phakopsora pachyrhizi* Sydow) in an early growth stage. **Functional & integrative genomics**, v. 7, p. 291-301, 2007.

PANTHEE, D. R. et al. Differential expression of genes in soybean in response to the causal agent of Asian soybean rust (*Phakopsora pachyrhizi* Sydow) is soybean growth stage-specific. **Theoretical and Applied Genetics**, v. 118, p. 359-370, 2009.

PASCHOLATI, S. F.; LEITE, B. Hospedeiro: mecanismos de resistência. In: Bergamin Filho, A; Kimati, H.; Amorin, L. (Ed.). Manual de Fitopatologia: Princípios e Conceitos. São Paulo Mecanismos bioquímicos de resistência a doenças. **Revisão Anual de Patologia de Plantas**, v. 2, p. 1-51, 1995.

PFÄFFL, M. W. A new mathematical model for relative quantification in real-time RT-PCR. **Nucleic Acids Research**, v. 29, p. 2002-2007, 2001.

POPP, M. P.; DILLON, C. R.; KEISLING, T. C. Economic and weather influences on soybean planting strategies on heavy soils. **Agricultural Systems**, v. 76, p. 969-984, 2003.

PROBER, J. M. et al. A System for Rapid DNA Sequencing with Fluorescent Chain-Terminating Dideoxynucleotides. **Science**, v. 238, p. 336-341, 1987.

RIBEIRO, A. S. et al. Genetic control of Asian rust in soybean. **Euphytica**, v. 157, p. 15-25, 2007.

ROCHE Applied Science. Disponível em: <<http://454.com/products-solutions/system-benefits.asp>>. Acesso em 31 ago. 2010.

ROCHE Applied Science. Disponível em: <<http://454.com/products-solutions/system-features.asp#titanium>>. Acesso em: 19 mar. 2011.

SANGER, F.; NICKLEN, S.; COULSON A. R. DNA sequencing with chain-terminating inhibitors. **Biochemistry**, v. 74, nº 12, p. 5463-5467, 1977.

SCHWEIGHOFER, A. et al. The PP2C-type phosphatase AP2C1, which negatively regulates MPK4 and MPK6, modulates innate immunity, jasmonic acid, and ethylene levels in Arabidopsis. **The Plant Cell**, v. 19, p. 2213–2224, 2007.

SIEBEN, A.; MACHADO, C. A. Histórico e contextualização sócio-econômica e ambiental da soja (*Glycine max*) no Brasil. **Geoambiente online**. Jataí-Go jul-dez 2007.

SILVA, D. C. G. **Mapeamento De Genes De Resistência Da Soja À Ferrugem Asiática E Análise Transcricional Na Interação Patógeno-Hospedeiro**. Tese (Doutorado) - Universidade Estadual Paulista “Julio de Mesquita Filho”, Jaboticabal, São Paulo, 2008.

SILVA, J. C.; MAIA, G. L.; MEYER M. C. Occurrence of asian soybean rust on irrigated fields in Maranhão. In: Abstracts of contributed papers and posters, **VII World Soybean Research Conference**. (Moscardi, F. et al. editors) Londrina: Embrapa Soybean, 2004.

SORIA-GUERRA, R. T. et al. Transcriptome analysis of resistant and susceptible genotypes of *Glycine tomentella* during *Phakopsora pachyrhizi* infection reveals novel rust resistance genes. **Theoretical and Applied Genetics**, v. 120, p. 1315-1333, 2010a.

SORIA-GUERRA, R. T. et al. Identifying differentially expressed genes in leaves of *Glycine tomentella* in the presence of the fungal pathogen *Phakopsora pachyrhizi*. **Planta**, v. 232, p. 1181-1189, 2010b.

TAIZ, L.; ZEIGER, E. **Fisiologia Vegetal**. Porto Alegre: Artmed, ed. 4, 2009.

TALLA, E.; TEKAIA, F.; BRINO, L.; DUJON, B. A novel design of whole-genome microarray probes for *Saccharomyces cerevisiae* which minimizes cross-hybridization. **BMC Genomics**, v. 4, n. 38, p. 1471-1475, 2003.

TORRES, M. A.; JONES, J. D. G.; DANGL, J. L. Reactive Oxygen Species Signaling in Response to Pathogens. **Plant Physiology**, v. 141, p. 373–378, 2006.

TREMBLAY, A. et al. Transcriptome analysis of a compatible response by *Glycine max* to *Phakopsora pachyrhizi* infection. **Plant Science**, v. 179, p. 183-193, 2010.

TREMBLAY, A.; LI, S.; SCHEFFLER, B. E.; MATTHEWS, B. F. Laser capture microdissection and expressed sequence tag analysis of uredinia formed by *Phakopsora pachyrhizi*, the causal agent of Asian soybean rust. **Physiological and Molecular Plant Pathology**, v. 73, p. 163-174, 2009.

van de MORTEL, M. V. D. et al. Distinct biphasic mRNA changes in response to Asian Soybean Rust Infection. **Molecular Plant-Microbe Interactions**, Saint Paul, v. 20, n. 8, p. 887-899, Aug. 2007.

van der HOORN, R. A. L.; KAMOUN, S. From Guard to Decoy: A New Model for Perception of Plant Pathogen Effectors. **The Plant Cell**, v. 20, p. 2009-2017, 2008.

VARSHNEY, R. K. et al. Next-generation sequencing technologies and their implications for crop genetics and breeding. **Trends in Biotechnology**, v. 27, n. 9, p. 522-530, 2009.

VELCULESCU, V. E. et al. Serial Analysis of Gene Expression. **Science**, v. 270, P. 484-487.

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: a revolutionary tool for transcriptomics. **Nature Reviews**, v. 10, p. 57-63, 2009.

WYNN, W. K. Appressorium formation over stomates by the bean rust fungus: Response to a surface contact stimulus. **Phytopathology**, v.66, p.136- 146, 1976.

YORINORI, J. T.; NUNES JÚNIOR, J.; LAZZAROTTO, J. J. **Ferrugem asiática da soja no Brasil: evolução, importância econômica e controle**. Londrina: Embrapa Soja. Documentos, n. 24, 36p, 2004.

YUEN, T. et al. Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. **Nucleic Acids Research**, v. 30, p. 47-48, 2002.

ZAMBOLIN, L. Manejo integrado da ferrugem asiática da soja. In: ZAMBOLIN, L. (Ed.). **Ferrugem asiática da soja**. Viçosa MG: Suprema, 139 p. 2006.

ZHOU, X. G. et al. The next-generation sequencing technology: A technology review and future perspective. **Science China. Life sciences**, v. 53, n. 1, p. 44–57, 2010.

## CAPÍTULO 2

### PERFIS TRANSCRICIONAIS DE GENÓTIPOS DE SOJA COM RESISTÊNCIA QUANTITATIVA E QUALITATIVA À FERRUGEM ASIÁTICA AO LONGO DO CICLO DA DOENÇA EM DIFERENTES FASES DA INFECÇÃO

#### ARTIGO

#### Perfis transcricionais de genótipos de soja com resistência quantitativa e qualitativa à ferrugem asiática ao longo do ciclo da doença em diferentes fases da infecção

Artigo a ser submetido à revista Genetics and Molecular Biology

#### RESUMO

A soja é uma das culturas de maior importância econômica para o Brasil, no entanto, sua produção tem sofrido perdas causadas por diferentes doenças no país, sendo que atualmente uma das mais devastadoras tem sido a Ferrugem Asiática da Soja (FAS) causada pelo fungo biotrófico *Phakopsora pachyrhizi*. Não existe ainda, comercialmente disponível, cultivares com resistência efetiva a esse fungo, mas sabe-se que existem genótipos capazes de resistir à doença, pois apresentam diferentes mecanismos de resistência; como é o caso de genótipos que carregam os genes *Rpp*, que conferem resistência qualitativa a FAS, bem como genótipos com resistência quantitativa. Até o momento, o perfil transcricional de genótipos resistentes em resposta a FAS, associando técnicas de hibridização subtrativa e sequenciamento Solexa ainda é inédito. Neste estudo, procurou-se caracterizar os genes envolvidos no mecanismo de resistência qualitativa e quantitativa ao fungo *P. pachyrhizi*, utilizando-se os genótipos PI561356 e BRS231, respectivamente, através do sequenciamento, utilizando diferentes plataformas, de três bibliotecas subtrativas para cada genótipo, agrupadas em “bulks”, contendo genes expressos ao longo do ciclo da doença: [“bulk” 1: 12 horas após a inoculação (hai), 24hai, 48hai; “bulk” 2: 72hai e 96hai e “bulk” 3: 192hai]. Entre os transcritos identificados nos diferentes “bulks” destacam-se aqueles envolvidos nos mecanismos de resposta aos diferentes tipos de estresses como os que codificam para proteínas cinases, genes que apresentam domínios de resistência “Leucine Rich Repeat” (LRR) envolvidos em apoptose e diversos domínios de fatores de transcrição. A partir da análise dos dados, verificou-se que muitos genes são ativados simultaneamente em ambos os genótipos após a infecção com o patógeno, no entanto genes exclusivos, provavelmente relacionados a mecanismos específicos de resposta nos diferentes genótipos também foram identificados. A análise da expressão de sete genes selecionados nas bibliotecas por RT-qPCR permitiu validar os resultados das bibliotecas nesses dois genótipos.

**Palavras-chave:** Sequenciamento de nova geração. *Phakopsora pachyrhizi*. *Glycine max*. Bibliotecas subtrativas.

## ABSTRACT

Among the major crops of economical interest for Brazil is soybean, which suffers substantial losses due to pathogen attack. Asian Soybean Rust (ASR), caused by the biotrophic fungus *Phakopsora pachyrhizi* is one of the most devastating diseases. There are no cultivars, commercially available, with effective genetic resistance against ASR yet, but some genotypes are able to resist the disease by activating several resistance mechanisms; this is the case of genotypes carrying *Rpp* genes, that confer qualitative resistance, and also genotypes presenting quantitative resistance. To date, the transcriptional profile of ASR resistant genotypes associating subtractive hybridization and Solexa sequencing wasn't done yet. In this study, we aimed to characterize genes involved in quantitative and qualitative resistance of soybean against *P. pachyrhizi*, using PI561356 and BRS231 genotypes, respectively. For each line, three cDNA subtractive libraries were sequencing, using different platforms, from samples grouped in bulks containing genes expressed along the disease cycle; bulks were defined according to the time after infection they were collected: Bulk 1 - 12hai, 24hai and 48hai; Bulk 2 - 72hai and 96 hai; Bulk 3 - 192hai). Among the detected transcripts deserve mention those involved in response to different stresses, like kinase proteins, genes related to apoptosis containing the LRR resistance domain and several transcription factors domains. Data analysis showed that many genes are activated simultaneously in both genotypes after pathogen infection. However, exclusive genes, possibly related to response mechanisms specific to each line, were also identified. The expression analysis by qRT-PCR of seven selected genes allowed the validation of the results from the subtractive libraries.

**Keywords:** Next generation sequencing. *Phakopsora pachyrhizi*. *Glycine max*. Subtractive libraries.

## 1 INTRODUÇÃO

O Brasil, segundo maior produtor mundial de grãos de soja [*Glycine max* (L.) Merrill], atingirá, na safra 2010/11, uma área plantada estimada em 24,08 milhões de ha, e uma produção estimada em 68,55 milhões de toneladas de grãos (CONAB, 2011). Economicamente, as exportações do complexo soja geraram US\$ 7,4 bilhões ao país em 2009/10, o que representa 10,2% do total nacional de exportações (CONAB, 2010). No entanto, a expansão da cultura ainda é limitada por perdas causadas por fatores bióticos e abióticos. Dentre esses, a ferrugem asiática da soja (FAS), causada pelo fungo biotrófico obrigatório *Phakopsora pachyrhizi*, é a maior ameaça ao cultivo de soja na América do Sul. Os danos causados na cultura da soja por esse fungo variam de 10 a 80% (Schneider *et al.*, 2005). Genótipos de soja com resistência a todos os isolados da população do patógeno ainda não estão

disponíveis, e os tratamentos com fungicidas constitui a principal medida de controle, o que onera o custo da produção.

Estudos com os genes de resistência qualitativa em soja vêm amplamente sendo realizados para delineamento de estratégias de melhoramento, visando o retardo na quebra da resistência causada por raças-específicas do patógeno. No entanto, muito pouco é conhecido sobre genótipos que apresentam resistência conferida por poligenes. Sabe-se que essa é uma herança mais difícil de ser quebrada (CAMARGO, 1995), porém os mecanismos e os genes envolvidos na resistência quantitativa ainda são pouco estudados. Sabe-se que muitos genes envolvidos na produção de metabólitos secundários, principalmente os fenilpropanóides, são ativados durante a interação incompatível e compatível de soja contra a ferrugem asiática, principalmente em genótipos com resistência qualitativa (van de MORTEL *et al.* 2007; PANTHEE *et al.* 2009; TREMBLAY *et al.*, 2010; SORIA-GUERRA *et al.*, 2010; PANDEY *et al.*, 2011). O estudo do transcriptoma de genótipos contendo resistência qualitativa e resistência quantitativa pode auxiliar na elucidação dos mecanismos de defesa contra essa doença.

O presente estudo buscou caracterizar, por meio de tecnologias de sequenciamento, o transcriptoma de folhas inoculadas com o fungo *P. pachyrhizi*, a partir de bibliotecas subtrativas supressivas de dois genótipos de soja, PI561356 [resistência qualitativa à FAS (CAMARGO, 2010)], e BRS231 [resistência quantitativa à FAS (RIBEIRO *et al.*, 2007)] para inferir sobre os possíveis mecanismos de defesa e rotas ativadas durante o processo de infecção em cada um dos genótipos estudados.

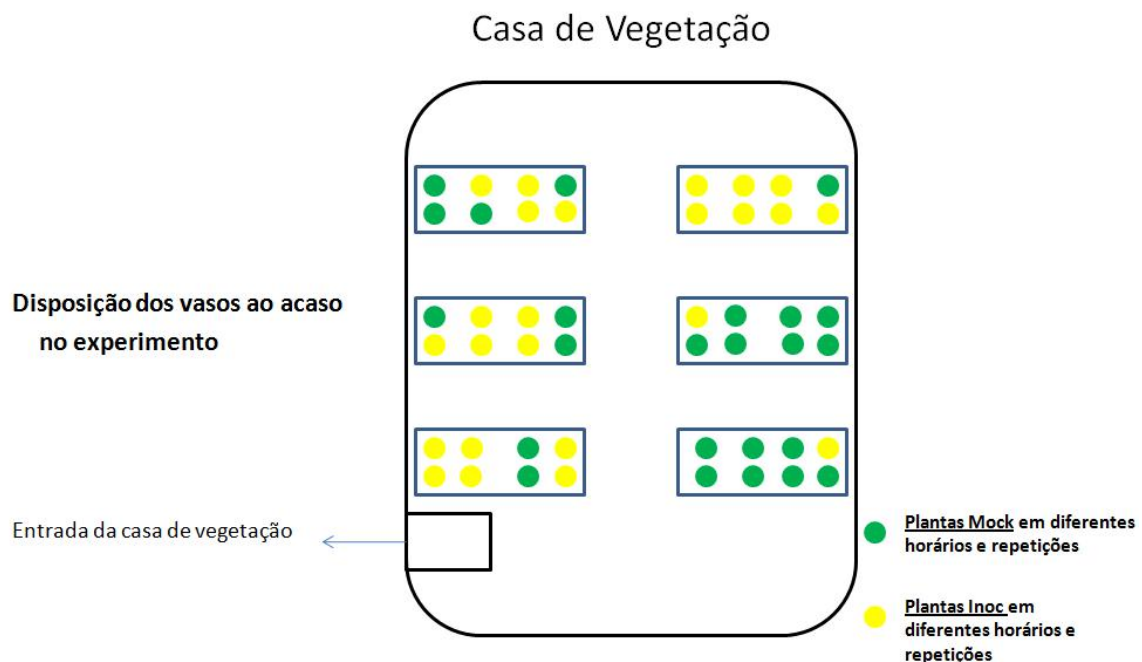
## **2 MATERIAL E MÉTODOS**

### **2.1 MATERIAL BIOLÓGICO E DESENHO EXPERIMENTAL**

Amostras de soja dos genótipos PI561356 (resistência qualitativa), BRS231 (resistência quantitativa) foram plantadas e mantidas em casa de vegetação, sob condições de temperatura ( $22\pm 1^{\circ}\text{C}$ ) e umidade ( $>60\%$ ) controladas, e fotoperíodo de 12h, até o estágio de desenvolvimento V2 (21 dias). O experimento seguiu delineamento inteiramente casualizado (DIC) com três repetições, em vasos separados, (cada repetição contendo duas replicatas no mesmo vaso) e estrutura de

tratamentos em fatorial com três fatores: tempo de coleta (12, 24, 48, 72, 96 e 192 horas após inoculação - hai), genótipos (PI561356 e BRS231) e tipo de infecção (inoculado com esporos de ferrugem asiática – “inoc”, e inoculado com água – “mock”). No total 72 vasos foram utilizados, contendo duas plantas, totalizando 144 plantas.

**Figura 1** - Representação do delineamento experimental em casa de vegetação.



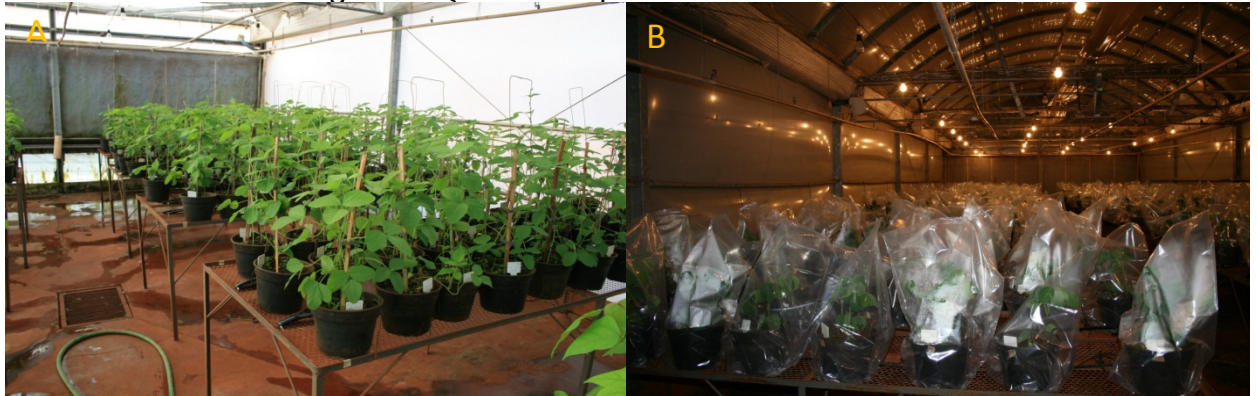
Um novo experimento para validar a hibridização subtrativa unidirecional via RT-qPCR utilizada na construção das bibliotecas foi posteriormente montado em casa de vegetação seguindo os mesmos parâmetros do experimento anterior.

## 2.2 PREPARO DOS INÓCULOS E INOCULAÇÃO

Primeiramente foram preparadas as amostras da condição mock: com a ajuda de um borrifador manual, foi pulverizada uma solução de água destilada e 0,05% (v/v) de Tween 20 (Uniqema) sobre as superfícies adaxiais e abaxiais das folhas. O procedimento foi repetido para a condição de inoculação, usando-se uma suspensão de uredósporos de *Phakopsora pachyrhizi* em água destilada contendo 0,05% de Tween 20, a uma concentração de  $6 \times 10^5$  uredósporos/mL. Os vasos de

ambos os tipos de inoculação foram cobertos com sacos plásticos, mantidos até o tempo de 12hai (Figura 2), para conservar a umidade e evitar a contaminação cruzada.

**Figura 2** - Experimento em casa de vegetação. A: plantas antes da inoculação distribuídas na casa de vegetação, B: plantas após a inoculação dos uredósporos, cobertas com sacos plásticos, para manter a umidade alta, facilitando a germinação dos esporos da FAS.

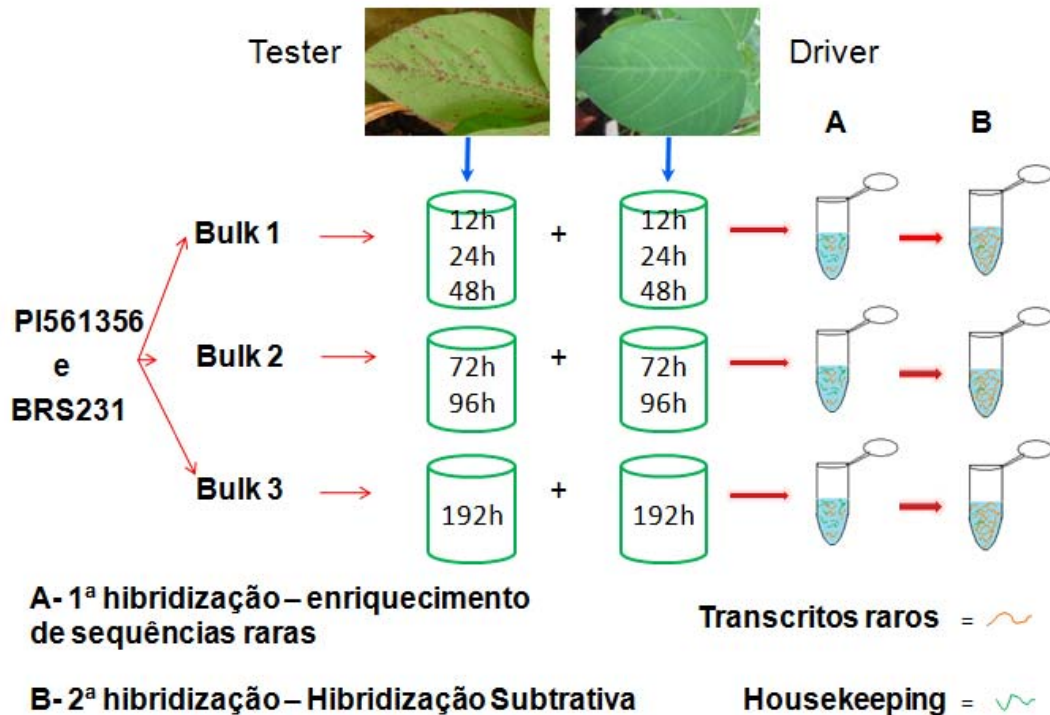


### 2.3 CONSTRUÇÃO DAS BIBLIOTECAS SUBTRATIVAS

O RNA total foi extraído de amostras de folhas usando o reagente TRIzol<sup>®</sup> (Invitrogen). O RNAm foi isolado com o kit FastTrack<sup>™</sup> MAG (Invitrogen). Após isolamento, o RNAm foi quantificado por espectrofotometria para avaliação da qualidade e quantidade de material purificado. A eficiência do isolamento de RNAm foi testada por meio de eletroforese em gel de agarose corado com 1% de Brometo de Etídio.

Para a construção das bibliotecas de hibridização subtrativa supressiva (SSH), parte das amostras de RNAm foram agrupadas em “bulks”, conforme o tempo de coleta e tipo de inoculação (Figura 3). Cada “bulk” foi composto de quantidades equivalentes a 2 µg de RNAm de cada amostra. As bibliotecas subtrativas foram construídas com o kit PCR-Select<sup>™</sup> cDNA Subtraction (Clontech). Seis bibliotecas SSH unidirecionais foram construídas e as amostras de folhas dos genótipos PI561356 e da BRS231 coletadas às 12, 24, 48, 72, 96 e 192 hai, foram utilizadas como “testers”; os respectivos tratamentos falso-inoculados (MOCK) foram utilizados como “drivers” (Figuras 3).

**Figura 3** - Esquema de construção das bibliotecas subtrativas para os genótipos PI561356 e BRS231.



Nas caixas verdes estão os horários após a inoculação da doença correspondentes aos “bulks” montados com os cDNA inoculado (“tester”) e falso inoculado (“driver”), em seguida o enriquecimento das seqüências que se dá ao longo das duas hibridizações por PCRs. O mesmo procedimento foi feito em separado, para BRS e PI resultando num total de 6 bibliotecas.

## 2.4 SEQUENCIAMENTO E BIOINFORMÁTICA

Os genótipos foram sequenciados por diferentes plataformas de sequenciamento, sendo a BRS231 sequenciada pela plataforma Sanger e a PI561356 sequenciada pela plataforma Solexa.

Os produtos de PCR das três bibliotecas subtrativas do genótipo PI561356 foram preparados para a clonagem “in vitro” e a subsequente reação de sequenciamento. As amostras foram sequenciadas com a tecnologia Solexa (Illumina) aplicadas em uma canaleta do sequenciador “Genome Analyzer” GAI utilizando “single reads” e gerando fragmentos de 76 pb, esse sequenciamento foi realizado através de prestação de serviço pela empresa FASTERIS AS, Genebra - Suíça.

A montagem do transcriptoma, as análises “in silico” e o banco de dados do GenoSoja foram realizadas em colaboração com o Laboratório de

Genômica e Expressão (LGE - <http://bioinfo03.ibi.unicamp.br/soja/>) da Unicamp. Os dados foram depositados no banco de dados do projeto GenoSoja/Embrapa Soja. A construção do banco de dados da soja, para servir como genoma de referência, foi realizado utilizando etiquetas de sequências expressas (“Expressed Sequence Tags” - ESTs) depositadas no banco de dados do “National Center of Biotechnology Information” (NCBI - <http://www.ncbi.nlm.nih.gov/>) e os dados do genoma da soja disponibilizados pelo DOE Joint Genome Institute, no site do Phytozome (<http://www.phytozome.net/soybean>).

As sequências produzidas foram ancoradas numa montagem prévia de ESTs, contendo as sequências disponíveis nos bancos de dados públicos, mais os Modelos de Genes (Gene models) do genoma referência Williams 82, com auxílio do programa SOAP – “Short Oligonucleotide Alignment Program” (Li *et al.*, 2008) usando parâmetros que permitiram o alinhamento com no máximo dois “mismatches” entre as sequências.

Os produtos de PCR das três bibliotecas subtrativas do genótipo BRS231 foram inseridos em vetor de clonagem PCR 2.1 TOPO T/A (Invitrogen), introduzidas por eletroporação em células eletrocompetentes da linhagem DH5 $\alpha$  de *Escherichia coli*, e cultivadas em meio LB contendo ampicilina (100 $\mu$ g/ml). Aproximadamente 3.500 clones individuais das três bibliotecas subtrativas foram tomados ao acaso, cultivados em meio líquido “Circle Grow”(CG) (BioSystems) para a minipreparação plasmidial. Para verificar o sucesso do isolamento dos plasmídeos e eliminar os clones contendo mais que um tipo de sequência, realizou-se a digestão dos mesmos com *EcoRI* (1 U/  $\mu$ g DNA) seguida de análise por eletroforese em gel de agarose a 1 % (m/v). O sequenciamento unidirecional (Sanger) foi realizado utilizando-se o iniciador M13, o kit “ABI Prism<sup>®</sup> BigDye<sup>™</sup> Terminator Cycle sequencing” v.2.0. (Applied Biosystems) e a separação dos fragmentos ocorreu em sequenciador automático ABI3700 (Applied Biosystems), através de prestação de serviço no laboratório de Bioquímica e Microbiologia de Plantas, Unesp/Jaboticabal.

Após a obtenção das sequências, estas foram agrupadas com o intuito de detectar as sequências oriundas do mesmo gene. A partir dos agrupamentos construiu-se as sequências consenso utilizando o programa CAP3. Os programas “Phred” e “Cross-Match” foram também utilizados para verificar a qualidade das sequências e remover as sequências do vetor de clonagem, respectivamente. Todas as sequências consenso e sequências únicas foram

comparadas com o banco de dados de proteínas (nr) do NCBI utilizando o BLASTx para identificação de homólogos.

Os dados de sequenciamento de ambos os genótipos foram anotados automaticamente no banco de dados do projeto GenoSoja, no qual esse trabalho está inserido, pela ferramenta AutoFACT (“Automatic Annotation and Classification Tool”; KOSKI *et al.*, 2005), que analisa dados de sequências de nucleotídeos e proteínas e determina a descrição funcional mais informativa, graças à combinação de relatórios de vários BLASTs provenientes de bancos de dados diversos. Posteriormente, foi feita uma nova anotação automática, por parte da equipe do Laboratório de Bioinformática da Embrapa Soja, utilizando a ferramenta Blast2GO.

O Software Blast2Go foi utilizado para gerar os gráficos para visualização das principais categorias funcionais bem como das rotas biológicas identificadas (GÖTZ *et al.*, 2008).

## 2.5 VALIDAÇÃO DOS GENES POR RT-QPCR

Foram selecionados sete transcritos que tem participação nos mecanismos de resistência à FAS presentes em ambas as bibliotecas para quantificação através de RT-qPCR a partir dos tempos individuais de coleta, a fim de verificar sua ativação em pelo menos um dos tempos de coleta e, conseqüentemente, validar a hibridização subtrativa utilizada para obtenção dos transcritos. Desses, quatro apresentam sequências com domínios relacionados a genes de resistência (três com domínios LRR: ACM89601.1, XP\_002532476.1, AAO15846.1; e um com domínio Serina/Treonina proteína quinase, Q2MHE4.1). Também foram selecionadas a proteína Mitogen Activated Protein Kinase Kinase - MAPKK (MPKK, ABK96640.1), a proteína com repetições de anquirina (ACU20172.1) e o fator de transcrição WRKY23 (Glyma15g11680).

Com o auxílio dos programas Primer Express 3.0 (Applied Biosystems) e Vector NTI Advance™ (Invitrogen) foram desenhados seis pares de “primers” nas sequências do genótipo BRS231, e um par de primer (para o gene WRKY23) foi desenhado no “gene model” da cultivar Wilians 82 depositada no banco de dados do Phytozome (<http://www.phytozome.net/>).

Para quantificar a expressão dos genes previamente selecionados, foi utilizado o método de RT-qPCR. Uma etapa prévia às reações para análise da expressão dos genes foi realizada, com o objetivo de escolher os melhores controles endógenos para a normalização das quantidades iniciais de cDNA. Cinco genes constitutivos foram testados:  $\beta$ -actina,  $\beta$ -tubulina, fator de alongação 1b (ELF-1b), gliceraldeído-3-fosfato desidrogenase (GAPDH) e RNAr 18S. As reações foram feitas em triplicatas, para cada uma das 3 repetições, e os dados foram analisados com a ajuda dos algoritmos NormFinder e geNorm. As reações foram realizadas utilizando-se o *kit Platinum<sup>®</sup> SYBR<sup>®</sup> Green qPCR SuperMix-UDG With Rox* (Invitrogen), e o equipamento 7300 Real Time PCR System (Applied Biosystems).

O RNA total (5  $\mu$ g), do segundo experimento montado para a validação dos genes das bibliotecas subtrativas, previamente tratado com DNase, foi utilizado para síntese de cDNA, segundo as especificações do fabricante do *kit M-MLV Reverse Transcriptase* (Invitrogen). O teste de eficiência foi realizado a partir de curvas de diluições seriadas do cDNA ( $10^{-1}$ ,  $10^{-2}$ ,  $10^{-3}$  e  $10^{-4}$ ) utilizando-se a fórmula  $E = 10^{(-1/\text{slope})} - 1 \times 100\%$  (PFAFFL, 2001). Em seguida, foi realizada a reação de RT-qPCR utilizando-se 1,0  $\mu$ L de cDNA como molde, 0,02  $\mu$ M de cada um dos oligonucleotídeos direto e reverso, 6,25  $\mu$ L de SYBR, e 4,25  $\mu$ L de água. As condições de amplificação atenderam aos seguintes parâmetros: 50°C por 2 min, 95°C por 10 min, 40 ciclos de 95°C por 15 seg e 62°C por 1 min. Ao final da reação, foi adicionada uma etapa de dissociação, na qual os produtos de amplificação de cada par de oligonucleotídeos iniciadores foram submetidos a temperaturas de 60 a 95°C. A especificidade dos produtos amplificados foi avaliada pela análise das curvas de dissociação geradas. As reações foram conduzidas utilizando os tempos individuais de coleta (12, 24, 48, 72, 96 e 192 hai).

Para a análise de expressão foi utilizado o método de quantificação relativa, utilizando as amostras controle “mock” em cada tempo correspondente como calibrador (LIVAK & SCHMITTGEN, 2001). Para a análise estatística da expressão dos genes alvos, foi utilizado o Relative Expression Software Tool (REST) versão 2.0.7 (PFAFFL *et al.*, 2002).

### 3 RESULTADOS E DISCUSSÃO

#### 3.1 ANÁLISE COMPARATIVA ENTRE OS SEQUENCIAMENTOS POR SOLEXA E SANGER

Nas bibliotecas subtrativas do genótipo PI561356, submetidas ao sequenciamento Solexa, foram gerados 14.886.500 sequencias de 76pb de comprimento, agrupadas em 3.277 contigs Solexa (Tabela 1), permitindo identificar, após o mapeamento contra EST e Modelos de Genes previamente identificados a partir de bancos de públicos, 1.370 modelos de genes (Glyma), 1736 “contigs” e 171 “singlets” (dados não mostrados). Em contrapartida, as bibliotecas do genótipo BRS231, sequenciadas pelo método de Sanger, geraram apenas 3.168 sequencias, que, apos as análises de agrupamento, permitiu a obtenção de 213 “contigs” e 1805 “singlets” (Tabela 1). A forma como as bibliotecas de ambos os genótipos foram analisadas podem ser observadas nas Figuras 4 para PI561356 e Figura 5 para BRS231.

**Tabela 1** -Quantidade de “reads”, “contigs” e “singlets” e tamanhos dos ESTs das bibliotecas subtrativas dos genótipos PI561356 e BRS231.

		Read	Contig	Singlet	Maior tamanho do contig ou sequência	Menor tamanho do contig ou sequência	Tamanho médio
PI561356	Bulk 1	5,185,015	1812	-	1348	100	334
	Bulk 2	5,000,616	748	-	879	100	286
	Bulk 3	4,700,869	717	-	1025	100	259
Total		14,886,500	3277				
BRS231	Bulk 1	1,056	63	788	1711	186	798
	Bulk 2	1,056	76	496	2024	235	791
	Bulk 3	1,056	74	521	1670	295	746
Total		3168	213	1805			

Diversos trabalhos demonstram a alta capacidade de sequenciamento por Solexa: bibliotecas de microRNAs em trigo sob estresse abiótico (calor) e biótico (míldio) geraram respectivamente 8.838.536 e 12.174.124 sequencias. Além disso, a análise do transcriptoma de raiz de soja infectada e não infectada por *Bradyrhizobium japonicum* (agente mutualista) geraram 4,88 e 6,84 milhões de sequencias (XIN *et al.*, 2010; LIBAUT *et al.*, 2010). De modo similar, neste trabalho ficou evidente o elevado poder do sequenciamento Solexa para gerar

grandes volumes de dados em comparação com o sequenciamento Sanger (Tabela 1). Porém como as sequências geradas são pequenas e “gaps” podem ser formados com certa constância, uma boa estratégia seria a união do sequenciamento Solexa aliado ao próprio sequenciamento Sanger ou pirosequenciamento 454, que são capazes de gerar sequências maiores, como demonstrado na tabela 1, onde as sequências geradas por Sanger são maiores que as da tecnologia Solexa.

No genótipo PI561356, 91,82% das sequências apresentaram similaridade com alguma proteína já descrita enquanto 8,18% das sequências não apresentaram similaridade com proteínas descritas, sendo classificadas como “no hits”. No genótipo BRS231, 79,97% das sequências apresentaram similaridade com alguma proteína já descrita e 24,03% das sequências não apresentaram similaridade com proteínas, sendo classificadas como “no hits” (Tabela 2).

Hanriot et al.,(2008) sequenciaram duas bibliotecas LongSage de hipotálamo de camundongos, uma através de Solexa e a outra através de Sanger para comparação dessas duas tecnologias de sequenciamento. 92,25 das tags Solexa ancoraram no genoma de referência enquanto que 90,28 das tags Sanger ancoraram. Os autores acreditam que as tags Solexa apresentam melhor qualidade do que as tags Sanger baseados também em outra evidência, pois ao permitir um mismatch, 2,29 das tags Solexa ancoraram no genoma, já as tags Sanger apresentaram mais mismatches e por isso 4,05 delas ancoraram também. Talvez o mesmo esteja acontecendo nas bibliotecas de PI561356 o que poderia explicar a menor quantidade de “no hits” e maior quantidade de informação que essas bibliotecas apresentaram, além do sequenciamento Solexa gerar um número muito maior de sequências que Sanger.

**Tabela 2** -Quantidade de sequências conhecidas e “no hits” nos genótipos PI561356 e BRS231.

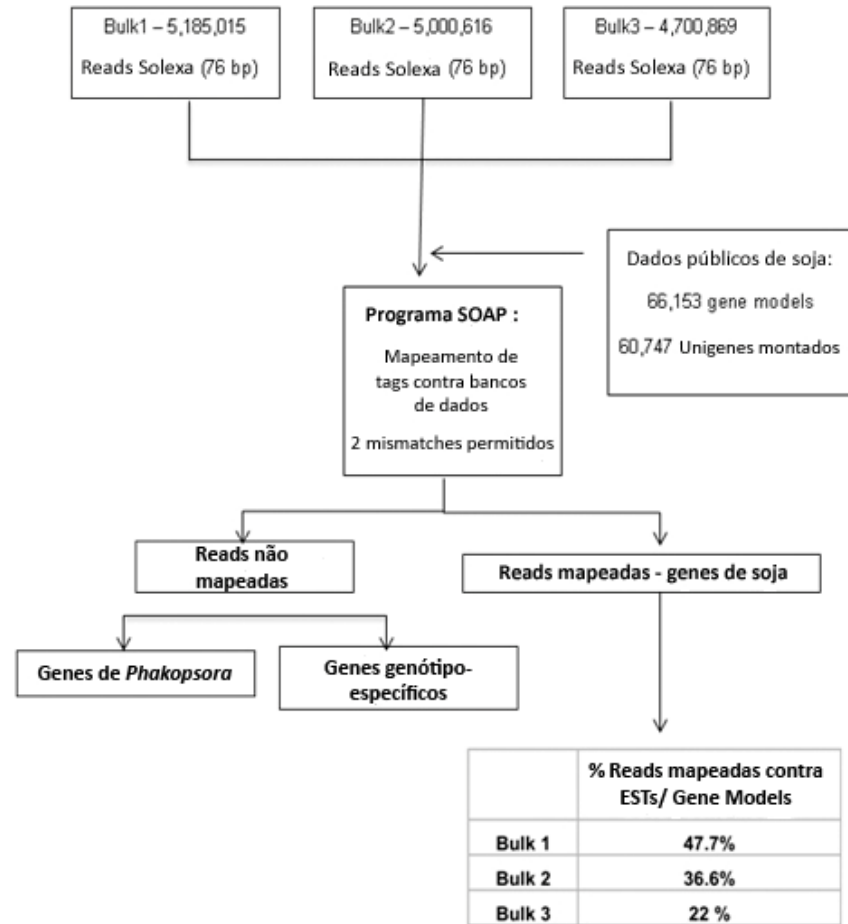
	PI561356				BRS231			
	Bulk 1	Bulk 2	Bulk 3	Total	Bulk 1	Bulk 2	Bulk 3	Total
Conhecidas	1675	695	659	<b>3029</b>	676	449	502	<b>1627</b>
<i>No hits</i>	137	53	58	<b>248</b>	175	123	93	<b>391</b>
% de <i>No hits</i>	8,18	7,62	8,8	<b>8,19</b>	25,89	27,39	18,52	<b>24,03</b>

Como esperado, tanto para o genótipo PI561356 como no BRS231 a maior parte da sequências foram anotadas com base nos Modelos de genes

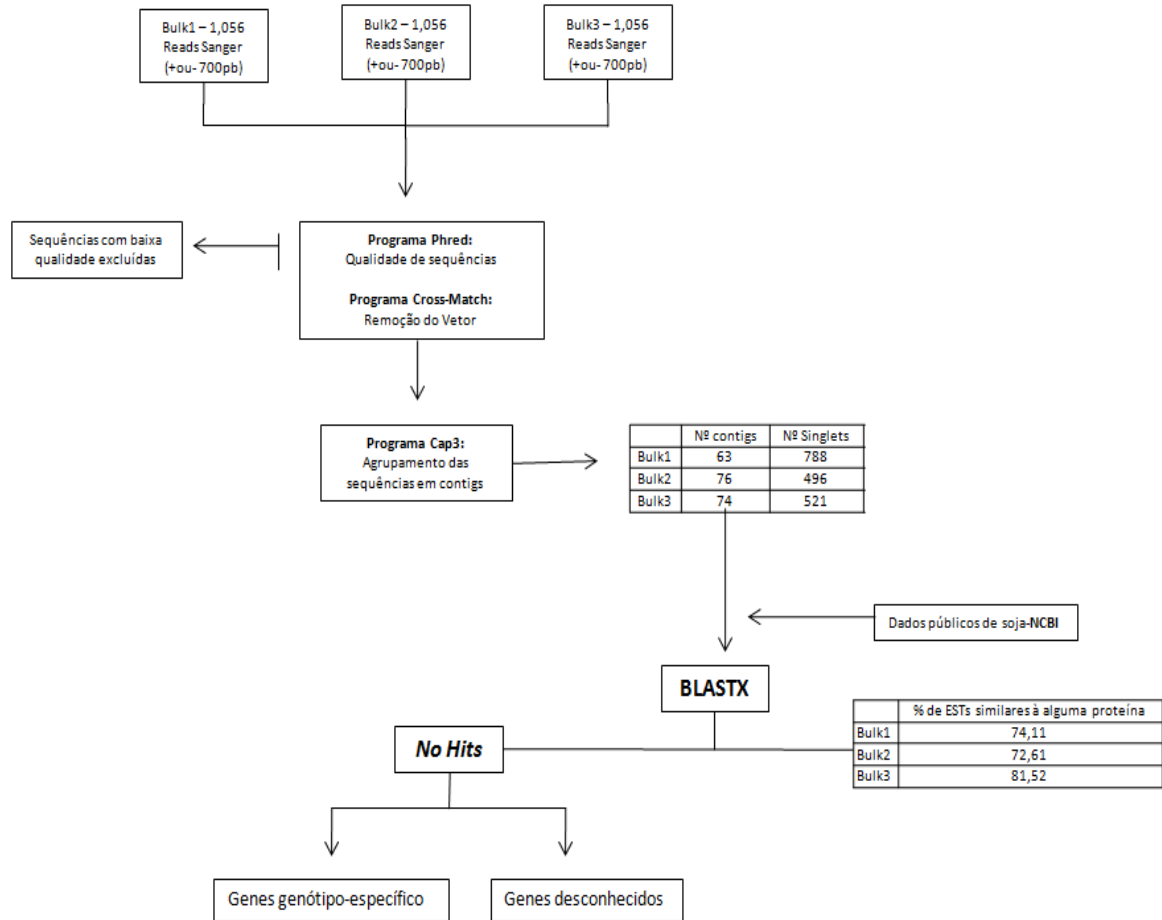
disponíveis a partir do sequenciamento do genoma de *Glycine max*, obtendo-se 1.629 e 512 ESTs respectivamente, o que já era esperado, pois essa é uma planta com quase todo o seu genoma anotado e depositado em bancos de dados públicos (Schmutz et al., 2010). Em seguida, no genótipo PI561356, o gênero com maior número de sequências similares foi *Vitis* com 398 ESTs, seguida por *Ricinus* com 290 ESTs e *Populus* com 228 ESTs. No genótipo BRS231 as sequências foram similares às dos gêneros *Populus* com 115 ESTs, *Vitis* com 139 ESTs e *Ricinus* com 123 ESTs. Os demais gêneros que apresentaram similaridade com PI561356 e BRS231 pertencem à mesma família da soja. Algumas das sequências de ambos os genótipos apresentaram similaridade com outros organismos, entre eles bactérias, vírus patogênicos de plantas, fungos (que podem ser genes de *P. pachyrhizi*, elicitores do fungo que ainda não foram estudados e podem trazer grandes contribuições para o estudo desta interação planta-patógeno), insetos, protozoários e mamíferos, porém, foi um número pequeno de sequências para PI561356 (26) e para BRS231 (123), o que pode se dever à grande capacidade de amplificação de sequências raras das SSH, uma vez que esses transcritos podem ser de contaminantes presentes nas folhas de soja.

Fica claro também na Figura 4 onde é mostrado o esquema das análises Solexa que muitas sequências não puderam ser mapeadas contra a montagem de ESTs e modelos de genes de soja, o que indica que muitas vezes, essa grande quantidade de sequências geradas pelas novas tecnologias de sequenciamento, podem não trazer a quantidade de informação equivalente a sua capacidade de geração de sequências devido a erros de montagem das pequenas “reads”, gerados por ferramentas de bioinformática, que, no entanto, vêm sendo constantemente aprimoradas para atender à demanda de dados gerados por essas novas tecnologias (HORNER *et al*, 2009). Por outro lado, essa grande quantidade de sequências não alinhadas em ESTs e modelos de genes de soja podem ser tanto genes desconhecidos de soja, bem como genes genótipo-específico, além de haver a possibilidade de serem genes do patógeno.

**Figura 4** - Esquema das análises realizadas nas sequências Solexa obtidas para PI561356.



**Figura 5** - Esquema das análises realizadas nas sequências Sanger obtidas para BRS231.

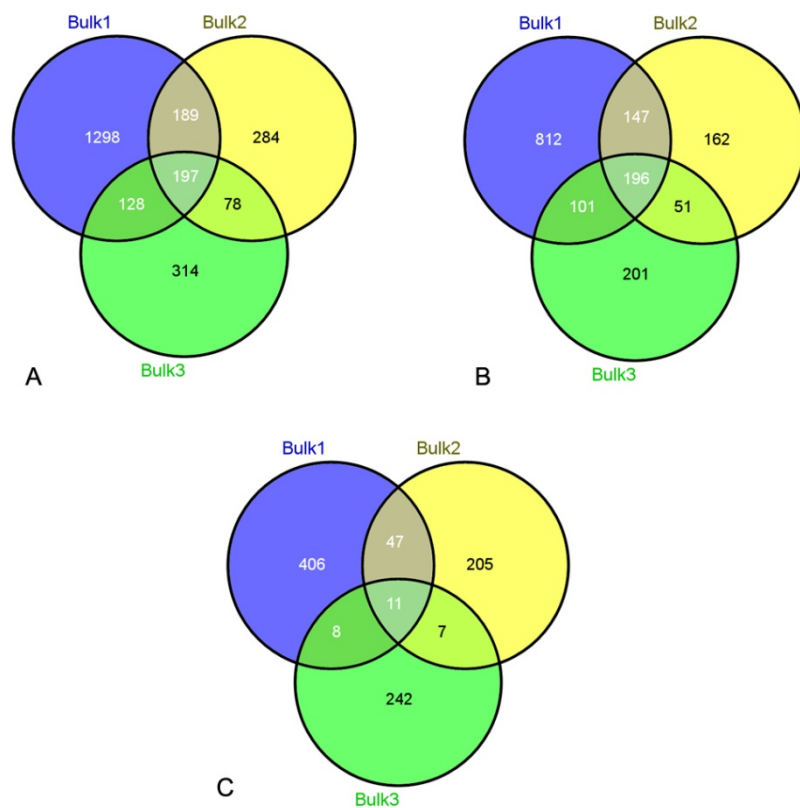


### 3.2 CATEGORIZAÇÃO FUNCIONAL

No total, foram anotadas 3.029 sequências consenso no genótipo PI561356, distribuídas nos três “bulks”. Quando essas sequências foram agrupadas de acordo com a anotação funcional obtida pela ferramenta AutoFact, a redundância em termos de anotação foi retirada, resultando assim em 1.670 unissequências. Vale ressaltar que genes diferentes podem ter mesma anotação seja por falta de informação em termos de anotação ou pelo fato de serem realmente genes relacionados funcional ou evolutivamente. A melhor estratégia para identificação dos unigenes seria a ancoragem das sequências obtidas contra o genoma da soja. Como no genótipo BRS231 as sequências foram obtidas por sequenciamento Sanger, os “singlets” correspondem às unissequências, que somam 926 distribuídas entre os “bulks”.

Das sequências geradas a partir das bibliotecas subtrativas do genótipo BRS231, 2.018 apresentaram similaridade com genes listados pela AutoFACT, mas apenas 916, que estão distribuídas entre os bulks, representam diferentes unissequências (Figura 6C).

**Figura 6** - A: Diagrama de Venn obtido através da comparação das sequências da PI561356 em cada um dos bulks obtidas através do alinhamento dos “contigs” do Solexa nos ESTs e Modelos de genes de soja (Glymas), B: unissequências da PI561356 após análise pelo AutoFact e C: unissequências da BRS231 após análise pelo AutoFACT.

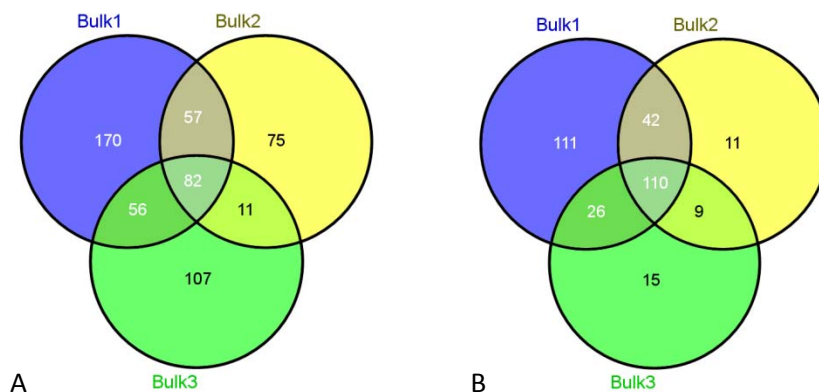


Diagramas obtidos a partir do programa Venny

Fonte: Oliveros, 2007.

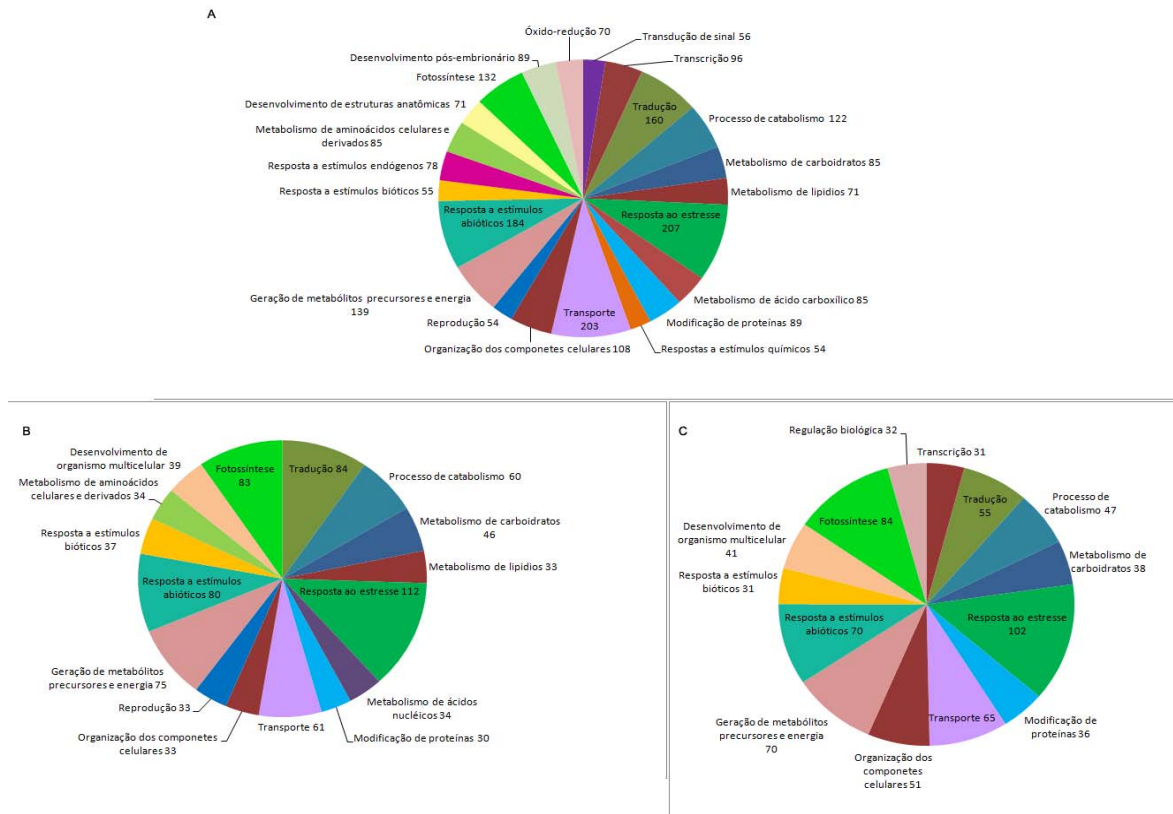
No genótipo PI561356 foram identificados o total de 326 e no genótipo BRS231 558 processos biológicos diferentes (com no mínimo um transcrito pertencente ao processo biológico), distribuídos entre os “bulks” (Figura 7).

**Figura 7** - Processos biológicos compartilhados e exclusivos entre os "Bulks" obtidos em ambos os genótipos. A: BRS231 e B: PI561356.

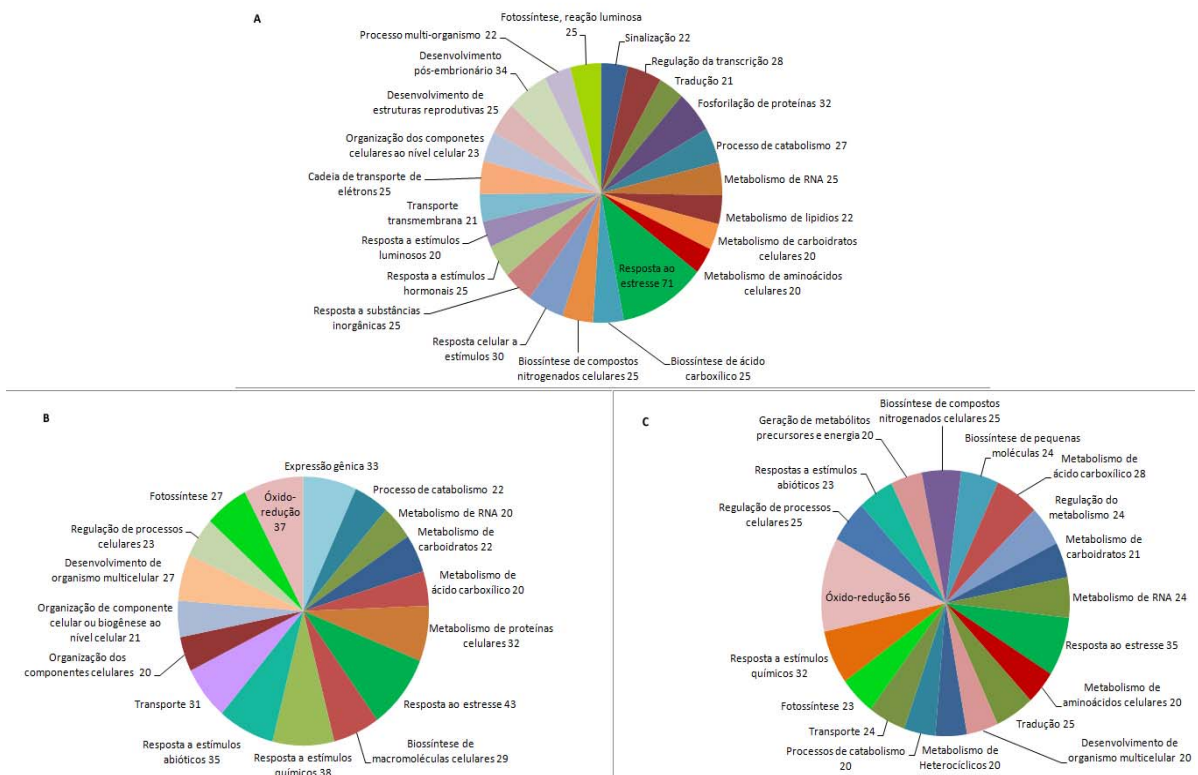


Após a caracterização funcional das ESTs, as categorias funcionais identificadas, relacionadas aos diferentes processos biológicos, foram agrupadas com auxílio do Programa Blast2GO. As sequências da PI561356 foram agrupadas em pelo menos 14 principais processos biológicos quando o número de 30 sequências foi utilizado como corte (cut-off). As categorias funcionais contendo o maior número de sequências identificadas em cada um dos três "bulks" foram aquelas relacionadas com transporte, transcrição, tradução, fotossíntese, resposta a estresse, processos catabólicos e geração de precursores de metabólitos e energia. (Figure 8). A mesma análise foi realizada com os contigs oriundos da BRS231, e foi possível identificar pelo menos as 20 categorias com um corte de 20 sequências, sendo que 14 delas são similares às identificadas na PI561356 como por exemplo: fotossíntese, geração de metabólitos precursores e energia, metabolismo de ácido carboxílico, metabolismo de lipídios, óxido-redução, resposta a estímulos abióticos, resposta ao estresse, tradução e transporte. No entanto, 28 categorias foram exclusivas da BRS231, dentre elas metabolismo de heterocíclicos, metabolismo de RNA e resposta a estímulos hormonais (Figura 9).

**Figura 8 - Gráficos da categoria Processos biológicos mais comuns encontrados no genótipo PI561356. Em A: Bulk1, em B: Bulk 2 e em C: Bulk3**



**Figura 9 - Gráficos da categoria Processos biológicos mais comuns encontrados no genótipo BRS231. Em A: Bulk1, em B: Bulk 2 e em C: Bulk3**



Na PI561356, as três categorias de processos biológicos que apresentaram maior número de sequências foram resposta ao estresse (GO:0006950) com 421 sequências, resposta a estímulos abióticos (GO:0009628) com 334, e transporte (GO:0006810). Já na BRS231, as três categorias com maior número de sequências foram resposta ao estresse (GO:0006950) com 149, óxido-redução (GO:0055114) com 93 sequências e resposta a estímulos químicos (GO:0042221) com 70 sequências.

Na categoria resposta ao estresse pode-se destacar processos biológicos mais especificamente relacionados a determinados tipos de estresse, reconhecimento de estímulos e resposta de defesa. Outras categorias que podem estar relacionadas a defesa do organismo também foram desmenbradas em categorias mais específicas (Tabela 3).

**Tabela 3 - Processos Biológicos relacionado a defesa e resposta a estress expressos em reposta a FAS nos genótipos PI561356 e BRS 231.**

Processo Biológico	Nº do Gene Ontology	Nº de ESTs em PI561356	Nº de ESTs em BRS231
<b>1. Defesa e Resposta a Estresse Bioticos e Abiótico</b>			
		<b>total</b>	<b>total</b>
Apoptose	GO:0006915	16	4
Biossíntese de alcalóides	GO:0009821	1	-
Biossíntese de flavonóides	GO:0009813	1	-
Biossíntese de lignina	GO:0009809	1	1
Biossíntese de terpenóides	GO:0016114	-	1
Biossíntese de 3-beta glucanase	GO:0006075	1	2
Catabolismo de peróxido de hidrogênio	GO:0042744	6	4
Deposição de calose na parede celular durante a resposta de defesa	GO:0052544	-	4
Metabolismo de fenilpropanóides	GO:0009698	1	3
Metabolismo de oxigênio e espécies reativas de oxigênio	GO:0006800	-	3
Metabolismo de superóxidos	GO:0006801	3	2
Biossíntese de ácidos gráxos	GO:0006633	21	7
Biossíntese de lípidios	GO:0008610	17	3
Biossíntese de oxilipídios	GO:0031408	2	4
Morte celular	GO:0008219	-	2
Remoção de radicais superóxidos	GO:0019430	-	1
Resistência sistêmica adquirida	GO:0009627	-	4
Resposta a bactéria	GO:0009617	4	1
Resposta a dessecação	GO:0009269	1	-
Resposta a estímulos bióticos	GO:0009607	6	3
Resposta a estresse oxidativo	GO:0006979	13	13
Resposta a estresse salino	GO:0009651	1	21
Resposta a fermentos	GO:0009611	1	9
Resposta a fungo	GO:0009620	-	2
Resposta a herbicidas	GO:0009635	-	12
Resposta a moléculas de origem bacteriana	GO:0002237	-	4
Resposta a nematóides	GO:0009624	1	5
Resposta a quitina	GO:0010200	-	6
Resposta ao calor	GO:0009408	-	6
Resposta ao frio	GO:0009409	3	29
Resposta de defesa	GO:0006952	20	9
Resposta de defesa contra bactéria	GO:0042742	1	17
Resposta de defesa contra fungo	GO:0050832	1	2
Resposta imune inata	GO:0045087	1	2
Rota de sinalização intracelular	GO:0023034	-	2
Transdução de sinais	GO:0007165	11	10
<b>2. Cinases</b>			
Fosforilação de proteínas	GO:0006468	70	29
<b>3. Regulação da expressão gênica</b>			
Transcrição	GO:0006350	33	2
Regulação da transcrição	GO:0045449	81	26
Regulação da transcrição , DNA-dependente	GO:0006355	81	24
Regulação negativa da transcrição	GO:0016481	-	3
Regulação positiva da transcrição	GO:0045941	-	3
Tradução	GO:0006412	251	29
Alongamento da tradução	GO:0006414	22	14
Iniciação da tradução	GO:0006413	16	12
Aminoacilação de tRNA para a tradução de proteínas	GO:0006418	19	1
<b>4. Modificação e catabolismo de proteínas</b>			
Modificação de proteínas pós-tradução	GO:0043687	15	2
Ciclo da ubiquitina	GO:0006512	14	-
Modificação de proteínas	GO:0006464	18	3
Catabolismo de proteínas	GO:0030163	13	2
Catabolismo de proteínas dependente de ubiquitina	GO:0006511	8	7

### 3.2.1 Categoria Funcional: Defesa e Resposta a Estresses Bióticos e Abióticos

As categorias funcionais agrupadas hierarquicamente sob as categorias de defesa e resposta a estresse bióticos e abióticos permitiram a identificação de pelo menos 36 processos biológicos diretamente envolvidos com a percepção e transdução de sinais, culminando com a ativação de genes relacionados a resposta de morte celular, biossíntese de metabólitos secundários, produção de hormônios, espécies reativas de oxigênio, entre outros (Tabela 3).

Apoptose e morte celular são respostas de defesas utilizadas contra patógenos para evitar que esses se espalhem para as células saudáveis da planta, e estão ligadas a rotas de sinalização que influenciam outros processos de defesa (GREENBERG e YAO, 2004).

Os processos biológicos catabolismo de peróxido de hidrogênio, remoção de radicais superóxidos, metabolismo de superóxidos, e metabolismo de oxigênio e espécies reativas de oxigênio estão todos envolvidos com outro processo biológico que é resposta ao estresse oxidativo. O estresse oxidativo, desencadeado pelas espécies reativas de oxigênio pode levar tanto a mecanismos de sinalização na planta contra estresses bióticos ou abióticos, bem como desencadear a morte celular (MITTLER, 2002).

A resposta de resistência do genótipo PI561356 (fenótipo com lesões RB) pode ser caracterizada pela ativação da resposta de hipersensibilidade (HR). Este é um tipo comum de resposta quando genes de resistência (genes R) são desafiados por moléculas codificadas pelos genes de avirulência do patógeno (genes *avr*). Vários processos biológicos, representados significativamente nas bibliotecas desse genótipo (Tabela 3) - apoptose, fosforilação de proteínas, transdução de sinais, resposta de defesa e estresse oxidativo – estão envolvidos nas cascatas de sinalização que culmina com HR e corroboram com a presença de pelo menos um gene R, provavelmente aquele estudado por Camargo (2010), que foi mapeado no cromossomo 18 da soja, próximo ao alelo *Rpp1*.

Uma forma de evitar a entrada do fungo da ferrugem é o espessamento da parede celular nas folhas da soja a partir de glicoproteínas ricas em hidroxiprolina (GO:0008150, processos biológicos), que conferem à parede celular maior resistência contra enzimas que a degradam e são provenientes de patógenos (HÜCKELHOVEN *et al.*, 2007). Essas glicoproteínas também foram

encontradas entre as bibliotecas de ambos os genótipos de soja estudados no “bulk” 1, o que pode ser um indicativo de que o espessamento da parede celular, como forma de deter a entrada do patógeno, está ocorrendo nessas plantas. A infecção de milho pelo agente causador do míldio, *Sclerospora graminicola*, mostrou que na cultivar resistente ao patógeno, a parede celular foi reforçada por uma combinação de polipeptídeos como a isoditrosina, bem como grandes quantidades de hidroxiprolina nas glicoproteínas e grandes quantidades de H<sub>2</sub>O<sub>2</sub> em contraste com as plantas suscetíveis (DEEPAK *et al.*, 2007). Para a confirmação de que está ocorrendo à formação deste tipo de barreira nos dois genótipos, estudos complementares por microscopia eletrônica e/ ou óptica trariam grandes contribuições para a compreensão destes mecanismos de resistência.

Muitos dos genes expressando nas bibliotecas do genótipo PI561356 também foram encontrados nas bibliotecas de BRS231, o que supõe que ambas as respostas de resistência podem estar sendo mediadas por genes em comum, como foi também observado na interação do tomateiro com o fungo *Oidium neolycopersici*, em que foram avaliados um genótipo suscetível e dois resistentes (qualitativa e quantitativa) (LI, BAI e JACOBSEN, 2006). Nesse estudo foram encontrados genes que estão envolvidos na sinalização por ácido salicílico, como *fenilalanina amonia-liase* (PAL) (GO:0009698, tabela 3), genes relacionados à resposta de hipersensibilidade, como *glutathione-S-transferase* (GO:0008152, processos de metabolismo) e genes relacionados à patogênese (“Patogen Related” - PR) como os que codificam para as proteínas PR-1 (GO:0006952, GO:0009607, tabela 3). Outro ponto importante foi o sincronismo com que esses genes são ativados, que pode culminar na diferença entre interações compatíveis e incompatíveis dos tipos poligênica e monogênica. Além disso, o ácido salicílico mostrou ser o principal hormônio sinalizador nesta interação planta-patógeno para ambos os tipos de resistência (LI, BAI e JACOBSEN, 2006), pela ativação de muitos genes envolvidos na rota do AS. Portanto, esses dados demonstram que a intensidade e o período em que um determinado gene é expresso em um genótipo pode ser a causa das diferentes respostas entre suscetibilidade, resistências qualitativa e quantitativa na planta durante a interação entre *P. pachyrhizi* e os genótipos de soja utilizados.

As proteínas PAL, 4-coumarate-CoA Ligase (GO:0008152, processos de metabolismo) e Chalcona sintase (GO:0008152, processos de

metabolismo) estão envolvidas diretamente na rota dos fenilpropanóides e seus derivados, as fitoalexinas, que são compostos tóxicos aos patógenos (DIXON e PAIVA, 1995; HAMBERGER e HALBROCK, 2004; TAIZ e ZAIGER, 2009). Tais proteínas foram encontradas nas bibliotecas de ambos os genótipos, com exceção da chalcone sintase, que só foi encontrada nos Bulk 1 e 3 da PI561356. Pandey *et al.* (2011) publicaram um trabalho de silenciamento de vários genes (pela técnica de VIGS) envolvidos no mecanismo de resistência a FAS, mediado por *Rpp2* de soja, e ao silenciar fatores de transcrição *GmWRKY36* e *GmMYB84*, demonstrou que os transcritos para as proteínas PAL, 4-coumarate-CoA Ligase e Chalcona sintase são reprimidos devido ao silenciamento desses fatores de transcrição que provavelmente os controlam, demonstrando a grande importância da ativação da rota dos fenilpropanóides na resposta a FAS. Essa rota provavelmente está sendo ativada nos genótipos em estudo devido ao aparecimento desses transcritos entre as bibliotecas, porém ao que tudo indica, com o aparecimento de mais transcritos envolvidos nessa rota, e em diferentes Bulks na PI561356 do que na BRS231, essa pode ser uma das diferenças entre resistência qualitativa e quantitativa, ou uma diferença gerada pelos métodos de sequenciamento, já que PI561356 foi sequenciada por Solexa que gera maior quantidade de sequências e pode obter maior variedade de transcritos. No entanto maiores estudos para elucidar esses mecanismos de resistência são necessários.

Transcritos para proteínas envolvidas no controle do estresse oxidativo das células também foram encontradas nas bibliotecas dos dois genótipo, destacando-se a Superóxido dismutase (GO:0019430, GO:0006801, GO:0055114, tabela 3) e Catalase (GO:0008219, GO:0042744, GO:0055114, GO:0006979, tabela 3). Superóxido dismutase é uma enzima importante para manter o equilíbrio redox da célula ao dismutar os radicais  $O_2^-$ , que são altamente oxidantes de compostos celulares como ácidos nucléicos e podem destruir as células, para  $H_2O_2$  (KLIEBENSTEIN, MONDE e LAST, 1998; MITTLER, 2002), que por sua vez pode também ser tóxico para as células quando em altas concentrações. A catalase é capaz de degradar o  $H_2O_2$  em  $H_2O + O_2$  (MHAMDI, *et al.*, 2010). Tais proteínas podem estar protegendo as estruturas celulares do estresse oxidativo.

Lipoxigenase (GO:0006633, GO:0008610, GO:0031408, tabela 3), também encontrada nos três Bulks da PI561356 e no Bulk1 da BRS231, é uma enzima responsável por degradar lipídeos envolvidos na produção de ácido

jasmônico e envelhecimento das membranas, sendo que um dos resíduos formados é o radical  $O_2^-$  (FEUSSNER e WASTERNAK, 2002). Um grande número de lipoxigenases e peroxidases foram também identificadas em um estudo de expressão gênica em plantas de soja (PI200492) infectadas pela ferrugem asiática (CHOI *et al.*, 2008), demonstrando o quanto as enzimas oxidativas estão envolvidas no mecanismo de defesa contra esse patógeno.

Proteínas com repetições de anquirina também foram encontradas entre as bibliotecas de ambos os genótipos. Essas repetições estão presentes no gene que codifica a proteína NPR1 de *Arabidopsis* (CAO *et al.*, 1997), que é um gene chave na resposta de resistência induzida pelo hormônio ácido salicílico (AS) e na supressão de genes responsivos ao ácido jasmônico (SPOEL *et al.*, 2003). Sabe-se que AS é um dos mais importantes hormônios na resposta a patógenos biotróficos como é o caso da FAS (SCHWEIGHOFER *et al.*, 2007). Portanto, a presença de genes que codificam para a proteína anquirina, que pode estar envolvida na supressão de hormônios como ácido jasmônico e etileno durante a infecção por *P. pachyrhizi*, pode favorecer a resposta de defesa da planta a esse patógeno.

A identificação de proteínas com atividade fosfatase (GO:0006470, desfosforilação de aminoácidos de proteínas) entre as bibliotecas dos genótipos BRS231 (“bulk” 1) e PI561356 (todos os “bulks”), colaboram ou propiciam a ocorrência de sinalização por AS na infecção por FAS, pois proteínas serina/treonina fosfatases 2C podem inativar algumas Mitogen Activated Protein Kinase (MAPK) responsivas ao estresse, como MPK4 e MPK6, envolvidas na resposta contra ferimentos. As fosfatases também causam a diminuição da produção de etileno e ácido jasmônico, que são importantes na indução de resposta a ferimentos e patógenos necrotróficos, mas são antagonistas ao AS (SCHWEIGHOFER *et al.*, 2007). É possível que fosfatases causem inibição de proteínas cinases envolvidas na regulação de ácido absísico (ABA), suprimindo respostas de sinalização pelo AS (UMEZAWA *et al.*, 2009), já que o ABA também é um antagonista de AS, por regular a enzima isocorismato sintase que sintetiza um de seus precursores (ZABALA *et al.*, 2009).

Lipases (GO:0006629, metabolismo de lipídios), associadas ao etileno, podem estar envolvidas na sinalização de resistência sistêmica em *A. thaliana*, como demonstrado por Kwon *et al.* (2009) ao promover o acúmulo de

moléculas sinalizadoras e permitir que estas realizem a sinalização célula-a-célula. Uma lipase, com função antimicrobiana, induzida por AS, também foi induzida por etileno, e promoveu resistência ao patógeno necrotrófico *Alternaria brassicicola* em *A. thaliana* e também demonstrou estar relacionada à sinalização de resistência sistêmica (OH *et al.*, 2005). De acordo com esses dados é possível sugerir que os transcritos para lipases encontrados nas bibliotecas poderão codificar essas proteínas que possivelmente estão auxiliando na transmissão de informações entre as células dos genótipos analisados, podendo também estar associadas ao AS.

Genes *atp binding cassette (ABC) transporter* (GO:0009624, GO:0009611, GO:0009627, GO:0052544, GO:0042742, tabela 3), foram identificados nas bibliotecas de ambos os genótipos. Esses genes estão envolvidos no controle da morte celular em resposta de defesa e podem transportar algumas substâncias como lipídios, ácidos inorgânicos, metabolitos secundários, toxinas, drogas e peptídeos que estão relacionados com a resposta de defesa inicial (KOBAE *et al.*, 2006). Esse gene também está envolvido na resistência a metais pesados em *A. thaliana* (KIM *et al.*, 2007). Um tipo de *abc transporter* em soja, o *GmPDR12*, responde rapidamente ao ácido salicílico (EICHHORN *et al.*, 2006), o que pode explicar o aparecimento desses genes nas bibliotecas, já que lesões causadas pela hipersensibilidade são facilmente observadas nas plantas do genótipo PI561356 (resistência qualitativa). No entanto tais lesões não são vistas na BRS231 (resistência quantitativa), sugerindo que algum outro mecanismo, ou a interação com outra proteína, ou até mesmo o nível de expressão seja um diferencial entre esses dois genótipos levando a diferentes respostas.

Quanto às proteínas que apresentam domínios de resistência LRR (GO:0006915, GO:0006952, GO:0007165, GO:0045087, tabela 3) e que também foram encontradas em ambos os genótipos, existem proteínas integrais de membrana com o domínio LRR extracelular que interage primeiramente com moléculas elicitoras, induzindo a resposta imune basal da planta que, a princípio, evita a entrada do patógeno na célula. Porém, se o patógeno “burla” essa resposta primária e penetra na célula, proteínas intramembrana com domínios NBS-LRR, que também reconhecem a molécula elicitora, uma vez ativadas, induzem a resposta de hipersensibilidade que leva à morte celular programada. As NBS-LRR podem conter domínios N-terminais variáveis como, por exemplo, CC (“coiled coil”) ou TIR (“*Drosophila Toll*” e receptor de Interleucina-1 humana), sendo assim, as NBS-LRR

funcionam como o sistema imune adaptativo de mamíferos, pois cada NBS-LRR reconhece moléculas de patógenos específicos (PADMANABHAN *et al.*, 2009). A morte celular programada, resultante da atividade de proteínas com este domínio, é uma característica evidente nas plantas do genótipo PI561356, que apresentam resistência do tipo qualitativa (genes R), que culmina na morte das células e aparecimento de lesões do tipo RB. No entanto, genes que contêm esses domínios também foram identificados na BRS231, a qual não apresenta a típica reação de hipersensibilidade e a consequente morte celular na resposta ao fungo da ferrugem.

### 3.2.2 Categoria Funcional: Cinases

Na PI561356, aproximadamente 99 ESTs apresentaram similaridade com proteínas cinases, enquanto que na BRS231 foram identificadas aproximadamente 50 ESTs. A grande maioria dessas sequências tanto em PI561356 quanto em BRS231 correspondem a transcritos para proteínas cinases (Figura 10) sem maiores especificações de suas funções. Apesar de já existirem vários estudos sobre proteínas cinases, a real função e a forma de interação com outras proteínas ainda é desconhecida para a grande maioria (DING *et al.*, 2009). Portanto, estudos mais aprofundados, como os realizados por Ding *et al.* (2009), para caracterização das interações de proteínas cinases em arroz através do sistema duplo híbrido, podem ser realizados também para caracterizar a função dessas proteínas em soja, visto que são proteínas chaves na transdução de sinais para a ativação de defesa durante ataques de patógenos.

Como pode ser observado na Figura 10 os transcritos para a classe de proteínas cinases denominadas MAPK apareceram nas bibliotecas de ambos os genótipos. Em mamíferos, as MAPK respondem principalmente pela transdução de sinais mitogênicos (daí provém o seu nome), enquanto que em plantas essas proteínas estão envolvidas em uma ampla variedade de estímulos como respostas hormonais, regulação do ciclo celular, resposta a estresse abiótico e biótico (TENA *et al.*, 2001). Ntf4, uma MAPK de tabaco, até então relacionada somente à regulação da germinação de pólen, também está envolvida na resposta ao patógeno *Phytophthora cryptogea*, pois em plantas transgênicas que superexpressam essa proteína, e foram tratadas com um elicitor do patógeno, ocorre o aceleração da morte celular (REN *et al.*, 2006).

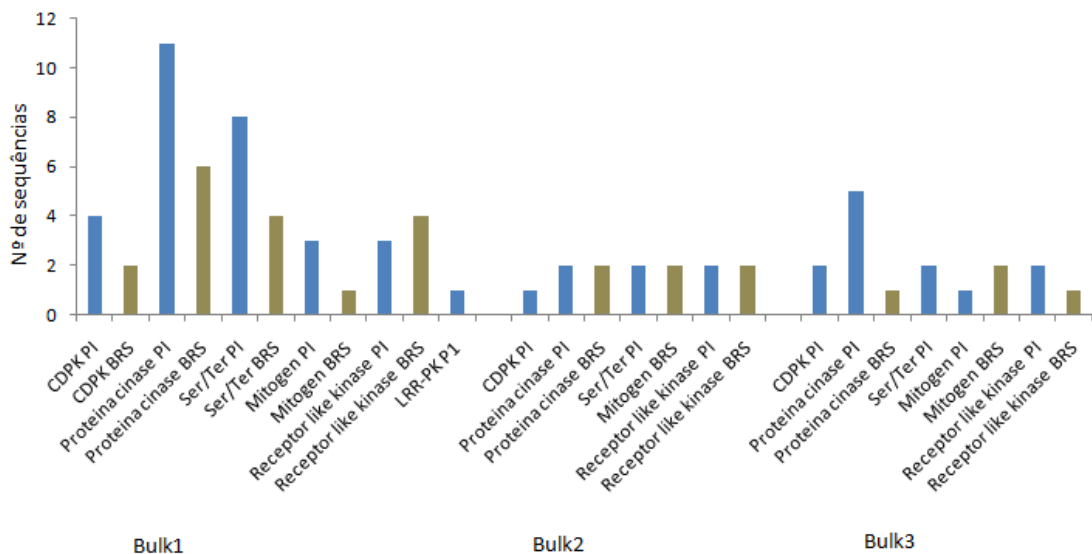
Através da análise de protoplastos de *A. thaliana*, tratados com o elicitor de bactéria Flg22, Asai *et al.* 2002 identificaram uma possível rota de transdução de sinais na defesa contra estresse biótico. Essa rota seria constituída pela cascata de sinalização entre MEKK1, MKK4/MKK5, MPK3/MPK6, que por sua vez ativariam fatores de transcrição envolvidos no controle da expressão de proteínas de defesa como o WRKY29. Isso demonstra a importância dessas proteínas na defesa contra patógenos.

CDPK (proteínas cinases dependentes de cálcio) são predominantemente de plantas e apresentam o mesmo domínio para ligação ao cálcio (EF-“hand”) que as proteínas cinases dependentes de cálcio/calmodulina encontradas em animais (STONE e WALKER, 1995; ROMEIS *et al.*, 2001). Uma NADPH oxidase de batata apresenta domínio EF-“hand”, indicando que esta pode ser regulada por cálcio. Essa enzima é importante para a defesa da planta, pois participa da produção de espécies reativas de O<sub>2</sub>. Estudos de espectrometria de massa revelaram que a proteína St CDPK5 é capaz de fosforilar as serinas 82 e 97 da região N terminal dessa NADPH e portanto regular a queima oxidativa durante interações planta-patógeno (KOBAYASHI *et al.*, 2007). As CDPK foram detectadas em todas as bibliotecas da PI561356, enquanto em BRS231 só aparecem na biblioteca 1 (Figura 10). Talvez essas proteínas estejam mediando transdução de sinais ou até mesmo regulando proteínas como NADPH oxidase em todos os horários de defesa da planta na PI561356, enquanto na BRS231 essa ativação somente seria necessária nos estágios iniciais da infecção (12, 24 e 48 hai). Porém métodos de validação seriam necessários para sustentar essa hipótese.

As proteínas “receptor-like kinase” (RLK) estão envolvidas em processos de desenvolvimento da planta como desenvolvimento meristemático, além da sinalização de rotas relacionadas à defesa da planta, atuando no reconhecimento de elicitores como flagelinas. RLKs podem ou não apresentar domínio extracelular, e quando apresentam geralmente é o domínio LRR (AFZAL, WOOD e LIGHTFOOT, 2008). O gene *xa21*, que confere resistência a *Xanthomonas oryzae* pv *oryzae* em arroz, é um exemplo de uma RLK (SONG *et al.*, 1995). Especula-se que a função da proteína “receptor-like kinase” seria reconhecer moléculas elicitoras do patógeno pelo domínio LRR e desencadear a fosforilação e transdução sinais da cascata via MAPK através do domínio RLK (AFZAL, WOOD e LIGHTFOOT, 2008), atuando como os receptores de padrões de reconhecimento na

imunidade de plantas (NÜRNBERGER e KEMMERLING, 2006). Portanto essas proteínas, que foram encontradas em todas as bibliotecas de ambos os genótipos (Figura 10), podem estar desempenhando também o papel de reconhecimento do elicitores não específicos e na transdução de sinais perante o patógeno.

**Figura 10** - Principais cinases envolvidas na transdução de sinais em resposta a estresse, encontradas em PI561356 (colunas em azul) e em BRS231 (colunas em cinza).



### 3.2.3 Categoria Funcional: Regulação da expressão gênica

Genes relacionados com a regulação da expressão gênica são importantes nos mecanismos de resistência visto que podem ativar e/ou reprimir muitos outros genes simultaneamente. PI561356 apresentou uma maior variedade em domínios de fatores de transcrição do que BRS231, e a quantidade de fatores de transcrição também foi maior no primeiro genótipo (Tabela 4). Uma das respostas para essa abundância de transcritos na PI561356 comparada com BRS231 seria justamente o fato de que o sequenciamento Solexa gera maior número de sequências, que como consequência, abrangeria o transcriptoma muito mais do que o sequenciamento Sanger, e também por provavelmente apresentar sequências com maior qualidade permitindo um melhor alinhamento das sequências e maiores informações dos transcritos (HANRIOT et al., 2008).

Em PI561356 os fatores de transcrição mais abundantes foram os domínios “zinc finger” (dedos de zinco), que podem receber outras denominações

devido a algumas diferenças nesse domínio (TAKATSUJI, 1998). Fatores de transcrição com esse domínio já foram previamente descritos na interação *P. pachyrhizi* – soja por Choi *et al.* (2008) e por van de Mortel *et al.*, (2007). Um “zinc finger” isolado de *A. thaliana* contendo o domínio RING-H2, apresentou indução quando a planta foi submetida e elicitores como quitina e celulase providas de fungos, indicando que esse gene está envolvido na resposta contra patógenos (SALINAS-MONDRAGÓN, GARCIDUEÑAS-PIÑA e GUZMÁN, 1999). Isso demonstra o quanto os produtos desses transcritos estão envolvidos na regulação de proteínas durante a defesa da planta perante o patógeno.

Outros fatores de transcrição que aparecem entre as bibliotecas de ambos os genótipos foram os WRKY. Os fatores de transcrição WRKY são caracterizados pela presença do domínio WRKY, cerca de 60 aminoácidos altamente conservados presente em todos os membros da família. Além disso, os WRKY apresentam domínios dedos de zinco e alguns apresentam também outros domínios como os zíperes de leucina, e possuem a habilidade de se ligar especificamente ao domínio (T)(T)TGAC(C/T), conhecido como W Box, presente na região promotores dos genes que regulam (EULGEM *et al.*, 2000).

Essas proteínas medeiam a resposta transcricional (resposta secundária) mediante os elicitores do patógeno, pois se ligam a sítios presentes no promotor de muitos genes associados a defesa, como os genes *PR* (EULGEM *et al.*, 2000). Diferentes membros da família dos WRKY foram previamente descritos como induzidos em resposta a infecção com FAS nos genótipos PI230970 e Embrapa 48 de soja (van de MORTEL *et al.*, 2007), em que análises de microarranjos demonstraram que 46 sondas correspondentes a essa família de genes foram diferencialmente expressas, porém, 18 foram diferencialmente expressas na condição inoculada no genótipo resistente à FAS (PI230970). Recentemente, Pandey *et al.* (2011) silenciaram os fatores de transcrição pertencentes as classes dos WRKY e MYB sendo eles: *GmWRKY36*, *GmWRKY40*, *GmWRKY45* e *GmMYB84* em plantas de soja contendo o alelo *Rpp2* que confere resistência e resposta de hipersensibilidade a FAS. As plantas com esses genes silenciados tornaram-se suscetíveis a ferrugem asiática, demonstrando que esses fatores de transcrição estão diretamente envolvidos na resposta contra o fungo por ativar genes envolvidos nas rotas de defesa. Os fatores MYB também estão presentes nas bibliotecas de ambos os genótipos (Tabela 4) e também já foram previamente

descritos na interação FAS- soja (van de MORTEL *et al.*, 2007; CHOI *et al.*, 2008; PANTHEE *et al.*, 2009).

Os fatores de transcrição NAC bem como outro domínio NAC denominado NAM, também já foram encontrados em interações compatíveis e incompatíveis de *P. pachyrhizi* com soja (van de MORTEL *et al.*, 2007; TREMBLAY *et al.*, 2010). Os Fatores de transcrição NAC estão presentes em ambos os genótipos somente nos “bulks” 1 e 3 (Tabela 4), talvez por ativar genes específicos no combate aos estágios iniciais e tardios de infecção.

Os domínios BEL-Like Homeodomain, KNOX homeobox, NDX homeobox e Auxin response factor (ARF) foram encontrados exclusivamente em PI561356 (Tabela 4). Esses dois primeiros domínios também foram encontrados por Choi *et al.*, (2008) que utilizou um genótipo de soja contendo *Rpp1*. Portanto esse dois domínios podem conter informações importantes, quanto aos genes que regulam e que estão envolvidos na resposta de resistência mediada por genes R nesse mesmo grupo de ligação. Já os NDX homeobox são fatores de transcrição relacionados ao desenvolvimento dos nódulos para fixação de nitrogênio em legumes (MUKHERJEE, BROCCIERI e BÜRGLIN, 2009). Estes fatores de transcrição podem não estar envolvidos somente neste processo biológico da soja, como a sua presença na biblioteca SSH de PI561356 pode indicar uma possível função na regulação de genes durante a defesa contra a ferrugem. Isto pode ser comprovado através da quantificação do nível de expressão desse gene em plantas falso inoculadas e inoculadas com a doença, por meio de RT-qPCR. A função dos ARFs é regulada por outro domínio ARF, o AUX-IAA, que leva a degradação de ambos por meio de degradação mediada por ubiquitina. ARFs induzem genes responsivos à auxina e AUX-IAA controla negativamente essa indução (SWAMINATHAN, PETERSON e JACK, 2008). Dentre os transcritos encontrados para ARF, dois são AUX-IAA e quatro são ARF. Apesar de trabalhos demonstrarem que o AS (acredita-se que AS seja o principal hormônio na resposta contra FAS, como já discutido anteriormente) inibe a expressão de vários genes induzidos por auxina (WANG *et al.*, 2007), muito ainda há para se descobrir acerca das relações hormonais durante a interação da planta com cada patógeno, pois muitos genes que são induzidos por hormônios antagonistas podem ser induzidos por hormônios antagonizados, como é o caso da lipase estudada por Oh *et al.* (2005).

O fator de transcrição GATA, encontrado exclusivamente em BRS231, apresenta domínio dedo de zinco (TAKATSUJI, 1998). Esse fator de transcrição teve sua função relacionada com a regulação da expressão de genes mediados por luz em plantas (REYES, MURO-PASTOR e FLORENCIO, 2004). GATA já foi descrito anteriormente por van de Mortel *et al.* (2007) em soja infectada por ferrugem, o que indica que esse fator de transcrição também pode estar relacionado com regulação de genes de defesa contra ferrugem, podendo ser validado a expressão de GATA através de RT-qPCR.

**Tabela 4 - Fatores de Transcrição mais representativos em PI561356 e BRS231.**

	PI561356				BRS231			
	Bulk1	Bulk2	Bulk3	Total	Bulk1	Bulk2	Bulk3	Total
Auxin response factor	6	-	-	6	-	-	-	-
Helix-loop-helix	7	-	2	9	2	-	-	2
BEL-Like Homeodomain	6	1	-	7	-	-	-	-
BZIP	6	-	2	8	3	1	-	4
Ethylene responsive TF	3	-	1	4	1	-	-	1
GATA	-	-	-	-	1	1	-	2
KNOX homeobox	3	-	-	3	-	-	-	-
Leucine zipper	3	4	-	7	4	1	-	5
MYB	4	1	2	7	3	4	-	7
NAC	7	-	2	9	2	-	1	3
NDX homeobox	2	-	-	2	-	-	-	-
Transcription Factor (TF)	6	5	5	16	3	8	7	18
Zinc Finger	18	3	5	26	4	-	-	4
WRKY	-	-	1	1	2	1	2	5
Total	71	14	20	105	25	16	10	51

#### 3.2.4 Categoria Funcional: Modificação/Catabolismo de proteínas

Sequências com similaridade a genes que codificam para E3 ubiquitina ligase (GO:0006464, GO:0006512, tabela 3) foram encontradas nas bibliotecas subtrativas de ambos os genótipos. Sabe-se que essa proteína é necessária para remover ou modificar a maioria dos peptídeos anormais e alguns reguladores de vida curta. Ela também influencia muitos processos como o ciclo

celular, transdução de sinais, transcrição e respostas contra estresse, incluindo defesa (CRAIG *et al.*, 2009). O gene *ACRE276*, uma E3 ubiquitina ligase de tabaco, está envolvida na resposta de hipersensibilidade e morte celular, pois quando plantas de tabaco reconhecem o elicitador, o peptídeo Avr9, elas apresentaram lesões nas folhas. Os transcritos do gene *ACRE276* foram rapidamente induzidos e constantes até a morte celular perante Avr9, porém, quando esse gene foi silenciado a célula perdeu a capacidade de resposta de hipersensibilidade (YANG *et al.*, 2006). É possível que essas proteínas estejam também associadas, principalmente, aos reguladores de vida curta e a transdução de sinais nos genótipos estudados, sendo que esses reguladores são responsáveis pela indução de genes expressos em condições de desafio, como no caso de ataque de patógenos.

A E3 ubiquitina ligase também participa do sistema de ubiquitinação para degradação de várias proteínas. NPR1 que, como já discutido, é uma proteína essencial, pois é coativadora da rota de AS, é degradada por meio do sistema de ubiquinização. É sabido que vários patógenos têm como alvo de seus elicitores proteínas envolvidas no processo de ubiquitinação, justamente para degradar essas proteínas chaves da própria planta em respostas contra patógenos, portanto eles utilizam da própria maquinaria da planta em seu benefício (TRUJILLO e SHIRASU, 2010). Foi encontrado no genoma de *A. thaliana*, pelo menos, 1400 proteínas preditas para E3 ligase (LECHNER *et al.*, 2006), portanto acredita-se que muitas dessas proteínas possam funcionar como proteínas chamarizes, uma vez que são alvos dos elicitores, ou também essas E3 ligases podem simplesmente reconhecer os elicitores para sua degradação via ubiquitinação (TRUJILLO e SHIRASU, 2010). Essa pode ser também uma das explicações para o aparecimento desses transcritos nas bibliotecas dos genótipos estudados.

A metionina sulfóxido redutase é uma proteína envolvida no reparo de danos causado pelo estresse oxidativo, pois geralmente a metionina é o aminoácido que mais sofre oxidação. Esses transcritos foram encontrados nas bibliotecas de BRS231, mas há evidências de que possam ser do patógeno, pois no trabalho realizado por El Hassouni *et al.*, (1999) relaciona essa proteína como sendo uma das defesas do patógeno *Erwinia chrysanthemi* contra as espécies reativas de oxigênio produzidas pela planta *Saintpaulia ionantha* em sua defesa.

Com relação ao processo de catabolismo de proteínas, foram encontrados transcritos para subunidades de 26s proteassoma, e protease

dependente de zinco em bibliotecas da PI561356, enquanto que para BRS231, foi encontrada subunidade regulatória de 26s protease, subunidade de proteassoma, além de ubiquitina.

As subunidades de 26s proteassoma e subunidade regulatória de 26s protease estão envolvidas no sistema de ubiquitina/26 proteassoma (UPS) que emprega um importante papel na degradação de proteínas da qual E3 ubiquitina ligase também faz parte. Esse sistema emprega um grande papel durante o processo de morte celular programada, e muitos componentes desse sistema já foram relacionados à defesa contra patógenos, como é o caso da própria E3 ligase (DIELEN *et al.*, 2010). Com o aparecimento dos transcritos para proteínas envolvidas nesse sistema de degradação de proteínas pode-se inferir que em ambos os genótipos esteja ocorrendo a degradação de proteínas através do sistema de ubiquitina/26 proteassoma durante o ataque do patógeno, o que pode culminar em morte celular programada, uma das respostas de defesa contra doenças.

### 3.3 SEQUÊNCIAS DESCONHECIDAS “UNKNOW” E “NO HITS”

Em um estudo realizado por Kruger *et al.* (2002), que construiu bibliotecas de cDNA em trigo infectado por *F. graminearum*, das 3.546 sequências não redundantes, 49% correspondiam a proteínas com função desconhecidas. Muitas sequências desconhecidas “*unknow*” e “no hits” foram também encontradas tanto na BRS231 como na PI561356 (Tabela 2). Portanto, é importante a investigação dessas sequências, já que elas podem codificar proteínas envolvidas nas rotas de sinalização e defesa da planta e que ainda não foram descritas. Devido à grande quantidade de sequências “no hits” que apareceram principalmente nas plantas do genótipo BRS231, algumas estratégias poderiam ser utilizadas para trazer informações mais detalhadas a respeito da expressão e validar um grande número de genes de uma só vez, como o “northern-blot”, microarranjos de cDNA e RNA-Seq. Após a identificação desses genes, podem ser delineadas estratégias para validar a função dos mesmos, e verificar se estes genes estão inseridos nos mecanismos de defesa das plantas. Metodologias que são amplamente empregadas para este fim são o silenciamento e a superexpressão gênica, através de técnicas transientes, como o VIGS e de incorporação definitiva, como na transformação genética.

### 3.4 RT-QPCR: IDENTIFICAÇÃO DE GENES NORMALIZADORES PARA ESTUDOS DE EXPRESSÃO GÊNICA EM RESPOSTA A FAS

Para evitar erros de normalização dos genes-alvo durante a quantificação relativa por RT-qPCR, foi realizada a análises de cinco genes constitutivos para a escolha dos dois melhores normalizadores. Os genótipos foram analisados juntos pelos dois algoritmos (geNorm e NormFinder), porém, como havia a hipótese de que o normalizador possa se expressar de forma diferencial em cada genótipo, essa análise também foi feita em separado.

Os resultados de expressão e estabilidade dos endógenos obtidos através do algoritmo NormFinder mostraram que os genes mais estáveis foram  $\beta$ -actina e GAPDH para a PI561356 e  $\beta$ -actina e Elf1b para BRS231. Quando os dois genótipos foram analisados em conjunto,  $\beta$ -actina e Elf1b também foram os genes mais estáveis (Tabela 5).

**Tabela 5** - Valores de estabilidade obtidos através do programa NormFinder, onde os melhores genes normalizadores são os que apresentam os valores mais próximos a 0.

Gene	Valor de estabilidade		
	PI561356	BRS231	Genótipos juntos
Btub	0,614	0,565	0,536
Bact	<b>0,493</b>	<b>0,334</b>	<b>0,348</b>
Elf1b	0,795	<b>0,334</b>	<b>0,421</b>
GAPDH	<b>0,509</b>	2,372	1,944
RNA18s	0,748	2,498	1,945

Com o uso do algoritmo geNorm o resultado foi mais discrepante para PI561356, em que  $\beta$ -tubulina e GAPDH foram os genes mais estáveis (Figura 11B). Já para BRS231,  $\beta$ -actina e Elf1b (os mesmo apontados pelo NormFinder), foram os mais estáveis, enquanto que para os dois genótipos em conjunto, o resultado foi similar ao apresentado pelo NormFinder, com  $\beta$ -actina e Elf1b apresentando maior estabilidade (Figura 11A).

Como a intenção dessa análise era buscar os dois melhores normalizadores para serem utilizados ao mesmo tempo em ambos os genótipos, optou-se pelo uso da  $\beta$ -tubulina e  $\beta$ -actina, pois nas análises em separado geNorm

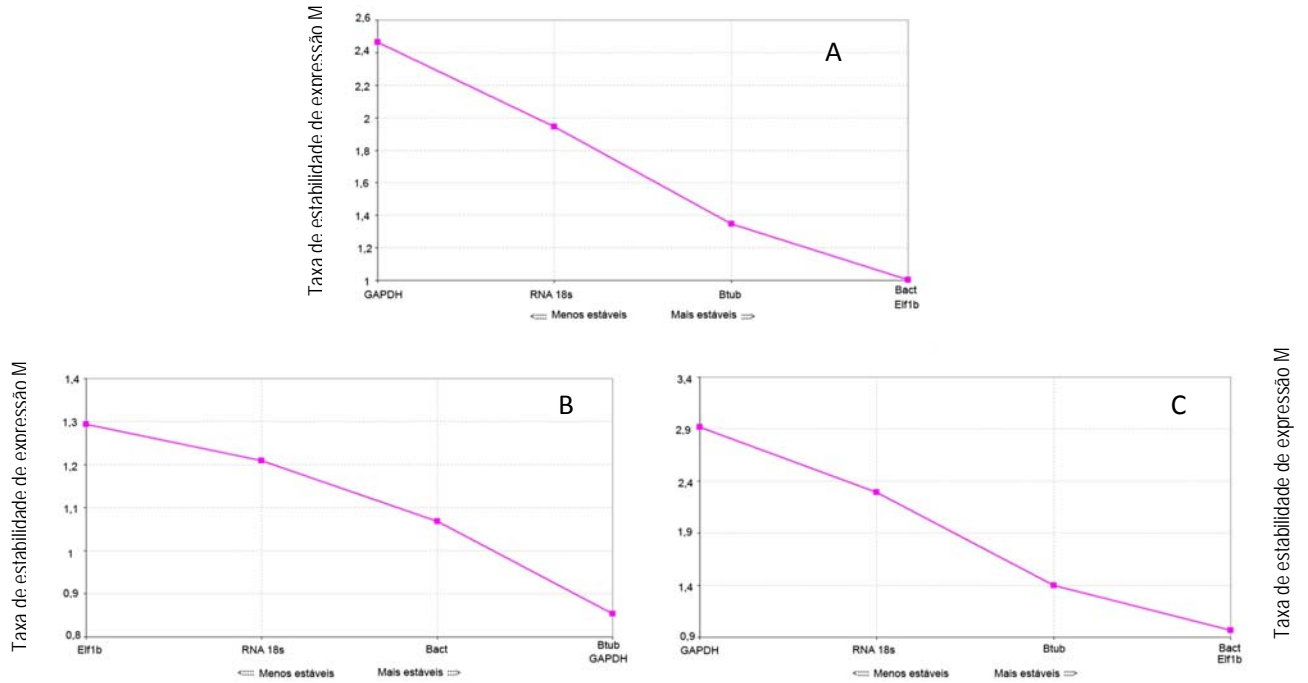
apontou Elf1b como sendo o normalizador menos estável na PI561356 (Figura 11B) e GAPDH como sendo o menos estável em BRS231(Figura 11C).  $\beta$ -tubulina é o terceiro gene mais estável segundo tanto na análise da BRS231 quanto na análise conjunta (Figura 11C e A); já a  $\beta$ -actina é o terceiro gene mais estável na PI561356 (Figura 11B). NormFinder também aponta Elf1b como o menos estável na PI561356 e GAPDH como o segundo menos estável em BRS231, já a  $\beta$ -tubulina mostra-se como o terceiro gene mais estável em todas as análises realizadas por esse algoritmo (Tabela 5).

### 3.5 RT-QPCR: VALIDAÇÃO DOS GENES IDENTIFICADOS NAS BIBLIOTECAS VIA HIBRIDIZAÇÃO SUBTRATIVA

Sete genes pertencentes a pelo menos três das principais categorias funcionais identificadas nas bibliotecas de ambos os genótipos foram selecionados para a validação dos resultados obtidos (Tabela 4). Visto que os mRNAs sequenciados em ambas as bibliotecas são oriundos de subtração unidirecional, sendo as amostras falso inoculadas utilizadas como “drivers”, espera-se que os genes identificados, em sua grande maioria, sejam aqueles induzidos após a infecção com FAS em pelo menos um dos tempos de coletas e/ ou montagem dos bulks.

Os oligonucleotídeos iniciadores para esses genes foram predominantemente desenhados na região predita como 3'UTR, com exceção de KR4, cuja sequência não apresenta essa região conhecida (Tabela 6). A quantificação relativa de todos os genes está representada na Figura 12.

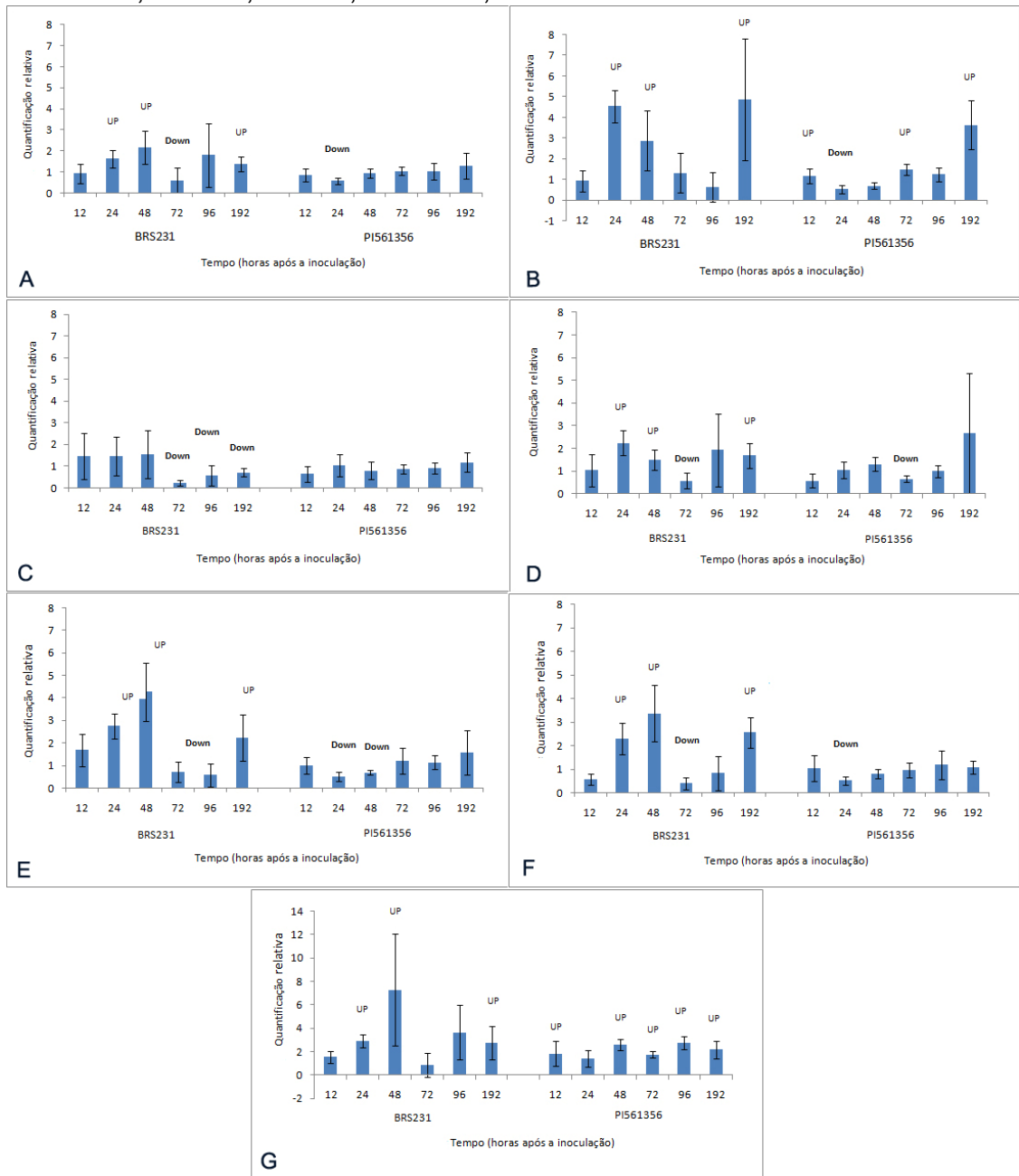
**Figura 11** - Gráficos obtidos através do programa geNorm apontando os genes mais estáveis com os menores valores de M. Em A: Ambos os genótipos quando analisados conjuntamente, B: em PI561356 quando analisada em separado, e em C: BRS231 quando analisada separadamente.



**Tabela 6 -** Transcritos escolhidos para a validação das bibliotecas SSH. A identificação corresponde as sequências das bibliotecas SSH de BRS231 depositadas no banco de dados da Embrapa Soja e ao modelo genico do Phytozome.

	Categoria Funcional	Primer	Identificação	Sequência	Amplicon	Eficiência
Ankyrin-3, Ankyrin-G	Ligação a proteína	Ankyrin	PLACA08F_H12_b_096.ab1	5'-TCAAAGCCACCAACCGCA-3' 3'-CCGAGGTAAGTACGAGTTCCA-5'	82pb	97,64%
disease resistance protein	Ligação a proteína	DRP	PLACA24F_G02_b_008.ab1	5'- TTTGAAGGTGGTATGTGATGAA-3' 3'- GAACCAGGTCATGTCTTAGGAGCTAA-5'	148pb	93,34%
resistance protein KR4	Apoptose	KR4	PL10F_A09_b_065.ab1	5'-TTGCCCTCCCTGGACTTTTG-3' 3'-GGTGCCCATAGAGGAGAAGTTCA-5'	148pb	97,31%
leucine-rich repeat-containing protein, putative	Membrana plasmática	LRR	PL06F_C10_b_070.ab1	5'- ATCCGTCATTAACCTACCACCACT-3' 3'-GCCCTTTCTTCGAGACT-5'	103pb	91,31%
Mitogen-activated protein kinase kinase 4	Cinase	MPKK	PL33F_H04_b_032.ab1	5'-GGACGACGGCTCAGCTCTT-3' 3'- AGGTGATTCCAAGATCCAAGTCATAA-5'	110pb	95,18%
Serine/threonine-protein kinase HT1, High leaf temperature protein 1	Cinase	Serine/Threonine	PLACA21F_G12_b_088.ab1	5'-CATCACTGGCACTTGTCACTTCAT-3' 3'-GTAGATAAGGCAACCGCTGGA-5'	101pb	83,70%
Pfam:03106 WRKY DNA-binding domain	Fator de transcrição	WRKY23	Glyma15g11680	5'-CACTCTCAGCTTCAGCACCATTC-3' 3'-GCGGCTCTGACCCAAAGT-5'	142pb	99,79%

**Figura 12** - Níveis de expressão gênica nos dois genótipos de soja (BRS231 e PI561356) em seis horários após inoculação com FAS. A: Ankyrin, B: DRP, C: KR4, D:LRR, E: MPKK, F: Serina/Treonina e G: WRKY23.



Dos sete genes avaliados, seis tiveram os níveis de expressão significativamente induzidos em pelo menos um dos tempos analisados para o genótipo BRS231, enquanto para o genótipo PI561356 apenas 2 genes puderam validar a subtração realizada. Isso pode ter ocorrido devido ao fato dos primers terem sido desenhados nas regiões 3' UTR dos transcritos para BRS231. Era sabido que os primers estavam na região 3' UTR devido a um alinhamento contra os modelos de genes da cultivar norte americana Williams 82 depositada no banco de

dados do Phytozome. O alinhamento demonstrou alguns “gaps” nessa região e, portanto, assim como havia esses “gaps” entre BRS231 e Willians 82, pode haver também entre BRS231 e PI561356, tornando os primers validáveis em BRS231 e não em PI561367. As regiões UTR são regiões passíveis de sofrer mutações e, portanto, constituem locais ideais para desenho de primers quando se quer apenas um gene dentre tantos de uma família gênica. Por outro lado, estão sujeitas à variação entre genótipos, o que pode ter sido a causa dos resultados obtidos neste experimento.

O gene que codifica a proteína com domínios anquirina foi regulado positivamente em 24, 48 e 192 hai, e negativamente em 72 hai no genótipo BRS231 (Figura 12 A). Já na PI561356, esse mesmo gene apresentou expressão diferencial negativa em 24 hai (Figura 12 A). Os domínios de anquirina foram encontrados no gene NPR-1 que codifica uma proteína chave na transdução de sinal mediada por SA, hormônio principalmente ativo em resposta a fungos biotróficos. Esses domínios são especialmente importantes durante a interação proteína-proteína.

O transcrito com similaridade aos genes que codificam para as proteínas *Disease resistance protein* DRP e KR4, uma proteína contendo domínio LRR, foram selecionados por possuírem domínios comumente presentes em genes de resistência. O gene DRP foi induzido nas plantas do genótipo BRS231 nos horários 24, 48 e 192 hai (Figura 12 B), portanto ele foi diferencialmente expresso não somente no “bulk” 3 como apresentado na Tabela 4, mas também em alguns dos horários iniciais da infecção por ferrugem. Na PI561356, esse gene apresentou-se “up”-regulado em 12, 72 e 192 hai, e “down”-regulado em 24 hai (Figura 12 B), o que significa que ele é expresso durante todo o período de infecção das plantas.

O gene KR4 foi “down”-regulado, na BRS231 em 72, 96 hai e 192 hai, não tendo sido observada expressão diferencial em nenhum dos tempos analisados PI561356 (Figura 12 C).

O gene que codifica o domínio LRR foi “up”-regulado na BRS231 em 24, 48 e 192 hai, e “down”-regulado em 72 hai (Figura 12 D); já na PI561356 esse gene foi regulado negativamente em 72 hai (Figura 12 D). Este domínio é amplamente presente nos genes de resistência identificados até o momento, como por exemplo, RPS2 (proteína com domínio NBS-LRR) de *Arabidopsis*, que confere resistência a *Pseudomonas syringae* (KIM et al., 2009); nos genes NtLRR1 e NtLRR2 de tabaco que conferem resistência a *P. syringae* e ao vírus do mosaico do

tabaco (XU et al., 2009); nos genes Lr10 e RGA2 que conferem resistência a ferrugem da folha do trigo (LOUTRE, et al., 2009); e também no gene R1 contra a doença causada por *Phytophthora infestans* em batata (BALLVORA et al., 2002).

KR4 e DRP são proteínas que foram associadas a resistência ao vírus do mosaico da soja que apresentam domínio NBS-LRR em dois genótipos de soja diferentes (WANG et al., 2004; HAYES et al., 2004). Em outro estudo, um gene contendo o domínio NBS-LRR também foi associado a resistência a ferrugem asiática (MEYER et al., 2009). Esse gene, quando silenciado no genótipo resistente PI459025B, reverteu o fenótipo de resistência (lesão RB) para o de suscetibilidade (TAN), e é um forte candidato a ser o alelo de resistência *Rpp4*.

O gene que codifica uma proteína cinase MPKK foi "up"-regulado na BRS231 em 24, 48 e 192 hai, e "down"-regulado em 96 hai, enquanto que na PI561356 ele foi "down"-regulado em 24 e 48 hai (Figura 12 E). É sabido que as diferentes combinações de MAPK, MPKK e MPKKK são responsáveis pela percepção de sinais em nível de membrana e imediatamente medeiam a transdução de sinais intracelularmente. O exemplo amplamente conhecido é a ativação de cascata de sinalização mediada por MAPK (*AtMMPK3/6* e *AtMKK4/5*), em *A. thaliana*, em resposta ao peptídeo elicitor bacteriano Flg2. Além de serem ativadas por Flg22, as MAPK cinases também podem desencadear a transdução de sinal ativada por outros estresses como salino e oxidativo, bem como elicitores fungicos, em *Arabidopsis* e tabaco. (PETERSEN et al., 2000; QIU et al., 2008; PEDLEY e MARTIN, 2005). A ativação de MAP cinases no genótipo BRS231 pode ser explicada pelo reconhecimento de elicitores fungicos inespecíficos, tais como fragmentos de quitina, peptídeoglicano, entre outros, capazes de ativar respostas basais de resistência, uma vez que até o momento nenhum gene maior de resistência a ferrugem foi descrito neste genótipo. Sabe-se que a ativação das respostas a infecção a patógenos mediada por tais elicitores, também denominadas padrões moleculares associados ao microrganismo ou ao patógeno (MAMPs/PAMPs), inicia-se pela percepção e reconhecimento dos mesmos por proteínas receptoras presentes na membrana denominados de receptores de padrões de reconhecimento (PRRs). As cascatas de transdução de sinal desencadeadas após a percepção de MAMPs/PAMPs frequentemente envolvem a participação de MAPK cinases, como *AtMMPK3/6* em *Arabidopsis*, em resposta a

Flg2, e NtMEK2 em tabaco em resposta a elicitores fúngicos (TSUDA e KATAGIRI, 2010).

O gene que codifica serina/treonina foi "up"-regulado na BRS231 em 24, 48 e 192 hai e reprimido em 72 hai. Na PI561356 esse gene apresentou valores significativos em 24 hai onde foi "reprimido (Figura 12 F). Serina/Treonina cinase é um importante domínio encontrado em proteínas de resistência que desencadeia transdução de sinais e em proteínas do tipo RLK. Em arroz, esse domínio está presente no gene *xa21* que confere resistência a *X. oryzae* pv *oryzae* (SONG *et al.*, 1995). Em cevada, foi isolada uma proteína de resistência a ferrugem do caule que continha uma estrutura não usual de proteína de resistência, pois a proteína em questão, RPg5, apresentava três domínios de resistência: NBS, LRR e Serina/Treonina cinase (BRUEGGEMAN *et al.*, 2008). Porém como já discutido anteriormente, esse domínio pode ser encontrado também em proteínas RLK que atuam como PRR reconhecendo os PAMP/MAMP (NÜRNBERGER e KEMMERLING, 2006). Portanto, a proteína codificada por esse gene pode estar também atuando no reconhecimento dos padrões moleculares do patógeno na BRS231, já que esta não possui um gene R, e desencadeando a transdução de sinais para que outras proteínas sejam traduzidas para a defesa da planta, uma vez que o gene está presente em horários iniciais de resposta de defesa a doença. Também pode atuar como mediador de transdução de sinais em horários mais tardios de infecção, visto que é induzido nesses horários também.

O gene WRKY23 teve sua expressão aumentada na BRS231 em 12, 24 e 48 hai e em todos os tempos na PI561356, exceto em 24 hai (Figura 12 G). Já foi demonstrado que fatores de transcrição WRKY, como o WRKY70, podem auxiliar na modulação hormonal contra patógenos biotróficos, como é o caso da FAS, pois inibem respostas por outros hormônios que não o AS (LI *et al.*, 2004). Recentemente, o envolvimento desses fatores de transcrição foram associados a resposta a FAS, em que, dentre outros genes, três fatores de transcrição dessa família, *GmWRKY36*, *GmWRKY40*, *GmWRKY45*, foram silenciados em soja resistente a ferrugem asiática contendo o gene *Rpp2* (PANDEY *et al.*, 2011). Esses genes quando silenciados tornaram as plantas suscetíveis à ferrugem, principalmente nas plantas contendo *GmWRKY36* silenciado, onde foi vista a diminuição da transcrição de flavina monoxigenase, transportador de prolina, *GmWRKY12*, *GmWRKY19* e *GmPR1*, que são genes envolvidos na defesa da

planta bem como na transcrição de genes que podem estar também envolvidos na defesa.

As W-box (sequências conservadas de reconhecimento dos WRKY) estão presentes no promotor de diversos genes responsáveis pela defesa da planta contra patógenos como PR-10 (RUSHTON *et al.*, 1996), quitinase (YAMAMOTO *et al.*, 2004) e também dos genes para NPR1 (YU *et al.*, 2001), que auxiliam na indução de AS. Portanto, visto que o gene WRKY23 foi regulado positivamente na condição de infecção, estudos para caracterizar a rota em que esse gene está inserido (através de silenciamento e superexpressão gênica) e quais genes são responsáveis pela indução da transcrição, podem auxiliar na compreensão dos mecanismos de defesa das planta contra à FAS, para que se possa delinear estratégias de transgenia através das sequências promotoras reconhecidas por esse fator de transcrição.

## CONCLUSÃO

Aparentemente, muitos genes apresentaram respostas similares em ambos os genótipos, PI561356 e BRS231 após a infecção por FAS. Porém, a identificação de transcritos que codificam para a proteína chalcona sintase e chalcona isomerase somente em PI561356, além de 4-coumarate-CoA Ligase aparecer nos tempos iniciais (“bulks” 1 e 2) em PI561356 e somente no “bulk” 3 em BRS231 pode trazer evidências de que a rota de fenilpropanóides está sendo ativada mais cedo em PI561356 do que em BRS231. Os domínios de fatores de transcrição exclusivos em PI561356 (BEL-Like Homeodomain, KNOX homeobox, NDX homeobox e Auxin response factor), bem como GATA em BRS231, podem indicar um diferencial na resposta de resistência, uma vez que esses fatores de transcrição podem estar ativando genes genótipo específicos ou até mesmo genes ainda não caracterizados pertencentes aos “no hits”. Portanto as sequências que não apresentam similaridade (“unknow” e “no hits”) com proteínas já descritas, principalmente em BRS231, podem estar contribuindo para essa diferença também.

Quanto aos genes validados, WRKY23 e DRP mostraram-se importantes para ambos os genótipos, pois foram induzidos em diferentes horários nos mesmos. KR4 foi regulado negativamente em BRS231 e não foi diferencialmente expresso na PI561356. Os demais genes apresentaram-se induzidos em BRS231 em algum dos horários analisados, porém foram regulados negativamente em PI561356. Portanto, esses genes devem pertencer ao mecanismo de resistência encontrado em BRS231, ou então, podem ser superexpressos em PI561356 em horários que não foram quantificados neste trabalho.

Deve-se investigar melhor o nível de expressão de outros genes que codificam proteínas envolvidas nos mecanismos de resistência, como genes envolvidos em rotas de sinalização e estresse oxidativo, outros fatores de transcrição e genes envolvidos na rota dos fenilpropanóides, pois como demonstrado, os mesmos genes estão presentes em genótipos diferentes, porém existem diferenças na expressão, e podem ser essas diferenças, o ponto chave causador de divergência entre os tipos de resistência.

A utilização de resistências qualitativa e quantitativa pode ser utilizada em estratégias de melhoramento, pois, a união desses dois mecanismos de

resistência conferiria resistência mais duradoura contra o patógeno. No entanto, são necessários mais estudos moleculares para caracterização dos genes envolvidos nos diferentes mecanismos de resistência, de forma, a saber, o que causa a diferença entre os mecanismos de resistência: se são diferentes genes expressos durante a infecção do patógeno, ou diferentes rotas metabólicas ativadas por diferentes respostas hormonais induzidas na planta.

## REFERÊNCIAS BIBLIOGRÁFICAS

- AFZAL, A. J.; WOOD, A. J.; LIGHTFOOT, D. A. Plant Receptor-Like Serine Threonine Kinases: Roles in Signaling and Plant Defense. **Molecular Plant-Microbe Interactions**, v. 21, n. 5, p. 507–517, 2008.
- ASAI, T. et al. MAP kinase signaling cascade in *Arabidopsis* innate immunity. **Nature**, v. 415, p. 977-983, 2002.
- BALLVORA, A. et al. The R1 gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. **The Plant Journal**, v. 30, n. 3, p. 361-371, 2002.
- BRUEGGEMAN, R. et al. The stem rust resistance gene Rpg5 encodes a protein with nucleotide-binding-site, leucine-rich, and protein kinase domains. **PNAS**, v. 105, n. 39, p. 14970–14975, 2008.
- CAO, H. et al. The *Arabidopsis* NPR1 Gene That Controls Systemic Acquired Resistance Encodes a Novel Protein Containing Ankyrin Repeats. **Cell**, v. 88, p. 57–63, 1997.
- CAMARGO, L. E. A. Análise genética da Resistência e da Patogenicidade. In: Bergamin Filho, A; Kimati, H.; Amorin, L. (Ed.). **Manual de Fitopatologia: Princípios e Conceitos**. São Paulo. Agronômica CERES, ed. 3; v. 1, p. 470-492, 1995.
- CAMARGO, P. O. **Estudo da Herança de Caracteres Quali-Quantitativos e Mapeamento Genético de Alelos Resistentes à Ferrugem Asiática da Soja Presentes nas Pi561356 e Pi594754**. Tese (Mestrado) - Universidade Estadual de Londrina, Londrina, Paraná, 2010.
- CHOI, J. J. et al. Expression patterns in soybean resistant to *Phakopsora pachyrhizi* reveal the importance of peroxidases and lipoxygenases. **Functional & integrative genomics**, v. 8, p. 341-359, 2008.
- CONAB. **Boletim Indicadores da Agropecuária**- Nº 8. 2010 Disponível em <<http://www.conab.gov.br>>. Acesso em 12/10/2010.
- CONAB. **Acompanhamento de safra brasileira: grãos 2010-2011, quarto levantamento, janeiro 2011 / Companhia Nacional de Abastecimento**. – Brasília: Conab, 2011.
- CRAIG, A. et al. E3 ubiquitin ligases and plant innate immunity. **Journal of Experimental Botany**, v. 60, n. 4, p. 1123–1132, 2009.
- DEEPAK, S. et al. Role of hydroxyproline-rich glycoproteins in resistance of pearl millet against downy mildew pathogen *Sclerospora graminicola*. **Planta**, v. 226, p. 323–333, 2007.

- DIELEN, A. S. et al. The ubiquitin/26S proteasome system in plant–pathogen interactions: a never-ending hide-and-see game. **Molecular Plant Pathology**, v. 11, n. 2, p. 293-308, 2010.
- DING, X. et al A Rice Kinase-Protein Interaction Map. **Plant Physiology**, v. 149, p. 1478-1492, 2009.
- DIXON, R. A.; PAIVA, N. L. Stress-Induced Phenylpropanoid Metabolism. **The Plant Cell**, v. 7, p. 1085-1097, 1995.
- EL HASSOUNI, M. et al. The minimal gene set member *mrsA*, encoding peptide methionine sulfoxide reductase, is a virulence determinant of the plant pathogen *Erwinia chrysanthemi*. **PNAS**, v. 96, p. 887-892, 1999.
- EICHHORN, H. et al. Isolation of a novel ABC-transporter gene from soybean induced by salicylic acid. **Journal of Experimental Botany**, v. 57, n. 10, p. 2193–2201, 2006.
- EULGEM, T. et al. The WRKY superfamily of plant transcription factors. **Trends in Plant Science**, v. 5, n. 5, p. 199-206, 2000.
- FEUSSNER, I.; WASTERNAK, C. The Lipoxygenase Pathway. **Annual review of plant biology**, v. 53, p. 275-97, 2002.
- GÖTZ, S. et al. High-throughput functional annotation and data mining with the Blast2GO suite. **Nucleic Acids Research**, v. 36, n. 10, p. 3420-3435, 2008.
- GREENBERG, J. T.; YAO, N. The role and regulation of programmed cell death in plant–pathogen interactions. **Cellular Microbiology**, v. 6, n. 3, p. 201-211, 2004.
- HAMBERGER, B.; HAHLBROCK, K. The 4-coumarate:CoA ligase gene family in *Arabidopsis thaliana* comprises one rare, sinapate-activating and three commonly occurring isoenzymes. **PNAS**, v. 101, n. 7, p. 2209-2214, 2004.
- HANRIOT, L. et al. A combination of LongSAGE with Solexa sequencing is well suited to explore the depth and the complexity of transcriptome. **BMC Genomics**, v.9, 2008.
- HAYES, A. J. et al. Recombination Within a Nucleotide-Binding-Site/Leucine-Rich-Repeat Gene Cluster Produces New Variants Conditioning Resistance to Soybean Mosaic Virus in Soybeans. **Genetics**, v. 166, p. 493–503, 2004.
- HORNER, D. S. et al. Bioinformatics approaches for genomics and post genomics applications of next-generation sequencing. **Briefings In Bioinformatics**, v. 11, n. 2, p. 181-197, 2009.
- HÜCKELHOVEN, R. Cell wall-associated mechanisms of disease resistance and susceptibility. **Annual review of Phytopathology**, 45:101–127, 2007.
- KIM, D-Y. et al. The ABC transporter AtPDR8 is a cadmium extrusion pump conferring heavy metal resistance. **The Plant Journal**, v. 50, p. 207–218, 2007.

KIM, M. G. et al. The *Pseudomonas syringae* type III effector AvrRpm1 induces significant defenses by activating the Arabidopsis nucleotide-binding leucine-rich repeat protein RPS2. **The Plant Journal**, v. 57, p. 645-653, 2009.

KLIEBENSTEIN, D. J.; MONDE, R. A.; LAST, R. L. Superoxide Dismutase in *Arabidopsis*: An Eclectic Enzyme Family with Disparate Regulation and Protein Localization. **Plant Physiology**, v. 118, p. 637-650, 1998.

KOBAE, Y. et al. Loss of AtPDR8, a Plasma Membrane ABC Transporter of *Arabidopsis thaliana*, Causes Hypersensitive Cell Death Upon Pathogen Infection. **Plant Cell Physiology**, v. 47, n. 3, p. 309-318, 2006.

KOBAYASHI, M. et al. Calcium-Dependent Protein Kinases Regulate the Production of Reactive Oxygen Species by Potato NADPH Oxidase. **The Plant Cell**, v. 19, p. 1065-1080, 2007.

KOSKI, L. B. et al. AutoFACT: An Automatic Functional Annotation and Classification Tool. **BMC Bioinformatics**, v. 6, n. 151, 2005.

KRUGER, W. M. et al. Functional and Comparative Bioinformatic Analysis of Expressed Genes from Wheat Spikes Infected with *Fusarium graminearum*. **Molecular Plant-Microbe Interactions**, v. 15, n. 5, p. 445-455, 2002.

KWON, S. J. et al. GDSL lipase-like 1 regulates systemic resistance associated with ethylene signaling in *Arabidopsis*. **The Plant Journal**, v. 58, p. 235-245, 2009.

LECHNER, E. et al. F-box proteins everywhere. **Current Opinion in Plant Biology**, v. 9, p. 631-638, 2006.

LI, J.; BRADER, G.; PALVA, E. T. The WRKY70 Transcription Factor: A Node of Convergence for Jasmonate-Mediated and Salicylate-Mediated Signals in Plant Defense. **The Plant Cell**, v. 16, p. 319-331, 2004.

LI, C.; BAI, Y.; JACOBSEN, E. Tomato defense to the powdery mildew fungus: differences in expression of genes in susceptible, monogenic- and polygenic resistance responses are mainly in timing. **Plant Molecular Biology**, v. 62, p. 127-140, 2006.

LI, R. et al: short oligonucleotide alignment program. **Bioinformatics**, v. 24, n. 5, p. 713-714, 2008.

LIBAULT, M. et al. Complete Transcriptome of the Soybean Root Hair Cell, a Single-Cell Model, and Its Alteration in Response to *Bradyrhizobium japonicum* Infection. **Plant Physiology**, v. 152, p. 541-552, 2010.

LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. **Methods**, v. 25, p. 402-408, 2001.

- LOUTRE, C. et al. Two different CC-NBS-LRR genes are required for Lr10-mediated leaf rust resistance in tetraploid and hexaploid wheat. **The Plant Journal**, v. 60, p. 1043-1054, 2009.
- NÜRNBERGER, T.; KEMMERLING, B. Receptor protein kinases - pattern recognition receptors in plant immunity. **Trends in Plant Science**, v. 11, n. 11, p. 519-522, 2006.
- MEYER, J. D. F. et al. Identification and Analyses of Candidate Genes for Rpp4-Mediated Resistance to Asian Soybean Rust in Soybean. **Plant Physiology**, v. 150, p. 295-307, 2009.
- MHAMDI, A. et al. Catalase function in plants: a focus on *Arabidopsis* mutants as stress-mimic models. **Journal of Experimental Botany**, p.1-24, 27 set 2010.
- MITTLER, R. Oxidative stress, antioxidants and stress tolerance. **Trends in Plant Science**, v.7, n.9, 2002.
- MUKHERJEE, K.; BROCCHERI, L.; BÜRGLIN, T. R. A Comprehensive Classification and Evolutionary Analysis of Plant Homeobox Genes. **Molecular biology and evolution** ., v. 26, n. 12, p. 2775-2794, 2009.
- OH, I. S. et al. Secretome Analysis Reveals an *Arabidopsis* Lipase Involved in Defense against *Alternaria brassicicola*. **The Plant Cell**, v. 17, p. 2832-2847, 2005.
- OLIVEROS, J.C. (2007). An interactive tool for comparing lists with Venn Diagrams. Disponível em: <<http://bioinfogp.cnb.csic.es/tools/venny/index.html>>. Acesso em: 15/02/2011.
- PADMANABHAN, M. et al. The leucine-rich repeat domain in plant innate immunity: a wealth of possibilities. **Cellular Microbiology**, v. 11, n. 2, p. 191–198, 2009.
- PANDEY, A. K.; YANG, C.; ZHANG, C.; GRAHAM, M. A.; HORSTMAN, H. D.; LEE, Y.; ZABOTINA, O. A.; HILL, J. H.; PEDLEY, K. F.; WHITHAM, S. A. Functional Analysis of the Asian Soybean Rust Resistance Pathway Mediated by Rpp2. **Molecular Plant-Microbe Interactions**, v. 24, n. 2, p. 194–206, 2011.
- PANTHEE, D. R. et al. Differential expression of genes in soybean in response to the causal agent of Asian soybean rust (*Phakopsora pachyrhizi* Sydow) is soybean growth stage-specific. **Theoretical and Applied Genetics**, v. 118, p. 359-370, 2009.
- PEDLEY, K. F.; MARTIN, G. B. Role of mitogen-activated protein kinases in plant immunity. **Current Opinion in Plant Biology**, v. 8, 541-547, 2005.
- PETERSEN, M. et al. *Arabidopsis* MAP Kinase 4 Negatively Regulates Systemic Acquired Resistance. **Cell**, v. 103, p. 1111-1120, 2000.
- PFAFFL, M. W.; HORGAN, G. W.; DEMPFFLE, L. Relative expression software tool (REST<sup>®</sup>) for group-wise comparison and statistical analysis of relative expression results in real time-PCR. **Nucleic Acids Research**, v. 30, n. 9, 2002.

QIU, J. L. et al. *Arabidopsis* Mitogen-Activated Protein Kinase Kinases MKK1 and MKK2 Have Overlapping Functions in Defense Signaling Mediated by MEKK1, MPK4, and MKS1. **Plant Physiology**, v. 148, p. 212-222, 2008.

REYES, J. C.; MURO-PASTOR, M. I.; FLORENCIO, F. J. The GATA Family of Transcription Factors in *Arabidopsis* and Rice. **Plant Physiology**, v. 134, p. 1718-1732, 2004.

REN, D. et al. Activation of Ntf4, a Tobacco Mitogen-Activated Protein Kinase, during Plant Defense Response and Its Involvement in Hypersensitive Response-Like Cell Death. **Plant Physiology**, v. 141, p. 1482-1493, 2006.

RIBEIRO, A. S. et al. Genetic control of Asian rust in soybean. **Euphytica**, v. 157, p. 15-25, 2007.

ROMEIS, T. et al. Calcium-dependent protein kinases play an essential role in a plant defence response. **The EMBO Journal**, v. 20, n. 20, p. 5556-5567, 2001.

RUSHTON, P. J. et al. Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes. **The EMBO Journal**, v.15, n.20, p. 5690-5700, 1996.

SALINAS-MONDRAGÓN, R. E.; GARCIDUEÑAS-PIÑA, C.; GUZMÁN, P. Early elicitor induction in members of a novel multigene family coding for highly related RING-H2 proteins in *Arabidopsis thaliana*. **Plant Molecular Biology**, v. 40, p. 579-590, 1999.

SCHMUTZ, J. et al. A Genome Sequence of the Palaeopolyploid Soybean. **Nature**, v. 463, p. 178-183, 2010.

SCHNEIDER, R. W. et al. First report of soybean rust caused by *Phakopsora pachyrhizi* in the continental United States. *Plant Disease* 89:774-774 (2005).

SCHWEIGHOFER, A. et al. The PP2C-type phosphatase AP2C1, which negatively regulates MPK4 and MPK6, modulates innate immunity, jasmonic acid, and ethylene levels in *Arabidopsis*. **The Plant Cell**, v. 19, p. 2213–2224, 2007.

SONG, W. Y. et al. Receptor Kinase-Like Protein Encoded by the Rice Disease Resistance Gene, Xa21. **Science**, v. 270, p. 184-186, 1995.

SORIA-GUERRA, R. T. et al. Transcriptome analysis of resistant and susceptible genotypes of *Glycine tomentella* during *Phakopsora pachyrhizi* infection reveals novel rust resistance genes. **Theoretical and Applied Genetics**, v. 120, p. 1315-1333, 2010.

SPOEL, S. H. et al. NPR1 modulates cross-talk between salicylate- and jasmonate-dependent defense pathways through a novel function in the cytosol. **The Plant Cell**, v. 15, p. 760–770, 2003.

STONE, J. M.; WALKER, J. C. Plant Protein Kinase Families and Signal Transduction. **Plant Physiology**, v. 108, p. 451-457, 1995.

SWAMINATHAN, K.; PETERSON, K.; JACK, T. The plant B3 superfamily. **Trends in Plant Science**, v.13, n.12, 2008.

TAIZ, L.; ZEIGER, E. **Fisiologia Vegetal**. Porto Alegre: Artmed, ed. 4, 2009.

TAKATSUJI, H. Zinc-finger transcription factors in plants. **CMLS**, v. 54, p. 582-596, 1998.

TENA, G.; ASAI, T.; CHIU, W. L.; SHEEN, J. Plant mitogen-activated protein kinase signaling cascades. **Current Opinion in Plant Biology**, v. 4, p. 392-400, 2001.

TREMBLAY, A. et al. Transcriptome analysis of a compatible response by *Glycine max* to *Phakopsora pachyrhizi* infection. **Plant Science**, v. 179, p. 183-193, 2010.

TRUJILLO, M.; SHIRASU, K. Ubiquitination in plant immunity. **Current Opinion in Plant Biology**, v. 13, p. 402-408, 2010.

TSUDA, K.; KATAGIRI, F. Comparing signaling mechanisms engaged in pattern-triggered and effector-triggered immunity. **Current Opinion in Plant Biology**, v. 13, p. 459-465, 2010.

UMEZAWA, T. et al. Type 2C protein phosphatases directly regulate abscisic acid-activated protein kinases in Arabidopsis. **PNAS**, v. 106, n. 41, p. 17588–17593, 2009.

van de MORTEL, M. V. D. et al. Distinct biphasic mRNA changes in response to Asian Soybean Rust Infection. **Molecular Plant-Microbe Interactions**, Saint Paul, v. 20, n. 8, p. 887–899, Aug. 2007.

WANG, B. et al. Characterization of an NBS-LRR resistance gene homologue from soybean. **Journal of Plant Physiology**, v.161, p. 815–822, 2004.

WANG, D. et al. Salicylic Acid Inhibits Pathogen Growth in Plants through Repression of the Auxin Signaling Pathway. **Current Biology**, v. 17, p. 1784-1790, 2007.

XIN, M. et al. Diverse set of microRNAs are responsive to powdery mildew infection and heat stress in wheat (*Triticum aestivum* L.). **BMC Plant Biology**, v. 10, n.123, 2010.

XU, Z. S. et al. Isolation and identification of two genes encoding leucine-rich repeat (LRR) proteins differentially responsive to pathogen attack and salt stress in tobacco. **Plant Science**, v. 176, p. 38-45, 2009.

YAMAMOTO, S. et al. Elicitor-induced activation of transcription via W box-related cis-acting elements from a basic chitinase gene by WRKY transcription factors in tobacco. **Biochimica et biophysica acta**, v. 1679, p. 279-287, 2004.

YANG, C-W. et al. The E3 Ubiquitin Ligase Activity of Arabidopsis PLANT U-BOX17 and Its Functional Tobacco Homolog ACRE276 Are Required for Cell Death and Defense. **The Plant Cell**, v. 18, p. 1084–1098, 2006.

YU, D.; CHEN, C.; CHEN, Z. Evidence for an Important Role of WRKY DNA Binding Proteins in the Regulation of NPR1 Gene Expression. **The Plant Cell**, v. 13, p. 1527-1539, 2001.

ZABALA, M. T.; BENNETT, M. H.; TRUMAN, W. H.; GRANT, M. R. Antagonism between salicylic and abscisic acid reflects early host–pathogen conflict and moulds plant defence responses. **The Plant Journal**, v. 59, p. 375–386, 2009.