



**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL**

**INSTITUTO DE BIOCÊNCIAS**

**PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E  
BIOLOGIA MOLECULAR**

**Análise funcional de genes que codificam proteínas WRKY, responsivos ao ataque do fungo *Phakopsora pachyrhizi* Sydow, agente causador da ferrugem asiática em soja [*Glycine max* (L.) Merrill]**

**Marina Borges Osorio**

**Porto Alegre, Março de 2010.**

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL**  
**INSTITUTO DE BIOCÊNCIAS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E**  
**BIOLOGIA MOLECULAR**

**Análise funcional de genes que codificam proteínas WRKY, responsivos ao ataque do fungo *Phakopsora pachyrhizi* Sydow, agente causador da ferrugem asiática em soja [*Glycine max* (L.) Merrill]**

**Marina Borges Osorio**

Dissertação submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS como requisito parcial para a obtenção do grau de Mestre em Ciências.

Orientadora: Dra. Márcia Margis-Pinheiro

Co-orientadora: Dra. Maria Helena Bodanese-Zanettini

Porto Alegre, Março de 2010.

## **INSTITUIÇÃO E FONTES FINANCIADORAS**

Este trabalho foi desenvolvido no Laboratório de Genética e Biologia Molecular Vegetal do Departamento de Genética da UFRGS e recebeu apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), assim como dos projetos GENOSOJA e BIOTECSUR.

## AGRADECIMENTOS

Às minhas queridas orientadoras, Márcia e Maria Helena: pela acolhida no Laboratório de Genética Vegetal da UFRGS, por terem me disponibilizado as melhores condições possíveis para a realização deste trabalho e principalmente, por servirem (sempre) como grandes exemplos, tamanha a competência e a dedicação com as quais exercem sua profissão.

À Bea: por no início, ter facilitado enormemente minha adaptação ao novo laboratório; e ao longo do mestrado, por ter sido como uma “terceira orientadora”, sempre prestativa e paciente, me dando dicas e sugerindo caminhos a seguir...

Aos colegas, professores e funcionários dos diversos núcleos que formam o Laboratório de Genética Vegetal, assim como do Laboratório de Genomas e Populações de Plantas (CBIOT), pela amizade e pela ajuda prestada ao longo destes dois anos...em especial ao João, por ter se mostrado sempre prestativo e atendido com boa vontade, sempre que uma “ajudinha” lhe era requisitada (e não foram poucas...).

À minha família e aos meus amigos pelotenses, que me deram o incentivo e o apoio fundamentais para meu sucesso nesta etapa...

Finalmente, a Deus, por ter me provido de paciência descomunal, sem a qual a escrita desta dissertação não seria possível.

Muito obrigada!

## RESUMO

A soja [*Glycine max* (L.) Merrill] é uma espécie da família Fabaceae que possui grande importância econômica mundial. A ferrugem asiática (ASR), causada pelo fungo *Phakopsora pachyrhizi* Sydow, tornou-se nos últimos anos uma das maiores ameaças às lavouras de soja ao redor do mundo. Cinco genes dominantes relacionados à resistência à ASR foram identificados. Entretanto, até o momento, nenhuma cultivar com resistência efetiva ao ataque pelo patógeno foi obtida, uma vez que todos os *loci* avaliados tiveram sua resistência quebrada por ao menos um isolado do fungo. A identificação de diversas fontes de resistência que possibilitem a construção de um “arsenal” contra a ASR no germoplasma comercial de soja, além do entendimento ao nível molecular dos mecanismos de defesa desta espécie contra *P. pachyrhizi*, podem contribuir significativamente no combate ao patógeno. Os fatores de transcrição WRKY são membros de uma superfamília com ampla distribuição no Reino Vegetal; embora suas funções regulatórias ainda não estejam bem definidas, inúmeros estudos realizados ao longo dos últimos anos têm reunido evidências de seu envolvimento nas respostas a ataques por patógenos. Estas requerem uma ampla reprogramação transcricional na célula vegetal, na qual fatores de transcrição WRKY parecem desempenhar um papel crucial no “ajuste-fino” da expressão de genes de defesa. O objetivo deste trabalho é caracterizar a função de dois genes da família WRKY em soja, envolvidos nas respostas ao ataque pelo fungo *P. pachyrhizi*, através de estratégias de genética reversa e estudos de expressão gênica. As seqüências genômicas e codificantes completas desses genes foram isoladas e identificadas respectivamente como *GmWRKY20* e *GmWRKY46*. Vetores para sua superexpressão e silenciamento em soja foram construídos e os processos de transformação genética desta espécie foram realizados por bombardeamento ou via sistema integrado “bombardeamento/*Agrobacterium*”, nos quais embriões somáticos foram utilizados como tecido alvo. Além disso, a análise do perfil de expressão desses genes foi realizada em diversos tecidos, fases do desenvolvimento, condições de estresse salino e em resposta à infecção por *P. pachyrhizi*.

## ABSTRACT

Soybean [*Glycine max* (L.) Merrill] is one of the most important legume crop worldwide. In the past few years, Asian Soybean Rust (ASR), caused by *Phakopsora pachyrhizi* Sydow, has become one of the major threats affecting the main soybean producers. Even though five dominant genes related to ASR resistance have been identified, there has been no report of a soybean variety presenting effectiveness towards the pathogen's attack, since the resistance ruled by every single one of these genes has already been overcome by at least one isolate around the world. The identification of resistance sources allowing the construction of an arsenal against ASR in commercial soybean germplasm, as well as the understanding of soybean's defense mechanisms against *P. pachyrhizi* at the molecular level, may therefore be extremely helpful regarding the pathogen's eradication. WRKY proteins belong to a superfamily of transcription factors with wide distribution amongst plants. Although their regulatory function is not yet clear, there have been several studies in the past few years raising evidences towards their involvement in pathogen's attack responses. These responses usually require a wide transcriptional reprogramming in the plant cell, at which WRKY proteins seem to play a central role fine-tuning the expression of defense related genes. The purpose of this study is to functionally characterize two soybean WRKY protein-encoding genes involved in soybean defense against *P. pachyrhizi*, by combining reverse genetics strategies with gene expression tools. Their genomic and complete coding sequences (cds) were isolated and the genes were identified as *GmWRKY20* and *GmWRKY46*, respectively. Besides, vectors for their silencing and overexpression in soybean were constructed and this plant species was genetically transformed by particle bombardment or by an integrated bombardment/*Agrobacterium* transformation system. Soybean somatic embryos were used as target tissue at both processes. In addition, an expression profile analysis of both *GmWRKY20* and *GmWRKY46* in several plant tissues and developmental stages as well as under salt stress and in response to *P. pachyrhizi* infection was accomplished.

## LISTA DE ABREVIATURAS

aa: aminoácido

ASR: ferrugem asiática (“Asian Soybean Rust”)

*At*: *Arabidopsis thaliana*

CaMV: vírus mosaico da couve-flor (“Cauliflower mosaic virus”)

CC-NB-LRR: “Coiled-Coil – Nucleotide Binding – Leucine-Rich Repeats”

cDNA: DNA complementar

CTAB: brometo de cetiltrimetilamônio

DNA: ácido desoxirribonucleico

D.O.: densidade óptica

dpi: dias após a inoculação

ESTs: etiquetas de seqüências expressas (“Expressed Sequence Tags”)

ET: etileno

GFP: proteína fluorescente verde (“Green Fluorescent Protein”)

*Gm*: *Glycine max*

hpi: horas após a inoculação

*hpt*: higromicina fosfotransferase

HR: resposta de hipersensibilidade (“Hypersensitive response”)

IPTG: “Isopropyl  $\beta$ -D-1-thiogalactopyranoside”

JA: ácido jasmônico

kb: quilobases

LB: meio de cultura Luria Broth

LRR-RLK: “Leucine-Rich Repeats – Receptor-like Kinase”

LRR-RLP: “Leucine-Rich Repeats – Receptor-like Protein”

MAPK: “Mithogen-activated Protein Kinase”

Mb: megabases

MKK7: “Mithogen-activated Protein Kinase Kinase 7”

MS: meio de cultura Murashinge e Skoog

Myr: milhões de anos atrás (“Million years ago”)

NCBI: “National Center for Biotechnology Information”  
NPR1: “Nonexpressor of PR genes 1”  
P35S: promotor 35S do CaMV  
pb: pares de base  
PCD: morte celular programada (“Programmed Cell Death”)  
PCR: Reação em Cadeia da Polimerase  
pH: potencial hidrogeniônico  
PIG: “Particle Inflow Gun”  
ProID: “rooting loci promoter” de *Agrobacterium rhizogenes*  
PRs: “Pathogenesis-Related proteins”  
pv: patovar  
RT-qPCR: PCR em Tempo Real quantitativo  
RLK: “Receptor-like Kinase”  
RNA: ácido ribonucleico  
RNAi: RNA de interferência  
ROS: espécies reativas de oxigênio  
rpm: rotações por minuto  
SA: ácido salicílico  
SAR: resistência sistêmica adquirida (“Systemic Acquired Resistance”)  
T35S: terminador 35S do CaMV  
T-DNA: DNA de transferência de *Agrobacterium tumefaciens*  
TIR-NB-LRR: “Toll/Interleukin1 Receptor – Nucleotide Binding – Leucine-Rich Repeats”  
TTSS: Sistema de Secreção Tipo III (“Type III Secretion System”)  
UTR: região não-traduzida  
UV: radiação ultravioleta  
WT : planta selvagem (“Wild-type”)  
X-Gal: “5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside”



## SUMÁRIO

1. <b>INTRODUÇÃO</b> .....	14
1.1. Estresses ambientais que afetam a cultura da soja e os mecanismos de resposta da planta .....	16
1.2. A genômica funcional como ferramenta para o melhoramento de plantas ....	23
1.3. A superfamília de fatores de transcrição WRKY .....	24
1.4. Transformação genética da soja e estudos funcionais .....	29
2. <b>OBJETIVOS</b> .....	32
2.1. Objetivo geral .....	32
2.2. Objetivos específicos.....	32
3. <b>METODOLOGIA</b> .....	33
3.1. Material.....	33
3.2. Seleção, análise <i>in silico</i> e isolamento dos genes de interesse .....	33
3.3. Obtenção de vetores para a transformação de soja.....	34
3.3.1. Construção de superexpressão .....	35
3.3.2. Construção de silenciamento .....	36
3.4. Transformação de <i>Escherichia coli</i> .....	38
3.5. Transformação de <i>Agrobacterium tumefaciens</i> .....	39

3.6.	Extração de DNA.....	40
3.7.	Extração de RNA total.....	40
3.8.	Síntese de cDNA .....	40
3.9.	Reação em cadeia da polimerase (PCR).....	41
3.10.	PCR em tempo real quantitativo (RT-qPCR) .....	42
3.11.	Análises da expressão gênica ao longo do desenvolvimento e em resposta a estresses.....	44
3.11.1.	Diferentes órgãos e fases do desenvolvimento .....	44
3.11.2.	Infecção por <i>P. pachyrhizi</i> .....	45
3.11.3.	Estresse salino .....	46
3.12.	Análises estatísticas .....	46
3.13.	Transformação genética de soja.....	46
3.13.1.	Indução de embriogênese somática e obtenção de tecido embriogênico proliferativo.....	47
3.13.2.	Transformação por biobalística (bombardeamento de partículas).....	47
3.13.3.	Transformação via sistema integrado “bombardeamento / <i>Agrobacterium</i> ” .....	48
3.13.4.	Seleção de transformantes, regeneração e aclimatação das plantas obtidas.....	48

3.13.5. Monitoramento da integração do transgene através da análise da expressão do gene repórter .....	51
<b>4. RESULTADOS .....</b>	<b>52</b>
4.1. Seleção, análise <i>in silico</i> e isolamento dos genes para estudo.....	52
4.2. Análises da expressão gênica .....	57
4.2.1. Diferentes órgãos e fases do desenvolvimento .....	57
4.2.2. Infecção por <i>P. pachyrhizi</i> .....	59
4.2.3. Estresse salino.....	61
4.3. Obtenção de vetores para a transformação de soja.....	63
4.3.1. Construção do vetor binário de superexpressão.....	63
4.3.2. Construção dos vetores de silenciamento .....	64
4.4. Obtenção de plantas transgênicas.....	67
4.4.1. Transformação de <i>Agrobacterium tumefaciens</i> com as construções de superexpressão e silenciamento.....	67
4.4.2. Transformação de embriões somáticos de soja via sistema integrado “bombardeamento / <i>Agrobacterium</i> ” .....	68
4.5. Análise dos tecidos geneticamente modificados e das plantas transgênicas ..	71
4.5.1. Monitoramento da integração do transgene através da análise da expressão do gene repórter .....	71

4.5.2.	Confirmação da integração do transgene por PCR.....	73
4.5.3.	Cálculo do número de cópias do transgene através de RT-qPCR .....	73
5.	<b>DISCUSSÃO</b> .....	75
5.1.	Análises <i>in silico</i> e isolamento dos genes em estudo .....	75
5.2.	Análises da expressão gênica .....	77
5.2.1.	Genes-referência .....	77
5.2.2.	Expressão em diferentes órgãos, fases do desenvolvimento e em resposta ao estresse salino .....	78
5.2.3.	Infecção por <i>P. pachyrhizi</i> .....	79
5.2.4.	Estrutura, expressão e função gênica .....	80
5.3.	Obtenção e análises de plantas transgênicas para o estudo funcional de <i>GmWRKY20</i> e <i>GmWRKY46</i> .....	84
5.3.1.	Expressão GFP.....	84
5.3.2.	Estimativa do número de cópias do transgene.....	88
5.3.3.	Revelação do fenótipo.....	90
6.	<b>CONCLUSÕES</b> .....	93
7.	<b>PERSPECTIVAS</b> .....	94
8.	<b>REFERÊNCIAS</b> .....	95

**ANEXO 1:** Sequências genômicas e codificantes de *GmWRKY20* e *GmWRKY46* e alinhamento dos “primers” utilizados nas construções RNAi e para as análises de expressão por RT-qPCR..... 102

**ANEXO 2:** Cálculo do número de cópias do transgene através da quantificação relativa pela análise da curva padrão..... 107

## 1. INTRODUÇÃO

Leguminosas são plantas pertencentes à família Fabaceae que, com aproximadamente 20.000 espécies, é a terceira maior família entre as plantas superiores. Dentre as leguminosas, a soja [*Glycine max* (L.) Merrill] é a espécie de maior importância econômica mundial. Segundo dados da FAO, a produção mundial de soja em 2008 ultrapassou 230 milhões de toneladas. Neste contexto, o Brasil aparece como o segundo maior produtor – atrás apenas dos Estados Unidos – tendo produzido, em 2008, quase 60 milhões de toneladas de grãos. A relevância desta espécie na agricultura é decorrente de sua capacidade de fixação de nitrogênio atmosférico através da simbiose com microrganismos; do seu alto teor protéico e lipídico, sendo utilizada há milênios na alimentação humana e animal; e mais recentemente, da possibilidade de utilização do óleo de soja para a produção de biocombustíveis, sendo esta utilidade de particular interesse do governo brasileiro (<http://www.biodiesel.gov.br/>).

Devido a sua grande importância, a comunidade científica internacional recomendou sua utilização como planta modelo entre as leguminosas para estudos genéticos e moleculares, na conferência “Cross-Legume Advances through Genomics (CATG)”, realizada em 2004, nos Estados Unidos (Gepts et al. 2005). Desde então, diversos esforços ao redor do mundo têm sido realizados visando a obtenção e a integração do maior volume de dados possíveis sobre o genoma da soja. A começar pelo seqüenciamento do genoma da cultivar Williams 82, financiado pelo Departamento de Energia dos Estados Unidos e realizado entre os anos de 2006 e 2008 através de “shotgun sequencing” (<http://www.jgi.doe.gov/CSP/>). Além disso, em 2007 foi criado o “International Soybean Genome Consortium” (ISGC) formado por pesquisadores da China, Japão, Coreia do Sul, Estados Unidos e Brasil. O Consórcio Nacional para Estudos do Genoma da Soja se integra ao ISGC através do projeto GENOSOJA, coordenado pelo pesquisador Ricardo Vilela Abdelnoor e financiado pelo CNPq.

O projeto GENOSOJA propôs o desenvolvimento de um consórcio entre diversos grupos de pesquisa do país para a caracterização do genoma da soja, tendo por objetivo a

integração de informações sobre a estrutura física do genoma e a expressão de genes e as proteínas codificadas por eles, dando ênfase às situações de estresses bióticos e abióticos que comprometem a produtividade da cultura no Brasil e no mundo. Entre os estresses alvos de estudo estão: a ocorrência de secas; doenças, como a ferrugem asiática; e pragas, como o ataque de diferentes espécies de nematóides. Ferramentas de bioinformática tornam possível a integração das informações geradas nos diferentes níveis do projeto – estrutural, transcricional, protéico e funcional – auxiliando na compreensão da função e dos mecanismos de controle da expressão de genes presentes na soja e envolvidos em processos de desenvolvimento e/ou defesa contra estresses ambientais.

Além do GENOSOJA, o Brasil também faz parte do projeto BIOTECSUR (Projeto de Cooperação Bi-Regional entre a União Européia & MERCOSUL, Referência EU 127119) que tem por objetivo a caracterização de genes e tecnologias derivadas de sua análise funcional, que possam agregar valor ao cultivo da soja frente aos estresses ambientais através da formação de uma rede de trabalho entre quatro países do MERCOSUL (Argentina, Brasil, Paraguai e Uruguai), representando um marco de sustentabilidade ambiental, social e econômica nestes países.

Recentemente, os primeiros resultados obtidos da análise do seqüenciamento do genoma da soja foram publicados (Schmutz et al. 2010). Compreendendo 950 Mb, o genoma da soja é o maior genoma vegetal seqüenciado por “shotgun” até hoje. Em termos comparativos, é duas vezes e meia maior que o genoma do arroz (389 Mb) e seis vezes maior que o genoma de *Arabidopsis thaliana* (157 Mb), espécie-modelo entre as dicotiledôneas. Dividido entre 20 cromossomos, 43% de suas seqüências se encontram em regiões eucromáticas, pouco repetitivas e com alta taxa de recombinação; entretanto 21,6% dos genes preditos com alta confiabilidade (“high confidence genes”) são encontrados nas regiões próximas aos centrômeros, ricas em seqüências repetitivas e transposons. O genoma da soja parece ter sofrido dois grandes eventos de duplicação (há 59 e 13 Myr) resultando em uma conformação atual com quase 75% dos genes presentes em múltiplas cópias, que vêm sendo extraordinariamente mantidas ao longo do tempo. Ressalta-se ainda, que dos

46.430 loci preditos como codificantes de proteínas, 12,2% foram identificados como sendo fatores de transcrição (em *Arabidopsis* este número é de apenas 7%). A distribuição geral destes entre as famílias de fatores de transcrição conhecidas é similar entre os genomas de soja e *Arabidopsis*, entretanto algumas famílias são mais esparsas e outras mais abundantes no genoma da soja, indicando que as vias regulatórias nesta espécie podem diferir daquelas descritas para *Arabidopsis*. Como referido anteriormente, a composição dos óleos produzidos em sementes de soja é de grande importância para sua utilização na indústria de biocombustíveis. Desta forma, chama-se atenção para a análise comparativa de genes que governam as vias de biossíntese e de acúmulo de lipídeos. O número de genes envolvidos nas diversas etapas destas vias aparenta ser bem maior em soja, sugerindo que estas sejam reguladas de maneira bem mais complexa nesta espécie, quando comparada a *Arabidopsis*.

### **1.1. Estresses ambientais que afetam a cultura da soja e os mecanismos de resposta da planta**

Os estresses ambientais são os grandes responsáveis por limitar um maior rendimento das culturas de soja. Estes podem ser de dois tipos: estresses abióticos (causados por alta salinidade e seca, principalmente) ou estresses bióticos (causados por pragas e doenças). Aproximadamente 40 doenças (bacterianas, fúngicas, causadas por vírus ou nematóides) que afetam as lavouras de soja já foram identificadas no Brasil; dentre elas, pode-se destacar o crestamento bacteriano da soja (*Pseudomonas savastanoi* pv. *glycinea*), o mosaico comum da soja (Vírus do Mosaico Comum da Soja, VMCS), o nematóide de cisto da soja (*Heterodera glycines*), a antracnose (*Colletotrichum dematium* var. *truncata*), a podridão de carvão (*Macrophomina phaseolina*) e a ferrugem asiática (*Phakopsora pachyrhizi*). A expansão da cultura para novas áreas, aliada a práticas inadequadas de manejo e à monocultura são os principais responsáveis pelo aumento de ocorrências e pela diversificação das doenças que afetam a soja (EMBRAPA Soja 2005).

A ferrugem asiática (ASR, “Asian Soybean Rust”) é uma doença causada pelo fungo *Phakopsora pachyrhizi* Sydow e foi identificada pela primeira vez no Japão, há mais de um século [Hennings VP (1903) A few new Japanese Uredinaceae. Hedwigia 42:107-

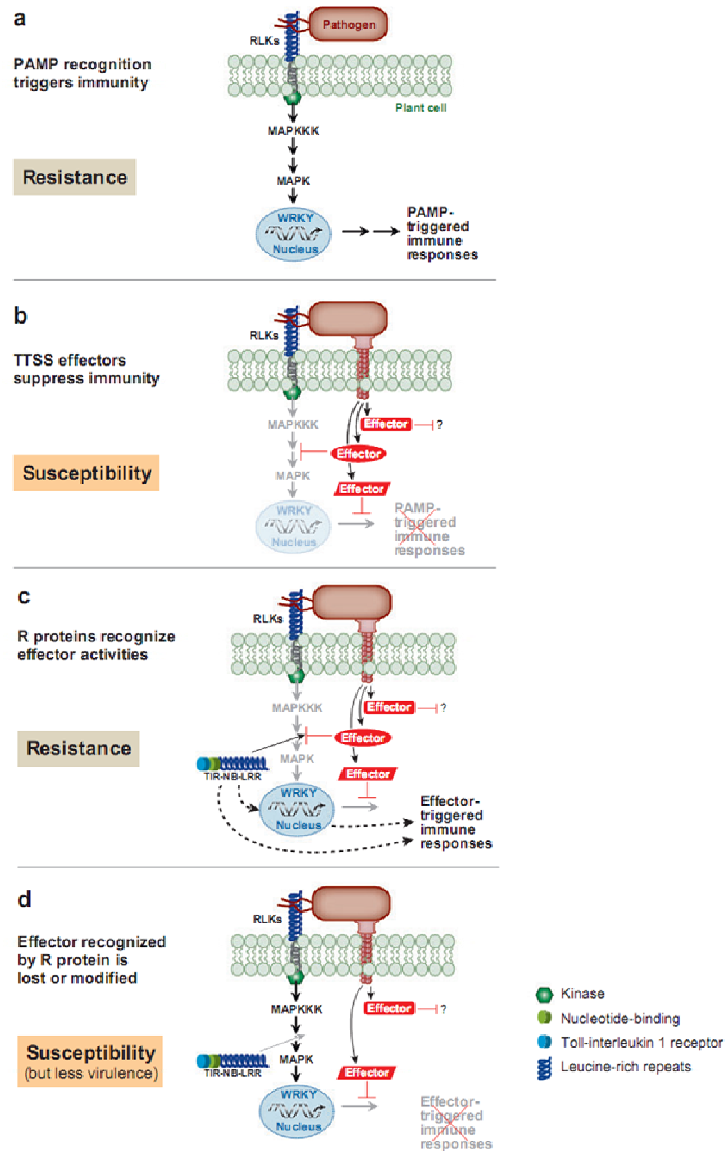


108, citado em Hartman, Miles & Frederick, 2005]. Entretanto, foi apenas a partir de 2001 que a doença começou a receber atenção por parte de agricultores e pesquisadores. Nesse ano foi descrito o primeiro caso de ASR em lavouras de soja do Paraguai, e alguns meses depois, no Brasil (Yorinori et al. 2005). Após sua dispersão pela América do Sul, ao final de 2004, a ASR chegou às lavouras dos Estados Unidos (Miles et al. 2006). Desde então, devido à sua alta severidade (perdas de até 80% na produção), se tornou uma das maiores ameaças às lavouras de soja ao redor do mundo. Estimativas apontam que a doença tenha causado perdas econômicas de mais de 2,2 bilhões de dólares na safra 2006/2007 no Brasil (Calvo et al., 2008). Além disso, a “USDA’s Risk Management Agency” (RMA), já há alguns anos põe em alerta os agricultores norte-americanos sobre o risco de contaminação de suas lavouras e recomenda práticas preventivas e de controle de uma possível epidemia (<http://www.rma.usda.gov/news/2006/04/soybeanrust.html>). O desenvolvimento da infecção resulta na desfolhação prematura e redução drástica no rendimento dos grãos. Atualmente, por não haver disponibilidade de cultivares comerciais com resistência genética durável, a principal forma de controle é através do uso de fungicidas. Estes, além de aumentarem substancialmente os custos da produção e serem uma fonte constante de contaminação ambiental, geralmente acabam sendo ineficientes, devido à agressividade do fungo e à dificuldade de detecção do mesmo nos estágios iniciais da infecção (Calvo et al., 2008; Silva et al., 2008).

O ciclo infeccioso de um fitopatógeno inicia-se pela sua penetração no tecido vegetal. As plantas dispõem de uma imunidade “pré-invasiva”, na qual lançam mão de um conjunto de barreiras físicas pré-existentes (cutícula, tricomas e parede celular) ou induzidas (deposição de calose, fechamento dos estômatos); além de barreiras químicas (liberação de quitinases, glucanases e compostos antimicrobianos) a fim de prevenir a invasão de microrganismos. Ao sobrepor estas barreiras, o patógeno é então exposto aos mecanismos de defesa da “imunidade pós-invasiva”, a qual pode ser ativada por PAMPs (“Pathogen-associated Molecular Patterns”) – sendo neste caso denominada PTI (“PAMP-triggered Immunity”) – ou ainda, ser ativada por efetores e denominada ETI (“Effector-triggered Immunity”)(Göhre et al. 2008). A PTI geralmente está associada à imunidade

não-hospedeira e à imunidade basal, que são respostas imediatas ao contato com o patógeno. Já a ETI está associada ao desenvolvimento de uma resposta de hipersensibilidade (HR) e à resistência sistêmica adquirida (SAR), sendo estas respostas de longa duração.

Juntas, a PTI e a ETI fazem parte do modelo atualmente conhecido como o “Dogma Central” da fitopatologia (Figura 1). Este importante modelo descreve um processo de co-evolução popularmente denominado “queda-de-braço evolucionária” (“evolutionary arms-race”), resumido a seguir. Além das barreiras físicas e químicas pré-existentes, as plantas possuem a capacidade de detectar componentes genéricos conservados dos microrganismos (MAMPs, “Microbe-associated Molecular Patterns”; ou PAMPs) desencadeando a PTI. Por sua vez, alguns microrganismos desenvolvem fatores de virulência (efetores) capazes de suprimir parte da resposta geral de defesa de seu hospedeiro. Estes acabaram por desenvolver genes de resistência (“R genes”) específicos, cujo produto detecta de forma direta ou indireta os fatores de virulência do patógeno, desencadeando a ETI. Finalmente, através da eliminação ou da modificação de seus efetores, o patógeno pode mais uma vez escapar à detecção pela planta (Bent et al. 2007).



**Figura 1. Modelo evolutivo da interação planta-patógeno.** O modelo apresentado utiliza a interação planta-bactéria como exemplo geral de interação planta-patógeno: (a) O reconhecimento de PAMPs (como a flagelina bacteriana) por receptores extracelulares RLKs prontamente ativa a imunidade basal, a qual inclui a sinalização por cascatas de MAPKs e a reprogramação transcricional mediada por fatores de transcrição WRKY. (b) Bactérias patogênicas utilizam o TTSS para secretar múltiplas proteínas efetoras que suprimem as respostas da imunidade basal do hospedeiro, permitindo a multiplicação do patógeno no apoplasto vegetal. (c) Proteínas de resistência vegetais (produtos de genes *R*, como proteínas TIR-NB-LRR) reconhecem a atividade de efetores e restabelecem a resistência da planta, desenvolvendo forte ETI. (d) O

patógeno evita a defesa mediada por genes *R* modificando ou eliminando os efetores que ativam esta defesa. Este estado lembra o apresentado em (b), exceto que neste caso o patógeno teve que perder ou alterar uma proteína efetora, ou ainda desenvolver uma nova. (Retirado de Bent & Mackey, 2007).

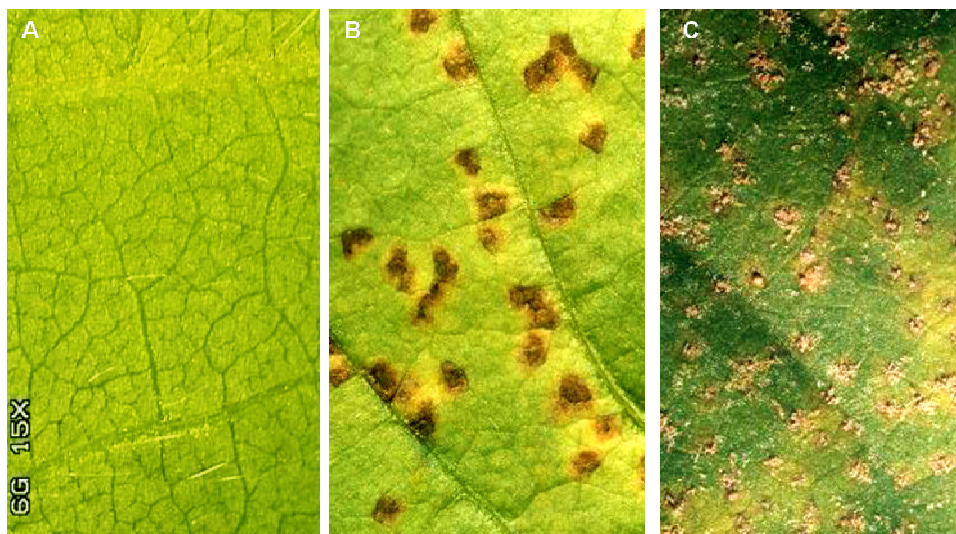
As estruturas ou moléculas conservadas pertencentes aos microrganismos que servem como alvo do reconhecimento de receptores eucarióticos são chamadas, de maneira geral, de elicitores. Frequentemente refere-se a estes através das siglas PAMPs ou MAMPs, previamente descritas. Flagelinas bacterianas, lipopolissacarídeos, quitinas fúngicas ou heptaglicosídeos de oomicetos são exemplos de MAMPs detectados pelas células vegetais; como estes estão geralmente envolvidos em processos biológicos básicos do microrganismo, sua evolução adaptativa a fim de evitar o reconhecimento pela planta é um processo muito lento. PRRs (“Pattern Recognition Receptors”) são os receptores presentes na membrana plasmática vegetal que estimulam cascatas de transdução de sinal ao reconhecerem MAMPs, induzindo respostas de defesa no núcleo. Dois tipos de PRRs são os mais conhecidos, LRR-RLKs ou LRR-RLPs e as respostas de defesa desencadeadas por estes geralmente envolvem: o fluxo de íons de  $Ca^{2+}$ , a formação de espécies reativas de oxigênio (ROS) e de óxido nítrico (NO) e a ativação de cascatas de transdução de sinal até o núcleo (envolvendo MAPKs e fatores de transcrição WRKY); resultando na deposição de calose, no fechamento dos estômatos e na síntese de compostos antimicrobianos (PRs), por exemplo. Os efetores liberados por patógenos, visando suprimir a PTI, são moléculas relacionadas à sua virulência que estão fortemente sujeitos à evolução adaptativa, por não desempenharem uma função fisiológica essencial nestes organismos. Se bem sucedidos, promovem a ocorrência de uma interação denominada “compatível” entre o patógeno e a planta, resultando no desenvolvimento do processo infeccioso. São também conhecidos como genes de avirulência (“*Avr* genes”), pois uma vez reconhecidos por receptores do hospedeiro nas interações incompatíveis, sua função de virulência é ofuscada por uma função dominante de avirulência e a infecção é evitada. Estes efetores geralmente são secretados no apoplasto ou no citoplasma vegetal, através do sistema de secreção tipo três (TTSS) em bactérias, ou por exocitose, no caso de fungos e oomicetos. Por fim, a ETI é desencadeada na célula vegetal quando o produto de “genes R” – proteínas receptoras de

membrana (do tipo LRR-RLKs ou LRR-RLPs), ou ainda intracelulares (CC-NB-LRR ou TIR-NB-LRR) – reconhece estas moléculas efetoras ou o produto de sua ação. As respostas de defesa decorrentes deste reconhecimento incluem, como referido anteriormente para a PTI, a geração de uma explosão oxidativa na célula, o fluxo de íons através da membrana plasmática e a expressão de genes de defesa no núcleo, ativados através de cascatas de transdução de sinal. Diferentemente da PTI, na ETI geralmente ocorre o desenvolvimento de uma HR, podendo esta eventualmente resultar em morte celular programada (PCD). Estas e outras evidências apontam atualmente para a ETI como sendo uma reativação da PTI, de maneira acelerada e potencializada (Bent & Mackey, 2007; Hükelhoven, 2007; Göhre & Robatzek, 2008).

O ciclo típico da infecção por *P. pachyrhizi* começa pela germinação dos urediniósporos, entre 1 e 2 horas após a inoculação (hpi) sob condições adequadas: escuro, alta umidade e temperatura permissiva, entre 15°C e 25°C (Tsukahara et al. 2008). Duas horas após a germinação é formado o apressório. Dentro deste é formado um cone, a partir do qual, 7 hpi uma hifa de penetração atravessa a epiderme do tecido vegetal. Quando esta atinge o espaço intercelular abaixo da epiderme, um septo é formado produzindo hifas primárias (entre 15 e 20 hpi). Os haustórios, órgãos especializados posicionados entre a parede celular e a membrana plasmática vegetal, pelos quais o fungo obtém nutrientes e secreta proteínas efetoras, são formados de 24 a 48 hpi. Neste mesmo período o fungo segue colonizando o mesófilo esponjoso, através da formação de hifas secundárias e haustórios adicionais. Novos urediniósporos são formados de 7 a 9 dias após a inoculação (dpi), sendo disseminados por até quatro semanas (van De Mortel et al. 2007).

Até hoje, cinco genes dominantes relacionados à resistência à ASR, “*Resistance to Phakopsora pachyrhizi genes*”, foram identificados em soja: *Rpp1*, *Rpp2*, *Rpp3*, *Rpp4* e *Rpp5*, sendo que cada um desses genes condiciona resistência a um grupo limitado de isolados de *P. pachyrhizi*. Três reações em resposta ao patógeno são possíveis (Figura 2). A reação imune (governada por *Rpp1*) na qual a planta não apresenta sintomas visíveis; a reação de hipersensibilidade (HR, governada por *Rpp2* a *Rpp5*) que possui como

característica a formação de lesões avermelhadas (“reddish-brown lesions”, RB) decorrentes de morte celular programada e que resultam na limitação da esporulação e do crescimento do fungo; e a reação suscetível, caracterizada por lesões marrons (“tan-lesions”, TAN) decorrentes da esporulação total das pústulas de *P. pachyrhizi*. Como já mencionado, até o momento a utilização de nenhuma destas fontes monogênicas gerou resistência durável em cultivares comerciais de soja, uma vez que todos os *loci* avaliados tiveram sua resistência quebrada por ao menos um isolado (Miles, Hartman, & Fredrick, 2005; Silva et al., 2008; Meyer et al., 2009). Além do mais, a baixa variabilidade genética apresentada entre as cultivares de soja existentes, tanto no Brasil quanto nos Estados Unidos, é um fator agravante na busca por novas fontes de resistência. Em 2006, Miles, Frederick & Hartman testaram todo o banco de germoplasma dos Estados Unidos contra cinco isolados de *P. pachyrhizi* e níveis de tolerância (ao menos parcial) foram encontrados em menos de 5 % da coleção disponível. A identificação de diversas fontes de resistência que possibilitem a construção de um “arsenal” contra a ASR no germoplasma comercial de soja, além do entendimento ao nível molecular dos mecanismos de defesa da soja contra este fungo, podem contribuir enormemente no combate ao patógeno.



**Figura 2. Reações da soja em resposta à ASR.** (A) reação imune, sem sintomas visíveis; (B) reação de hipersensibilidade, com lesões avermelhadas (RB); e (C) reação suscetível, com lesões marrons (TAN) (Modificado de Miles, Frederick, & Hartman, 2006).

## 1.2. A genômica funcional como ferramenta para o melhoramento de plantas

A genômica funcional aparece como uma importante ferramenta biotecnológica para se incrementar as bases genéticas utilizadas nos programas de melhoramento. A utilização combinada de técnicas de genética reversa, como a superexpressão de genes de interesse ou o silenciamento gênico por RNAi, aliada a análises de “high throughput”, como o microarranjo e o MPSS (“massively parallel signature sequencing”), têm feito da genômica funcional uma peça chave na era “pós-genômica”, de forma a possibilitar a identificação de genes e rotas metabólicas relacionadas aos mais variados processos biológicos, como o desenvolvimento de órgãos, ciclo celular e resposta a estresses ambientais, além de elementos-chave na regulação destas rotas (Waterhouse & Helliwell, 2003; Gutterson & Zhang, 2004).

Recentemente, três grupos independentes demonstraram – através de técnicas de genômica funcional “high throughput” – a modulação da expressão de vários genes de soja em resposta ao ataque do fungo *P. pachyrhizi* (Panthee et al., 2007; van De Mortel et al.,

2007; Choi et al., 2008; Panthee et al., 2009). O primeiro perfil do transcriptoma desta interação foi um microarranjo realizado por Panthee et al. em 2007, onde foram analisadas plantas jovens (estágio V2), 72 hpi com o patógeno. Em 2009, o mesmo grupo apresentou outra análise, desta vez comparando o perfil de expressão em duas fases distintas do desenvolvimento da soja (estágios V4 e R1), 72 hpi. O foco destes estudos foi demonstrar que a expressão diferencial de genes em resposta à infecção depende do estágio de desenvolvimento. Em 2008, Choi et al. obtiveram o perfil do transcriptoma através de construção de bibliotecas de SSH (“suppressive subtraction hybridization”) seguida de microarranjo, comparando a reação suscetível com a imune (governada pelo gene *Rpp1*) 1, 6, 12, 24 e 48 hpi. Neste trabalho foi ressaltada a importância, tanto de genes responsáveis pelo balanço oxidativo (lipoxigenases e peroxidases); quanto de membros de diversas famílias de fatores de transcrição, sugerindo um complexo padrão de regulação na célula vegetal para que a infecção seja evitada. Já no trabalho realizado por van De Mortel et al., em 2007, a reação suscetível foi comparada à reação de hipersensibilidade (governada pelo gene *Rpp2*), sendo as análises realizadas 6, 12, 18, 24, 36, 48, 72, 96, 120 e 168 hpi. Através de microarranjo pôde-se observar um grande número de genes diferencialmente expressos, ressaltando em especial a expressão diferencial de 46 fatores de transcrição pertencentes à família WRKY, indicando que estes ocupam papel-chave na resposta da soja à infecção pelo fungo causador da ferrugem asiática.

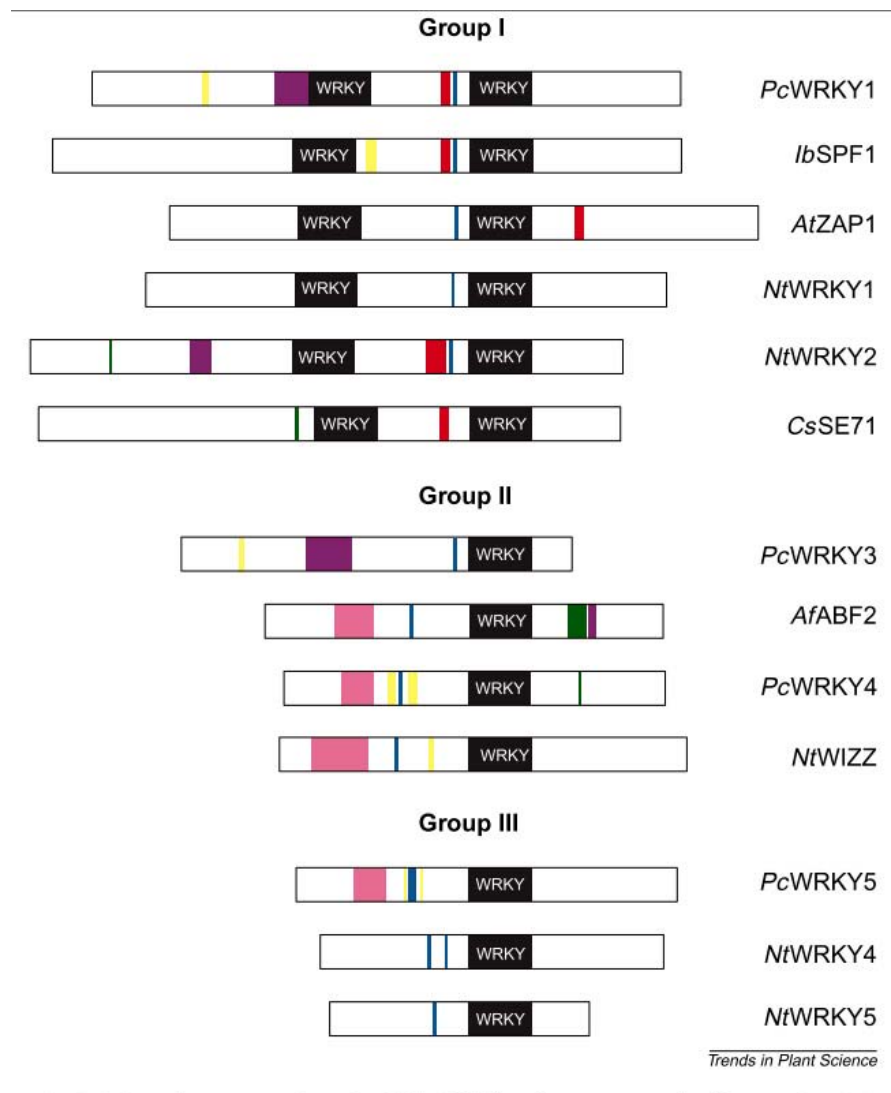
### **1.3. A superfamília de fatores de transcrição WRKY**

Os fatores de transcrição WRKY são membros de uma superfamília com ampla distribuição no Reino Vegetal, o que pode refletir sua importância na história evolutiva destes organismos (Eulgem et al., 2000). Membros desta família já foram identificados em diversos táxons vegetais: desde briófitas (*Physcomitrella patens* – 38 membros) e gimnospermas (*Pinus monticola* – 83 membros); até angiospermas monocotiledôneas (*Oryza sativa* – 105 membros) e dicotiledôneas (*Arabidopsis thaliana* – 74 membros, *Populus trichocarpa* – 100 membros) (Ulker & Somssich, 2004; Liu & Ekramoddoullah, 2009). O aparente sucesso desta superfamília no Reino Vegetal parece ser decorrente dos



sucessivos eventos de duplicação gênica sofridos pelos genomas das plantas, sendo que um grande número de genes WRKY duplicados foi mantido ao longo da evolução, tanto em espécies selvagens quanto nas cultivadas (Petitot, Lecouls, & Fernandez, 2008; Tao et al. 2009). Apesar disso, membros desta família foram recentemente identificados em espécies de outros reinos, tais como *Giardia lamblia* (protista), *Dictyostelium discoideum* (mixomicota) e *Chlamydomonas reinhardtii* (alga unicelular); sugerindo que possuam origem anterior ao surgimento das plantas, tendo sido perdidos em algum momento ao longo da evolução de fungos e animais – grupos nos quais não há registros de sua presença (Ülker et al. 2004).

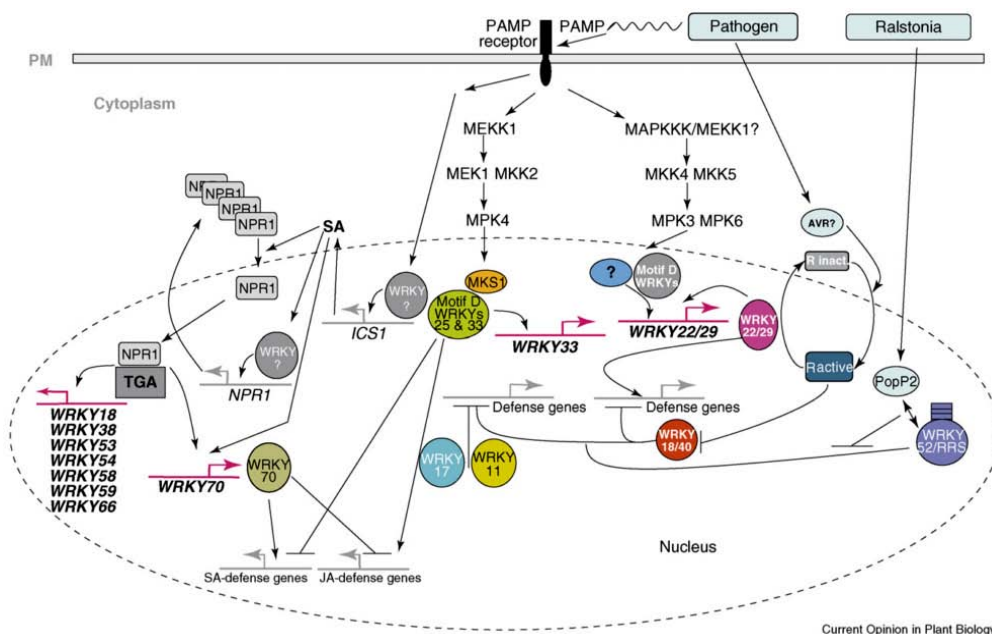
A principal característica compartilhada pelos membros desta família é seu domínio de ligação ao DNA, denominado “domínio WRKY”: tipicamente composto por 60 aminoácidos, dentre estes a seqüência conservada WRKYGQK (triptofano – arginina – lisina – tirosina – glicina – glutamina – lisina) na sua porção N-terminal, além de um motivo “zinc-finger-like” na sua porção C-terminal (Eulgem et al., 2000). O domínio WRKY possui como alvo principal de ligação, cis-elementos do tipo “W-boxes” (C/TTGACT/C), apesar de sítios alternativos terem sido identificados recentemente, o que reflete a grande variação apresentada entre os aminoácidos adjacentes ao motivo conservado WRKYGQK (Pandey et al. 2009). Baseado no número de domínios WRKY e nas características apresentadas por seus motivos “zinc-finger-like”, Eulgem et al., em 2000, dividiram os fatores de transcrição WRKY em três subgrupos (Figura 3): proteínas do grupo I apresentam dois domínios WRKY e potenciais ligantes a zinco seguindo o padrão C<sub>2</sub>-H<sub>2</sub> (C-X<sub>4-5</sub>-C-X<sub>22-23</sub>-H-X<sub>1</sub>-H) em seus motivos “zinc-finger-like”; no grupo II o padrão C<sub>2</sub>-H<sub>2</sub> permanece, mas estas proteínas possuem apenas um domínio WRKY; já no grupo III, o padrão C<sub>2</sub>-HC (C-X<sub>7</sub>-C-X<sub>23</sub>-H-X<sub>1</sub>-C) é observado e proteínas deste grupo também apresentam apenas um domínio WRKY.



**Figura 3. Representação esquemática dos domínios estruturais das proteínas WRKY de salsa (*Pc*), batata doce (*Ib*), *Arabidopsis* (*At*), tabaco (*Nt*), pepino (*Cs*) e aveia selvagem (*Af*). Elas são divididas em três grupos de acordo com o número e o tipo de domínio WRKY que contêm. Domínios WRKY estão representados em preto, prováveis sinais básicos de localização nuclear em azul e zíperes de leucina em rosa. Regiões ricas em serina-treonina estão em amarelo, regiões ricas em glutamina em roxo, regiões ricas em prolina em verde e regiões acídicas em vermelho (Retirado de Eugelm et al., 2000).**

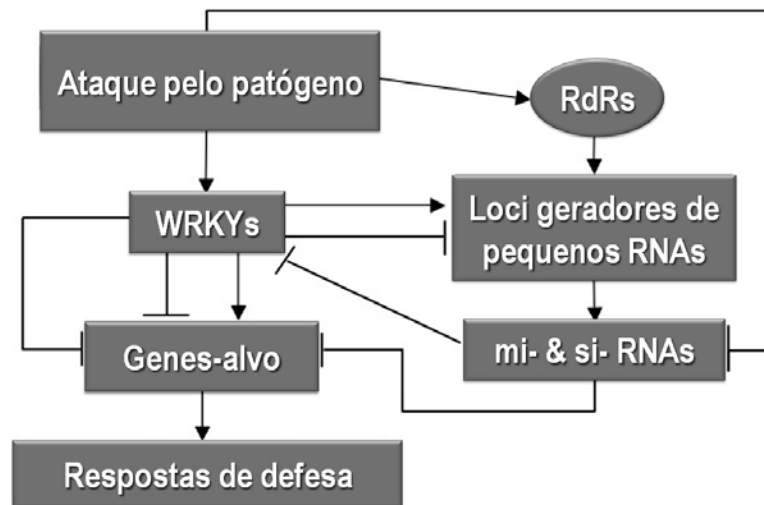
Embora as funções regulatórias destes genes ainda não estejam bem definidas, inúmeros estudos realizados ao longo dos últimos anos têm demonstrado evidências de sua participação nos processos de senescência, desenvolvimento de tricomas e metabolismo, além da resposta a estresses abióticos (seca, frio, irradiação por luz UV) e principalmente, em resposta aos mais diversos tipos de estresses bióticos (Berri et al., 2009; (Pandey et al. 2009).

Respostas a ataques por patógenos requerem uma ampla reprogramação transcricional e estudos sugerem que a maioria dos patógenos desencadeia uma rede de sinalização interconectada na célula vegetal envolvendo o balanço entre os hormônios ácido salicílico (SA) e ácido jasmônico (JA). Esta rede intrincada compreende ainda, a ação de ativadores e repressores transcricionais que realizam o ajuste fino da expressão de genes de defesa. Grandes famílias de fatores de transcrição são responsáveis por esta modulação; em particular, a presença de cis-elementos do tipo “W-boxes” na região promotora de genes co-regulados em respostas de defesa, indica que fatores de transcrição WRKY desempenham um papel amplo e crucial na resposta imune das plantas (Figura 4)(Kalde et al., 2003; Eulgem & Somssich, 2007; Berri et al., 2009; Pandey & Somssich, 2009). Além do mais, a presença destes cis-elementos em suas próprias regiões promotoras indicam a existência de um sistema de auto-regulação/regulação cruzada entre WRKYs através de mecanismos de “feedback” específicos (Pandey & Somssich, 2009). Por fim, estudos recentes (Navarro et al., 2008; Zhang, et al. 2008; Kuang, Padmanabhan, & Li, 2009) demonstraram que pequenos RNAs endógenos de plantas estão amplamente envolvidos na regulação pós-transcricional de genes envolvidos nas respostas a patógenos, apontando ainda, para a existência de um “interatoma” entre WRKYs e pequenos RNAs (Figura 5), uma vez que WRKYs são alvos preditos de diversos miRNAs, assim como estes também apresentam “W-boxes” em sua região promotora.



Current Opinion in Plant Biology

**Figura 4. Modelo hipotético da rede de interação entre WRKYs (de *A. thaliana*) envolvidos nas respostas ao ataque por patógenos.** A sinalização de defesa celular pode ser desencadeada pelo reconhecimento de PAMPs por receptores de membrana e subsequente ativação de cascatas de fosforilações envolvendo MAPKs (PTI), ou através da detecção de produtos de efetores do patógeno na célula vegetal por proteínas R (ETI). Em ambos os casos, rápidas alterações da expressão gênica subsequentes são mediadas pela ação de diversos fatores de transcrição, como WRKYs. A ETI pode ser desencadeada pela ativação de proteínas R (R inativa → R ativa) mediada por efetores e subsequente inibição de WRKYs supressores da defesa. A sinalização por SA ativada pelo patógeno libera NPR1 de complexos oligoméricos, resultando no acúmulo de monômeros de NPR1 no núcleo e sua associação com fatores de transcrição TGA em sítios promotores, sendo que conjunto de genes WRKY depende da ação de NPR1, tanto de forma positiva quanto negativa (Modificado de Eulgem & Somssich, 2007).



**Figura 5. Modelo do interatoma WRKYs - pequenos RNAs durante a reprogramação das respostas de defesa vegetal.** Durante o ataque do patógeno, *loci* geradores de pequenos RNAs podem estar sob o controle de fatores de transcrição WRKY; ao mesmo tempo, a abundância de WRKYs pode ser regulada por pequenos RNAs. RdRs, “RNA polimerases direcionadas à RNAs” (Modificado de Pandey & Somssich, 2009).

#### 1.4. Transformação genética da soja e estudos funcionais

A importância econômica da soja, aliada à sua baixa variabilidade genética e à incompatibilidade sexual desta espécie em cruzamentos interespecíficos e intergenéricos (Hu et al. 1995) têm feito da transformação genética de soja, não só uma importante alternativa às práticas de melhoramento convencional; mas também, servido como ferramenta a estas práticas, de forma a possibilitar o incremento das bases genéticas utilizadas nos programas de melhoramento. Hoje em dia, a transformação genética vegetal (tanto de forma estável quanto transiente), é pré-requisito para a realização de estudos moleculares, genéticos, bioquímicos e fisiológicos (Somers et al. 2003) que sejam realizados, por exemplo, a partir da superexpressão, do silenciamento ou da localização sub-celular dos genes em estudo. Entretanto, a existência de um sistema eficiente de transformação é essencial para que estes sejam bem sucedidos.

Dessa forma, diversos trabalhos têm sido desenvolvidos – principalmente em *Arabidopsis thaliana* – que além de ser uma espécie de fácil transformação (método “floral dip”; Clough & Bent, 1999) e rápido desenvolvimento, possui genoma de tamanho reduzido que acaba por facilitar as análises moleculares subsequentes. Mais especificamente, no estudo de genes envolvidos nos mecanismos de interação entre planta / patógeno pode-se citar diversos trabalhos bem sucedidos realizados através de técnicas de superexpressão e silenciamento. Navarro et al., em 2006, demonstraram pela primeira vez o envolvimento de um microRNA endógeno de plantas (miR393 de *Arabidopsis*) na via de defesa antibacteriana ativada por PAMPs (flg22 de *Pseudomonas syringae* pv. *tomato* DC3000). Quando induzido, o microRNA em questão atuaria inibindo a expressão de genes relacionados à sinalização de auxinas na célula vegetal. Esta via parece funcionar de forma antagônica à via de defesa contra patógenos biotróficos, mediada pelo ácido salicílico, e sua inibição resultou na restrição do crescimento bacteriano. Da mesma forma, Zhang et al., em 2007 demonstraram que a MKK7 de *Arabidopsis* está envolvida na regulação negativa do transporte polar de auxinas na célula. Além disso, a superexpressão de MKK7 aumentou a resistência aos patógenos *Pseudomonas syringae* pv. *maculicola* ES4326 e *Hyaloperonospora parasitica* Noco2 e seu silenciamento resultou no comprometimento, tanto da resistência basal quanto da resistência sistêmica adquirida (SAR) desta espécie (Zhang et al. 2007). Além disso, o papel regulatório dos próprios fatores de transcrição WRKY tem sido desvelado através de estudos funcionais utilizando plantas superexpressando, silenciadas ou mutantes para um ou mais genes WRKY. Em 2006, Xu et al. demonstraram haver uma interação física (formação de homo- e heterocomplexos) e funcional (redundante ou antagônica) entre *AtWRKY18*, *AtWRKY40* e *AtWRKY60* na resposta de *Arabidopsis* a diferentes patógenos. No mesmo ano, Wang, Amornsiripanitch & Dong demonstraram que 8 fatores de transcrição WRKY estavam entre os alvos diretos de NPR1 – que é um co-fator transcricional da via mediada por SA no desenvolvimento da SAR – e que estes atuam tanto de forma positiva quanto negativa na regulação da SAR.

A existência de um sistema eficiente de transformação é fundamental para a realização de estudos funcionais em uma espécie. No caso da soja, cujos processos de transformação genética se mostram bem mais complexos, não existem muitos relatos de estudos funcionais que se utilizem da transformação desta espécie. Em nosso laboratório o protocolo de transformação de soja via bombardeamento foi estabelecido por Droste, Pasquali & Bodanese-Zanettini em 2002. Além disto, foi desenvolvido um sistema de transformação de soja que combina a técnica de bombardeamento de partículas à transformação via *A. tumefaciens* (Droste et al., 2000). Em ambos os sistemas, embriões somáticos secundários são utilizados como tecido-alvo, sendo este bastante promissor como ferramenta para a realização de estudos funcionais nesta espécie.

Até o presente momento, diversas pesquisas em sistemas heterólogos de transformação (como é o caso de *Arabidopsis*) vêm gerando informações relacionadas à funcionalidade de genes de soja. Dentre elas pode-se citar o trabalho realizado por Zhou et al. em 2008, relacionando a função de *GmWRKY13*, *GmWRKY21* e *GmWRKY54* à tolerância a estresses abióticos através de sua superexpressão em *Arabidopsis*; ou o trabalho de Zhang et al., publicado no mesmo ano, que demonstrou que a superexpressão de *GmWRKY57* em tabaco conferiu tolerância à seca nas plantas transgênicas desta espécie. Como exceção, pode-se citar o trabalho desenvolvido por Mazarei et al. em 2007, no qual a participação de EREBPs (“Ethylene-responsive Element-binding Proteins”) nas vias de regulação mediadas por ET e JA foi demonstrada a partir da expressão destes fatores de transcrição em soja transformada geneticamente. Até o momento, nenhum trabalho envolvendo o estudo funcional de fatores de transcrição WRKY de soja através da transformação genética desta espécie foi publicado.

## **2. OBJETIVOS**

### **2.1. Objetivo geral**

O presente trabalho pretende caracterizar e determinar a função de genes que codificam proteínas WRKY em soja, relacionados com os processos de resposta ao ataque do fungo *P. pachyrhizi*, agente causador da ferrugem asiática, através de estratégias de genética reversa e estudos de expressão gênica.

### **2.2. Objetivos específicos**

- 1º. Isolar as seqüências genômicas e codificantes dos genes selecionados;
- 2º. Determinar o padrão de expressão desses genes em plantas de soja, através de PCR em tempo real (RT-qPCR);
- 3º. Construir vetores de transformação vegetal contendo os genes clonados, tanto para superexpressão quanto para o silenciamento (via RNAi) da expressão dos mesmos;
- 4º. Obter linhagens transgênicas de soja com as construções descritas no item c;
- 5º. Caracterizar em nível molecular, através de PCR e RT-qPCR, as linhagens de soja estavelmente transformadas.



### 3. METODOLOGIA

#### 3.1. Material

As cultivares IAS 5, BRSMG 68 (Vencedora), e MGBR 46 (Conquista) de soja [*Glycine max* (L.) Merrill] foram utilizadas para a obtenção de DNA, RNA e nos experimentos de transformação genética. Para a multiplicação dos vetores de clonagem e de transformação de plantas foram utilizadas as cepas de *Escherichia coli* X-L1 Blue (Stratagene) e Top10 (Invitrogen). A cepa de *Agrobacterium tumefaciens* LBA4404 foi utilizada no co-cultivo com a soja no processo de transformação genética vegetal.

#### 3.2. Seleção, análise *in silico* e isolamento dos genes de interesse

A seleção dos genes de estudo, envolvidos na resposta da soja à infecção pelo fungo *P. pachyrhizi*, ocorreu a partir de análise de microarranjo realizada por van De Mortel et al. em 2007, na qual genes codificando 46 fatores de transcrição da família WRKY apresentaram níveis alterados de expressão após a inoculação de folhas de soja com o referido patógeno. A seleção foi realizada com base nos seguintes critérios:

- 1º. Apresentar perfil de expressão diferencial ao longo do período analisado;
- 2º. Apresentar perfil de expressão diferencial entre a cultivar suscetível (Embrapa-48) e a linhagem resistente (PI970230) utilizadas no estudo;
- 3º. Apresentar perfil de expressão diferencial entre si.

Além desses, a disponibilidade destas seqüências (ao menos parciais) no Genbank também foi fator determinante para a seleção dos genes em estudo. Isto porque, no momento de sua seleção (início de 2008), os resultados obtidos no sequenciamento do genoma da soja ainda estavam sendo montados (esta montagem só foi finalizada no início de 2009), de forma que as informações obtidas no banco de dados onde este foi depositado (Phytozome - <http://www.phytozome.org/soybean>) se mostravam ainda incompletas.

Através dos códigos de acesso obtidos no microarranjo (Gene Chip<sup>®</sup> Soybean Genome Array, Affimetrix) foi realizada a análise *in silico* dos genes selecionados. Suas ESTs e seqüências codificantes parciais foram obtidas no NCBI GenBank (<http://www.ncbi.nlm.nih.gov/Genbank/>) e suas seqüências genômicas foram obtidas no Phytozome. Estas últimas eram verificadas à medida que atualizações no banco de dados eram realizadas.

O isolamento das seqüências genômica e codificante dos genes selecionados foi realizado através do desenho de oligonucleotídeos iniciadores (“primers”) específicos para cada gene, seguido de clonagem em vetor pGEM<sup>®</sup>-TEasy (Promega) seguindo o protocolo fornecido pelo fabricante. Os vetores obtidos da clonagem dos fragmentos isolados em pGEM<sup>®</sup>-TEasy foram inseridos em *Escherichia coli* XL-1Blue por choque térmico, como descrito no item 3.4. A fim de obter grande quantidade de cada produto clonado, colônias recombinantes foram cultivadas em meio LB líquido contendo 50 µg.ml<sup>-1</sup> de ampicilina e submetidas à extração de plasmídeo (Wizard<sup>®</sup> Plus SV Minipreps, Promega), de acordo com o manual do fabricante. Os plasmídeos purificados foram enviados para seqüenciamento, para a confirmação da identidade gênica.

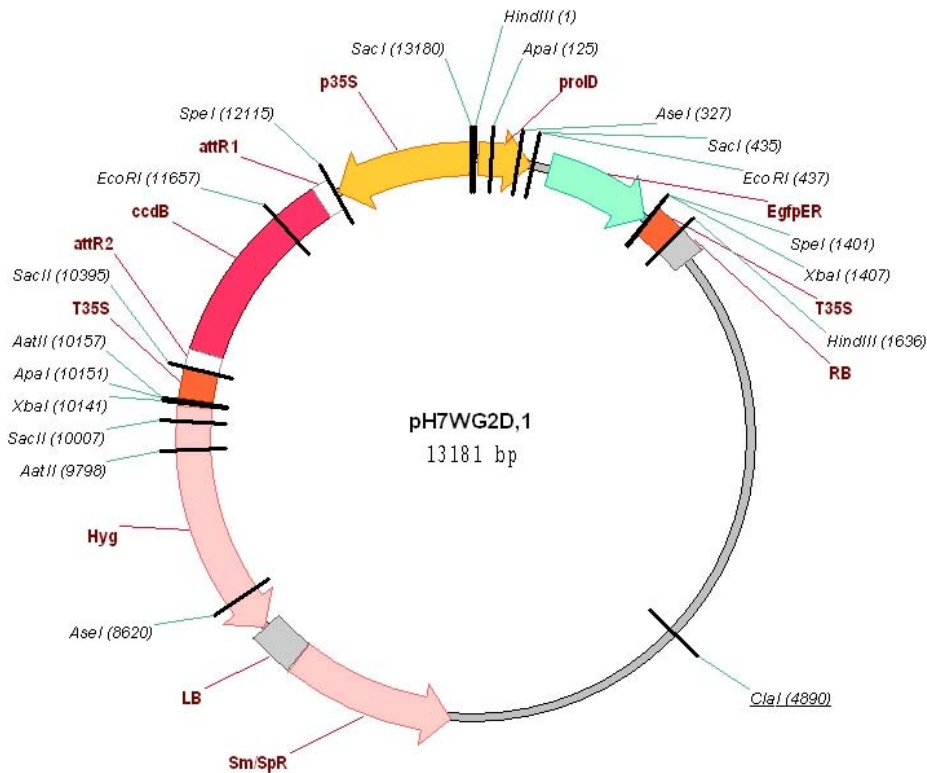
### **3.3. Obtenção de vetores para a transformação de soja**

Construções utilizando vetores de expressão em planta foram obtidas visando a superexpressão e o silenciamento dos genes em estudo. Estas foram obtidas através da tecnologia Gateway<sup>®</sup> de clonagem (Invitrogen), baseada na utilização de sítios de recombinação homóloga que garantem a inserção bem sucedida de fragmentos de DNA no vetor de interesse, sem a necessidade de utilização de enzimas de restrição. Primeiramente, o fragmento de DNA isolado foi amplificado com um “primer” direto específico, para que houvesse a adição da seqüência de nucleotídeos “CACC” em sua extremidade 5’, possibilitando a ligação adequada do fragmento de interesse ao vetor de entrada (pENTR Directional TOPO<sup>®</sup>, Invitrogen), uma vez que este possui uma extremidade “GTGG” proeminente. Esta ligação gerou uma construção que, a seguir, foi utilizada para efetivar a transferência do fragmento de interesse para o vetor de destino (neste caso os vetores de

expressão em plantas) através de uma “recombinação LR” utilizando a enzima LR clonase™ II (Invitrogen) de acordo com as especificações do fabricante. As construções obtidas contendo os fragmentos de interesse, tanto no vetor de entrada quanto nos vetores de destino, foram mantidas em *E. coli* Top 10 transformada por choque térmico, como descrito no item 3.4.

### **3.3.1. Construção de superexpressão**

O vetor de destino utilizado para a superexpressão foi o pH7WG2D,1 (Karimi, Inzé & Depicker, 2002 - <http://www.psb.ugent.be/gateway/>) (Figura 6). Este é um vetor binário no qual o fragmento de interesse é inserido entre o promotor e o terminador do 35S do CaMV (o promotor P35S é altamente ativo na maioria das células vegetais); genes de resistência à estreptomicina e à espectinomicina permitem a seleção em bactéria e a seleção em planta se deve à presença do gene *hpt* (higromicina fosfotransferase) que confere resistência à higromicina. Este plasmídeo possui ainda, um gene repórter que codifica uma “Green Fluorescent Protein” (*gfp*), cuja expressão é controlada pelo promotor rolD.

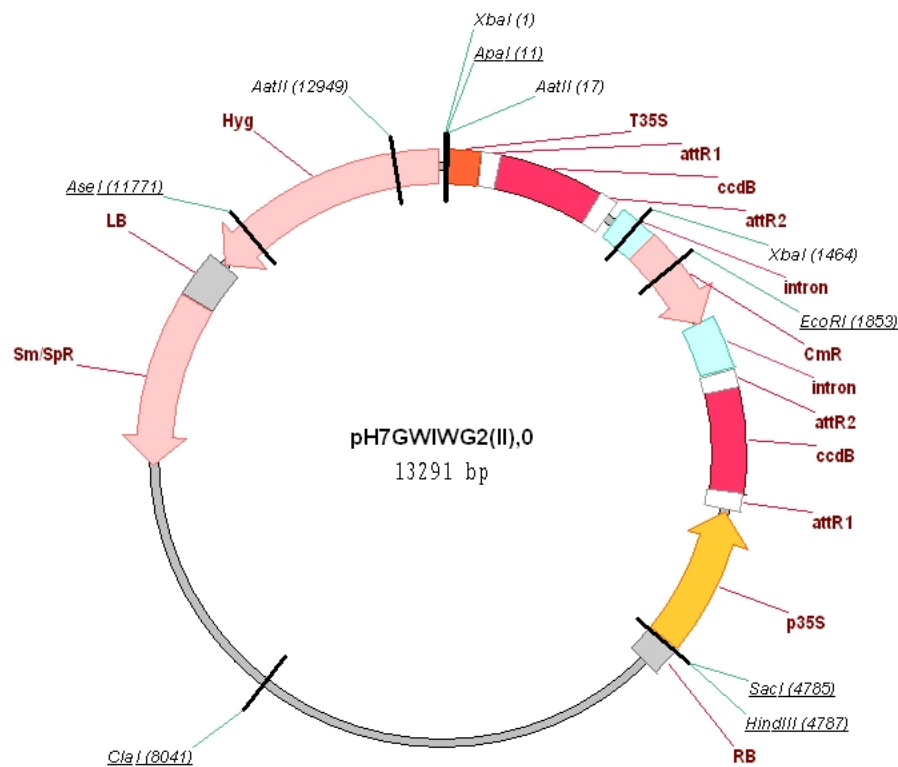


**Figura 6. Vetor de superexpressão pH7WG2D,1** (<http://www.psb.ugent.be/gateway/>). Neste vetor, o gene de interesse é inserido entre os sítios attR1 e attR2 através de recombinação homóloga, sendo sua expressão em planta controlada pelo promotor 35S de CaMV (em amarelo). Os genes de resistência a antibióticos estão representados em rosa: estreptomicina (Sm) e espectinomomicina (Sp) possibilitam a seleção de bactérias recombinantes, enquanto a seleção em planta é obtida pela resistência à higromicina (Hyg). O gene repórter *gfp* (em verde) é expresso sob o controle do promotor roID (em amarelo). RB e LB (em cinza) representam as bordas direita (“right border”) e esquerda (“left border”) do T-DNA (Modificado de Karimi, Inzé & Depicker, 2002).

### 3.3.2. Construção de silenciamento

Para o silenciamento por RNAi, o vetor de destino utilizado foi o pH7GWIWG2(II),0 (Karimi, Inzé & Depicker, 2002 - <http://www.psb.ugent.be/gateway/>) (Figura 7). Neste vetor binário, o fragmento de interesse é inserido duas vezes, de forma direta e inversa, havendo uma seqüência de íntron (de *Arabidopsis*) entre as duas inserções,

permitindo assim a formação de um “hairpin” RNA que resulte na ativação do silenciamento pós-transcricional de seqüências homólogas. A seleção em bactéria é feita devido à presença de genes de resistência à estreptomicina e à espectinomicina e a seleção em planta é obtida pela resistência à higromicina (*hpt*). Além disso, o plasmídeo possui um gene que confere resistência ao cloranfenicol (CmR) no meio da seqüência de íntron, permitindo a seleção apropriada deste.



**Figura 7. Vetor de silenciamento pH7GWIWG2(II),0 (<http://www.psb.ugent.be/gateway/>).** Neste vetor, o fragmento de interesse é inserido duas vezes, de forma direta e reversa, entre os sítios attR1 e attR2 e expresso em planta sob controle do promotor 35S de CaMV (em amarelo). Para que ocorra a formação apropriada do “hairpin”, uma seqüência de íntron (em azul) está situada entre os fragmentos clonados. Os genes de resistência a antibióticos estão representados em rosa: a seleção do íntron é realizada através da resistência ao cloranfenicol (Cm) em bactéria, estreptomicina (Sm) e espectinomicina (Sp) são utilizadas na seleção de colônias recombinantes e a seleção em planta é obtida pela resistência à

higromicina (Hyg). RB e LB (em cinza) representam as bordas direita (“right border”) e esquerda (“left border”) do T-DNA (Modificado de Karimi, Inzé & Depicker, 2002).

### 3.4. Transformação de *Escherichia coli*

Células termo-competentes de *E.coli* foram preparadas conforme protocolo descrito a seguir: colônias isoladas foram cultivadas em 3 ml de meio LB líquido, na presença do antibiótico adequado para a seleção de cada cepa – 10  $\mu\text{g.ml}^{-1}$  de tetraciclina para X-L1 Blue e 20  $\mu\text{g.ml}^{-1}$  de estreptomicina para Top10 – sob agitação constante de 225 rpm, a 37°C por 16 h. Uma pré-cultura foi feita (500  $\mu\text{l}$  em 20 ml de meio líquido com antibiótico), por 2 h e 30 min a 37°C, sob agitação (225 rpm); sendo esta transferida para tubos de centrifuga estéreis e centrifugada (5.000 rpm por 10 min). Após o descarte do sobrenadante, adicionou-se ao precipitado 5 ml de  $\text{CaCl}_2$  0,1 M, seguido de homogeneização (vórtex). As células foram mantidas no gelo por 20 min, sendo então repetidas a centrifugação (5.000 rpm por 10 min) e o descarte do sobrenadante. Em seguida, adicionou-se mais uma vez  $\text{CaCl}_2$  0,1 M (agora 0,5 ml) seguido da ressuspensão do precipitado no gelo. Este foi mantido no gelo até o momento da transformação, realizada por choque térmico, como descrito a seguir: a 100  $\mu\text{l}$  do precipitado bacteriano foram adicionados 300 ng do produto de cada ligação; a mistura foi mantida no gelo por 30 min, sendo então incubada a 42°C por 2 min e levada mais uma vez ao gelo, onde foi mantida por 2 min. A seguir, adicionou-se 400  $\mu\text{l}$  de meio LB líquido ao produto da transformação, sendo este incubado a 37°C por 40 min. Após esse período, as células bacterianas foram cultivadas em meio LB sólido acrescido dos antibióticos de seleção adequados para cada caso: para o vetor pGEM<sup>®</sup>-TEasy, adicionou-se ao meio de cultivo 50  $\mu\text{g.ml}^{-1}$  de ampicilina – além de 24  $\mu\text{l}$  de X-Gal (40  $\text{mg.ml}^{-1}$ ) e 10  $\mu\text{l}$  de IPTG (1M)<sup>1</sup>. Colônias de *E.*

---

<sup>1</sup> O Vetor pGEM<sup>®</sup>-TEasy possui mecanismo de seleção de recombinantes através do operon lac: a seqüência codificante do gene *lacZ* está inserida no sítio múltiplo de clonagem, resultando na produção da enzima  $\beta$ -galactosidase (colônias azuis). A inserção bem sucedida do fragmento de interesse nesta região resulta na interrupção do gene *lacZ*; neste caso a enzima não é produzida e as colônias bacterianas recombinantes são brancas.

*coli* Top10 recombinantes foram selecionadas pela adição de 50  $\mu\text{g.ml}^{-1}$  de canamicina para o vetor de entrada (pENTR Directional TOPO<sup>®</sup>); 20  $\mu\text{g.ml}^{-1}$  de estreptomicina e 75  $\mu\text{g.ml}^{-1}$  de espectinomicina para o vetor de superexpressão (pH7WG2D,1) e 20  $\mu\text{g.ml}^{-1}$  de estreptomicina, 75  $\mu\text{g.ml}^{-1}$  de espectinomicina e 100  $\mu\text{g.ml}^{-1}$  de cloranfenicol para o vetor de silenciamento [pH7GWIWG2(II),0].

### 3.5. Transformação de *Agrobacterium tumefaciens*

Células eletro-competentes de *A. tumefaciens* LBA4404 foram preparadas conforme o protocolo descrito a seguir: colônias isoladas foram cultivadas em 5 ml de meio LB líquido com rifampicina (50  $\mu\text{g.ml}^{-1}$ ) sob agitação constante de 225 rpm, a 28°C por 48 h. Pré-culturas desta foram feitas (500  $\mu\text{l}$  em 100 ml de meio líquido com rifampicina), com tempos diferentes. Após 16h, a densidade óptica (D.O.) destas culturas foi medida em espectrofotômetro, sendo utilizada a cultura que apresentasse D.O.<sub>600nm</sub> 0,5. A cultura escolhida foi então colocada em frascos GSA (30 ml de cultura por frasco de 40 ml) e centrifugada a 5.000 rpm por 15 min a 4°C. Descartou-se o sobrenadante e o precipitado foi ressuspensão em 150 ml de água milli-Q a 4°C. Este procedimento foi repetido três vezes, sendo o precipitado ressuspensão sucessivamente em 100 ml de água milli-Q a 4°C, 20 ml de glicerol 10% estéril e 2 ml de glicerol 10% estéril. As células foram divididas em alíquotas de 100  $\mu\text{l}$  e armazenadas em nitrogênio líquido até o momento da transformação. Esta foi realizada por eletroporação, onde aproximadamente 200  $\eta\text{g}$  do DNA recombinado foram homogeneizados a 100  $\mu\text{l}$  de bactéria competente, sendo a mistura submetida a uma diferença de potencial de 500 volts em eletroporador (BioRad Gene Pulser<sup>®</sup> II). Imediatamente após o choque, foram adicionados 400  $\mu\text{l}$  de LB líquido às células eletroporadas, sendo a suspensão cultivada por 1 h a 28°C sem agitação, seguida de 2 h a 28°C sob agitação constante de 225 rpm. A suspensão foi então plaqueada em meio sólido contendo rifampicina (50  $\mu\text{g.ml}^{-1}$ ) estreptomicina (20  $\mu\text{g.ml}^{-1}$ ) e espectinomicina (75  $\mu\text{g.ml}^{-1}$ ). O processo foi repetido para cada construção obtida.

### **3.6. Extração de DNA**

A extração de DNA genômico de folhas de soja foi realizada utilizando-se o tampão CTAB, conforme protocolo descrito por Doyle & Doyle em 1987, com algumas modificações. Após desidratação com sílica-gel (por 48 h, no mínimo), colocou-se aproximadamente 100 mg de cada amostra foliar em um tubo de micro-centrífuga, sendo este mergulhado diretamente no nitrogênio líquido e a amostra macerada com pistilo. A separação dos ácidos nucleicos de proteínas foi realizada em duas etapas: primeiramente adicionando-se fenol / clorofórmio (1:1) e em seguida clorofórmio / álcool isoamílico (24:1). Na etapa de precipitação com isopropanol, adicionou-se meio volume de acetato de amônio 7,5 M para a purificação do DNA precipitado. O DNA extraído foi solubilizado em 50 µl de água milli-Q estéril e as amostras foram quantificadas através do sistema Qubit™ (Quant-iT™ dsDNA BR Assay Kit, Invitrogen).

### **3.7. Extração de RNA total**

O material vegetal foi coletado e imediatamente congelado em nitrogênio líquido, sendo armazenado a -80°C até o momento da extração de RNA total. Esta foi realizada com o reagente Trizol (Invitrogen), de acordo com o protocolo do fabricante. Após a extração, a fim de eliminar contaminantes de DNA genômico, todas as amostras foram tratadas com DNase (Promega), sendo então quantificadas através do sistema Qubit™ (Quant-iT™ RNA Assay Kit, Invitrogen).

### **3.8. Síntese de cDNA**

A síntese de cDNA a partir do RNA total foi realizada utilizando-se o kit de transcrição reversa M-MLV (Promega), de acordo com as instruções do fabricante. Para cada reação, utilizou-se 10 µl de amostra de RNA tratado com DNase. Após a síntese, o volume final obtido de cada amostra (50 µl) foi diluído dez vezes (solução estoque) em água milli-Q estéril. Para uso nas reações de RT-qPCR, a solução estoque foi novamente diluída 10 vezes (diluição de 1:100).



### **3.9. Reação em cadeia da polimerase (PCR)**

As reações de amplificação dos DNAs genômicos, plasmidiais e cDNAs foram geralmente realizadas utilizando-se: 1 µl de MgCl<sub>2</sub> 50 mM, 2,5 µl de tampão 10x, 0,5 µl de dNTPS 10 mM, 1 µl de cada “primer” (direto e reverso) 10 mM, 0,25 µl de Taq DNA Polimerase (5 U/ µl), 17,75 µl de água milli-Q estéril e 1 µl de DNA. As condições utilizadas para a polimerização foram: desnaturação inicial a 94°C por 5 min, 35 ciclos de desnaturação a 94°C por 1 minuto, anelamento a 59°C por 1 minuto, extensão a 72°C por 1 minuto e extensão final a 72°C por 5 min. As amplificações foram submetidas à eletroforese e os fragmentos foram visualizados em gel de agarose (de 0,8 a 1,5%) corados com GelRed™ (Biotium, Inc) ou brometo de etídio. Uma lista com os “primers” utilizados nas reações de PCR encontra-se na Tabela 1.

**Tabela 1. Lista dos “primers” utilizados nas reações de PCR**

Primer	Direção	Seqüência
<i>GmWRKY20</i>	For	5'-CACCATGGAGAGTGA CTTGAG -3'
	Rev	5'- GTGCAATCAGCACAAAAATCC-3'
<i>GmWRKY46</i>	For	5'-CACCATGGAGAGTGA CTTGAG-3'
	Rev	5'- AGTGAAATGCAATCAGGAGAAA-3'
<i>RNAi1</i>	For	5'- CACCAAAGACATCCTTGGAGCCAA -3'
	Rev	5'- TGTGGTTGTCTTGTTCATGC-3'
<i>RNAi2</i>	For	5'- CACCCAGAGGCAAGCATACTGTTC -3'
	Rev	5'- TGTGGTTGTCTTGTTCATGC -3'
<i>5'UTRWRKY20</i>	For	5'- TTGCTTCCAGAGAAGAGCATC -3'
	Rev	-
<i>5'UTRWRKY46</i>	For	5'- TTCCA ACTCTGCTCCAGAAAA -3'
	Rev	-
<i>P35S</i>	For	5'- GGACCCCCACCCACGAGGAG-3'
	Rev	-
<i>Íntron1</i>	For	5'-GTCAAGCTGACCTGCAAACA-3'
	Rev	5'- TGCCTCTTCTTACGGCTTTC-3'
<i>Hpt1</i>	For	5'- GCGATTGCTGATCCCCATGTGTGAT-3'
	Rev	5'- GGTTTCCACTATCGGCGAGTACTT-3'

### 3.10. PCR em tempo real quantitativo (RT-qPCR)

As reações de RT-qPCR foram realizadas em aparelho StepOne Plus™ Real Time PCR System (Applied Biosystems™), conforme as seguintes condições: uma etapa de

desnaturação inicial de 5 min a 94°C, seguida por 45 ciclos de 10 s a 94°C, 15 s a 60°C e 15 s a 72°C. Após a amplificação, as amostras foram mantidas por 2 min a 40°C para reanelamento e, então, aquecidas de 55 até 99°C com uma taxa de incremento de 0.1°C/s para aquisição de dados e obtenção da curva de desnaturação. As reações foram feitas em um volume final de 20 µl, sendo os mesmos compostos de 2 µl de tampão PCR 10x (Invitrogen), 1,2 µl de MgCl<sub>2</sub> 50 mM, 0,4 µl de dNTPs 5 mM, 0,4 µl de cada par de “primers” 5 µM, 4,95 µl de água ultra pura, 1,0 µl de SYBR<sup>®</sup> Green (1:100.000, Molecular Probes Inc.) e 0,05 µl de Platinum Taq DNA polimerase (5 U/µl, Invitrogen). A este “mix”, adicionou-se: 10 µl de cDNA (diluição 1:100) nas análises de expressão gênica; ou 10 µl de DNA genômico em diluições seriadas (1:100, 1:1.000 e 1: 10.000), para a estimativa do número de cópias do transgene. A temperatura de anelamento de todos os “primers” foi ajustada para 60°C. Uma lista com os “primers” utilizados nas reações de RT-qPCR encontra-se na Tabela 2.

**Tabela 2. Lista dos “primers” utilizados nas reações de RT-qPCR**

Primer	Direção	Seqüência
<i>GmWRKY20rt</i>	For	5'- TTGCAAAGTTCAGAAAGTATCTTGTC -3'
	Rev	5'- GTGACCTGTTGTAGATCCCATC -3'
<i>GmWRKY46rt</i>	For	5'- CAATGCATCATCAACTTCCG -3'
	Rev	5'- CAAGACCACTTTCACAGCTCAC -3'
<i>GmFbox</i>	For	5'- AGATAGGGAAATGTTGCAGGT -3'
	Rev	5'- CTAATGGCAATTGCAGCTCTC -3'
<i>GmMetallo</i>	For	5'- ATGAATGACGGTTCCCATGTA -3'
	Rev	5'- GGCATTAAGGCAGCTCACTCT -3'
<i>GmLe1</i>	For	5'- TACCTATGATGCCTCCACCA -3'
	Rev	5'- GAGACCCCTATCCTCACCCA -3'

Nas análises de expressão gênica, os resultados foram expressos através de uma quantificação relativa a dois genes utilizados como controles endógenos da reação – “F-box

protein family” e “metalloprotease” (Libault et al. 2008) – através da metodologia  $2^{-\Delta\Delta Ct}$  descrita por Livak & Schmittgen em 2001.

A estimativa do número de cópias do transgene por RT-qPCR foi realizada através da “Quantificação relativa pela análise da curva padrão” (Shou et al., 2004). Nesta técnica, a diluição seriada das amostras possibilita a obtenção de curvas-padrão relativas, onde os valores de Ct entre o gene alvo e um controle endógeno são comparados. Este controle endógeno é um gene cujo número de cópias no genoma em estudo é pré-estabelecido, sendo utilizado para a normalização da reação – os valores obtidos nas diluições para o gene-alvo são divididos pelos valores obtidos para o controle endógeno (“*quantidade relativa ao gene endógeno*”,1). Além disso, é necessária a utilização de um “calibrador”, ou seja, uma planta onde o número de cópias do gene-alvo é conhecido. Os valores normalizados obtidos nas amostras analisadas são divididos pelos valores normalizados obtidos nas amostras do calibrador (“*número de cópias*”, 2) (Bubner & Baldwin, 2004; Shou et al., 2004). O controle endógeno utilizado foi o gene da lectina *Le1* (Schmidt et al. 2001) e plantas de soja WT foram utilizadas como calibradores.

### **3.11. Análises da expressão gênica ao longo do desenvolvimento e em resposta a estresses**

#### **3.11.1. Diferentes órgãos e fases do desenvolvimento**

Plantas de soja da cultivar MGBR 46 (Conquista) foram semeadas em casa de vegetação e as amostras coletadas como especificado na Tabela 3, sendo imediatamente congeladas em nitrogênio líquido e armazenadas a  $-80^{\circ}\text{C}$  até o momento da extração de RNA. De um grupo de plantas jovens, foram coletadas amostras de raiz e folha, 14 dias após a semeadura destas. De outro grupo, três meses após a semeadura foram coletadas amostras de flor e de folha. No mês seguinte foram coletadas – deste mesmo grupo de plantas maduras – amostras de caule, além das amostras de vagem e semente. Para cada tipo de amostra analisada, foram utilizadas quatro replicatas biológicas (quatro vasos), sendo cada uma destas composta por um “pool” de quatro plantas diferentes.

**Tabela 3. Órgãos utilizados nas análises de expressão gênica e as respectivas fases do desenvolvimento nas quais estes foram coletados.**

Fases do desenvolvimento	Órgãos
0 Planta jovem	Folha
	Raiz
1 Durante a floração	Folha
	Flor fechada (4-5 mm)
	Flor na fecundação (6-7 mm)
	Flor aberta
2 30 dias após a floração	Caule
	Semente imatura*
	Semente com embrião
	Vagem**

\* Sem embrião.

\*\* Sem semente.

### 3.11.2. Infecção por *P. pachyrhizi*

O experimento de infecção por *P. pachyrhizi* foi realizado em abril de 2009, na sede da Embrapa Soja (Londrina, PR). Plantas de soja da cultivar Embrapa – 48 (genótipo suscetível) e da linhagem PI 561356 (genótipo resistente) foram semeadas em casa de vegetação – seis vasos com cada genótipo, três plantas por vaso. Três semanas após o início do cultivo, realizou-se a inoculação com o patógeno: urediniosporos foram coletados e ressuspensos em água contendo gotas de Tween-20, para uma concentração final de  $2,6 \times 10^4$  esporos. ml<sup>-1</sup>. A solução foi borrifada nas folhas de metade das plantas de cada genótipo, sendo a outra metade utilizada como controle (“mock-inoculated”). A inoculação ocorreu após as 18h (à noite a umidade é maior, facilitando o processo de

infecção pelo fungo) e as plantas inoculadas permaneceram por doze horas cobertas por plástico. Trifólios foram coletados 1, 12, 24, 48, 96 e 192 hpi, sendo imediatamente congelados em nitrogênio líquido e armazenados a -80°C até o momento da extração de RNA.

### **3.11.3. Estresse salino**

Após duas semanas de cultivo em câmara de crescimento (em copos plásticos contendo vermiculita), plantas de soja foram tratadas com solução salina (150 mM de NaCl), duas horas após o início do fotoperíodo. Amostras de quatro copos, com duas plantas cada, foram utilizadas como controle (“mock”) e a mesma quantidade foi tratada. Amostras de raízes e folhas foram coletadas 6 h após o tratamento, sendo imediatamente submetidas à extração de RNA.

### **3.12. Análises estatísticas**

Os resultados obtidos nos experimentos de quantificação da expressão gênica (item 3.11) foram avaliados quanto a sua significância estatística através das seguintes análises: “one-way” ANOVA seguida do teste de Tukey HSD ( $p < 0,05$ ) (SAS/STAT® Software v.8) nas análises de expressão em diversos órgãos e fases do desenvolvimento; e teste - T de Student ( $p < 0,05$ ) (Microsoft® Office Excel 2007) nas análises de expressão na infecção por *P. pachyrhizi* e em condições de estresse salino.

### **3.13. Transformação genética de soja**

Uma lista com os meios de cultivo vegetal utilizados nas etapas descritas a seguir é apresentada na Tabela 4. Os conjuntos embriogênicos submetidos a cada experimento de transformação realizado são descritos na Tabela 5.

### **3.13.1. Indução de embriogênese somática e obtenção de tecido embriogênico proliferativo**

Vagens imaturas de soja foram colhidas de plantas cultivadas no campo e esterilizadas pela imersão em etanol 70% (1 minuto), seguida da imersão em solução de hipoclorito de sódio 4% contendo gotas de Tween-20 (15 min) e de três lavagens subsequentes, com água destilada autoclavada. Para a indução de embriogênese somática, sementes (3-5mm) foram excisadas das vagens esterilizadas. As metades dos cotilédones removidos foram utilizadas como explantes e colocadas em meio de indução D40. Após 30 dias, os explantes embriogênicos foram transferidos para meio de proliferação semi-sólido D20. Após 14 dias, os conjuntos de embriões secundários foram removidos dos cotilédones e transferidos para meio fresco. Sub-culturas destes para meio D20 fresco foram realizadas a cada 14 dias. A cultura *in vitro* foi mantida em câmara de crescimento a  $26\pm 1^{\circ}\text{C}$  de temperatura, com fotoperíodo de 16 h e intensidade luminosa de  $20\ \mu\text{mol m}^{-2}\text{s}^{-1}$ .

### **3.13.2. Transformação por biobalística (bombardeamento de partículas)**

Este processo de transformação foi realizado através de um acelerador de partículas de baixa pressão (PIG – Finer et al., 1992), segundo protocolo descrito por Finer & McMullen em 1991. Previamente ao bombardeio, placas de Petri com conjuntos de embriões globulares (15 conjuntos por placa) foram abertas em capela de fluxo laminar (por 15 min) para que estes fossem parcialmente desidratados, reduzindo assim a pressão de turgor do material vegetal (Vain et al. 1993). Às partículas de tungstênio diluídas em água estéril ( $100\ \text{mg}\cdot\text{ml}^{-1}$ ), adicionou-se  $2\ \mu\text{l}$  do DNA-alvo ( $250\ \eta\text{g}\cdot\mu\text{l}^{-1}$ ). A esta mistura, foram adicionados  $25\ \mu\text{l}$  de  $\text{CaCl}_2$  2,5 M e  $10\ \mu\text{l}$  de espermidina 0,1 M. Após homogeneização e incubação no gelo (5 min),  $45\ \mu\text{l}$  do sobrenadante foi descartado, sendo  $2\ \mu\text{l}$  do precipitado utilizado em cada disparo.

### **3.13.3. Transformação via sistema integrado “bombardeamento / *Agrobacterium*”**

Linhagens de *A. tumefaciens* recombinante foram preparadas de acordo com procedimento descrito por Droste et al., em 2000: colônias isoladas foram cultivadas por 48 h em meio LB líquido contendo rifampicina ( $50 \text{ mg.l}^{-1}$ ), canamicina ( $50 \text{ mg.l}^{-1}$ ) e acetoseringona ( $100 \text{ }\mu\text{M}$ ), a  $28^\circ\text{C}$  sob agitação contínua. Após centrifugação, as células foram ressuspensas em meio líquido D10 com  $100 \text{ }\mu\text{M}$  de acetoseringona (D.O.  $_{600\text{nm}}$  final de 0,3).

A seguir, realizou-se o protocolo de bombardeamento, como descrito no item 3.13.2. Neste caso, as partículas foram preparadas livres de DNA; servindo apenas para causar micro-fermentos nas células vegetais bombardeadas, facilitando a subsequente infecção por *A. tumefaciens*. Após serem bombardeados, os conjuntos de embriões foram incubados com a suspensão bacteriana durante 20 min, sendo a seguir, secos em papel filtro estéril (para a retirada do excesso de bactéria) e co-cultivados por 48 h em meio D20 com acetoseringona ( $100 \text{ }\mu\text{M}$ ). Ao co-cultivo sucedeu-se a lavagem dos embriões em água destilada autoclavada, seguida de secagem em papel filtro estéril e transferência para meio D20 contendo  $250 \text{ mg.l}^{-1}$  de cefotaxima e  $250 \text{ mg.l}^{-1}$  de vancomicina (Wiebke et al., 2006). Estes antibióticos foram adicionados aos meios de cultivo subsequentes, inclusive durante o período de seleção por higromicina, até que o tratamento completasse 49 dias, garantindo a eliminação total da bactéria utilizada na transformação.

### **3.13.4. Seleção de transformantes, regeneração e aclimatação das plantas obtidas**

Após serem submetidos a qualquer um dos processos de transformação genética referidos acima, os conjuntos embriogênicos de soja foram submetidos à seleção de transformantes através do cultivo vegetal em doses crescentes de higromicina. Esta seleção teve início 10 dias após a realização de cada experimento de transformação, quando se adicionou  $12,5 \text{ mg.l}^{-1}$  de higromicina ao meio de proliferação de embriões



(D20). Após 21 dias, a dose de higromicina foi aumentada para  $25 \text{ mg.l}^{-1}$ . Após três meses de seleção, ao total, conjuntos de embriões resistentes à higromicina (que permaneceram verdes) foram destacados dos embriões não resistentes e proliferados em meio D20 por três meses, para que os eventos de transformação obtidos fossem multiplicados. A seguir, os conjuntos proliferados foram transferidos para meio MSM6 (acrescido de 1% de carvão ativado) para iniciarem seu processo de histodiferenciação e maturação, onde permaneceram por 30 dias. Por mais 30 dias, os embriões permaneceram em meio MSM6 (sem carvão ativado) para que os processos de histodiferenciação e maturação fossem completados. Embriões maduros obtidos passaram por um processo de dessecação (Buchheim et al. 1989), permanecendo em placas sem meio de cultura por 6 h; sendo logo em seguida transferidos para meio MSO para a conversão em plântulas (em placas de Petri) e subsequente regeneração (em frascos). Plântulas regeneradas e enraizadas foram aclimatadas; primeiramente em copos plásticos contendo vermiculita e cobertos com filme plástico, sendo expostas de forma gradual às condições ambientais; e após 10 dias (no mínimo), transferidas para vasos com solo orgânico, sendo mantidas em câmara de crescimento com temperatura média de  $27 \pm 1^\circ\text{C}$ , fotoperíodo de 16 h e intensidade luminosa de aproximadamente  $60 \mu\text{mol m}^{-2} \text{ s}^{-1}$ .

**Tabela 4. Composição dos meios de cultivo vegetal utilizados na cultura de embriões somáticos e na transformação genética de soja.**

Meio de Cultivo	Componentes
D40 <sup>†</sup>	Sais do MS*, vitaminas do B5**, 40 mg.l <sup>-1</sup> de 2,4-D, 3% de sacarose, 0,3% de Phytigel <sup>TM</sup> , pH 7,0
D20 <sup>††</sup>	Sais do MS*, vitaminas do B5**, 20 mg.l <sup>-1</sup> de 2,4-D, 3% de sacarose, 0,3% de Phytigel <sup>TM</sup> , pH 6,4
D10 <sup>◇</sup>	Sais do MS*, vitaminas do B5**, 10 mg.l <sup>-1</sup> de 2,4-D, 3% de sacarose, pH 6,4
MSM6 <sup>◇◇</sup>	Sais do MS*, vitaminas do B5**, 6% de sacarose, 0,3% de Phytigel <sup>TM</sup> , pH 6,4
MSO <sup>◇◇</sup>	Sais do MS*, vitaminas do B5**, 3% de sacarose, 0,3% de Phytigel <sup>TM</sup> , pH 6,4

<sup>†</sup> Bailey, Boerma & Parrott, 1993

<sup>††</sup> Wright et al., 1991

<sup>◇</sup> Droste, Pasquali & Bodanese-Zanettini, 2000

<sup>◇◇</sup> Finer & McMullen, 1991

\* Murashige & Skoog, 1962

\*\* Gamborg, Miller & Ojima, 1968

**Tabela 5. Conjuntos embriogênicos submetidos aos experimentos de transformação realizados.**

Cultivares / construções utilizadas	Experimento I (bombardeamento / <i>Agrobacterium</i> )		Experimento II (bombardeamento)		Experimento III (bombardeamento / <i>Agrobacterium</i> )		
	IAS5	Venc	IAS5	Venc	IAS5	Venc	Conq
pH7WG2D,1 + <i>GmWRKY20</i>	150	150	-	-	-	-	-
pH7GWIWG2(II),0 + RNAi1	-	-	150	150	75	45	30
pH7GWIWG2(II),0 + RNAi2	-	-	-	-	75	30	45

### **3.13.5. Monitoramento da integração do transgene através da análise da expressão do gene repórter**

Como previamente mencionado, o gene *gfp* está presente entre as bordas do T-DNA do vetor de superexpressão utilizado (pH7WG2D,1). Este é utilizado como gene repórter, uma vez que o monitoramento da integração do T-DNA ao genoma vegetal pode ser feito através da visualização da expressão de *gfp* sob luz azul. Esta foi realizada em microscópio de fluorescência Olympus<sup>®</sup> (filtro BP; comprimentos de onda de excitação e emissão de 488 e 505-530 nm, respectivamente), com câmera fotográfica acoplada. As imagens foram obtidas através do software QCapture Pro<sup>™</sup> 6 (QImaging<sup>®</sup>).

## 4. RESULTADOS

### 4.1. Seleção, análise *in silico* e isolamento dos genes para estudo

Em 2007, van De Mortel et al. detectaram, através de análise de microarranjo, a expressão diferencial de 46 genes que codificam proteínas WRKY em resposta à infecção por ASR. Alguns destes foram pré-selecionados para estudo, de acordo com seu perfil de expressão. Entretanto, em alguns casos a “etiqueta” obtida através do código do microarranjo era insuficiente para a subsequente identificação *in silico* do gene, de forma que apenas dois genes, os quais possuíam uma porção considerável de sua região codificante disponível no GenBank foram escolhidos para o presente estudo. Através de seus respectivos códigos do chip de microarranjo [GmaAffx.51816.1.S1\_at e Gma.1256.1.S1\_at, Gene Chip<sup>®</sup> Soybean Genome Array (Affimetrix)], as seqüências das ESTs referentes a estes genes (BE804769 e BI967912) foram obtidas no GenBank. A busca por sua identidade gênica no GenBank (ferramenta “blastn”), resultou em duas seqüências parciais codificantes (EU019582.1 e EU019564.1) para ambas as ESTs, respectivamente anotadas como *GmWRKY46* e *GmWRKY20*. Curiosamente, quando estes dois fragmentos são alinhados, eles se complementam de forma parcial, aparentemente representando as extremidades amino- e carboxi-terminal da mesma proteína (Figura 8) e sua seqüência traduzida possui todas as características estruturais comuns aos membros da família dos fatores de transcrição WRKY. Além disso, ao se realizar uma busca por suas respectivas seqüências genômicas no Phytozome, ambas as ESTs apresentaram alta similaridade com duas regiões distintas do genoma (Figura 9): “scaffold\_78” e “scaffold\_92”, que foram identificadas respectivamente como Glyma05g36970 e Glyma08g02580, após o término da organização dos dados do genoma. Estes resultados indicavam que, apesar do resultado obtido no GenBank, as ESTs e seus respectivos perfis de expressão se referem a dois genes distintos, com alto grau de identidade.

```

CLUSTAL W (1.81) multiple sequence alignment

WRKY20      MESDLSWERNTLINELIQGMEVARKLKLADLKLPSVDTRDLLLQRIILSSYEKALLILRCS
WRKY46      -----

WRKY20      NASTSELQGMNQATPTLLPE SPLSVHGSPLREDVHGAIMDHHDSKKRKITRRWMDRVRVS
WRKY46      -----

WRKY20      CESGLEGPHEHDGYNWRKYGQKDI LGAKYPRSYRCTFRSTQGCWATKQVQRSDPTVFD
WRKY46      -----QKDI LGAKYPRSYRCTFRSTQGCWATKQVQRSDPTMFD
                      *****:*****

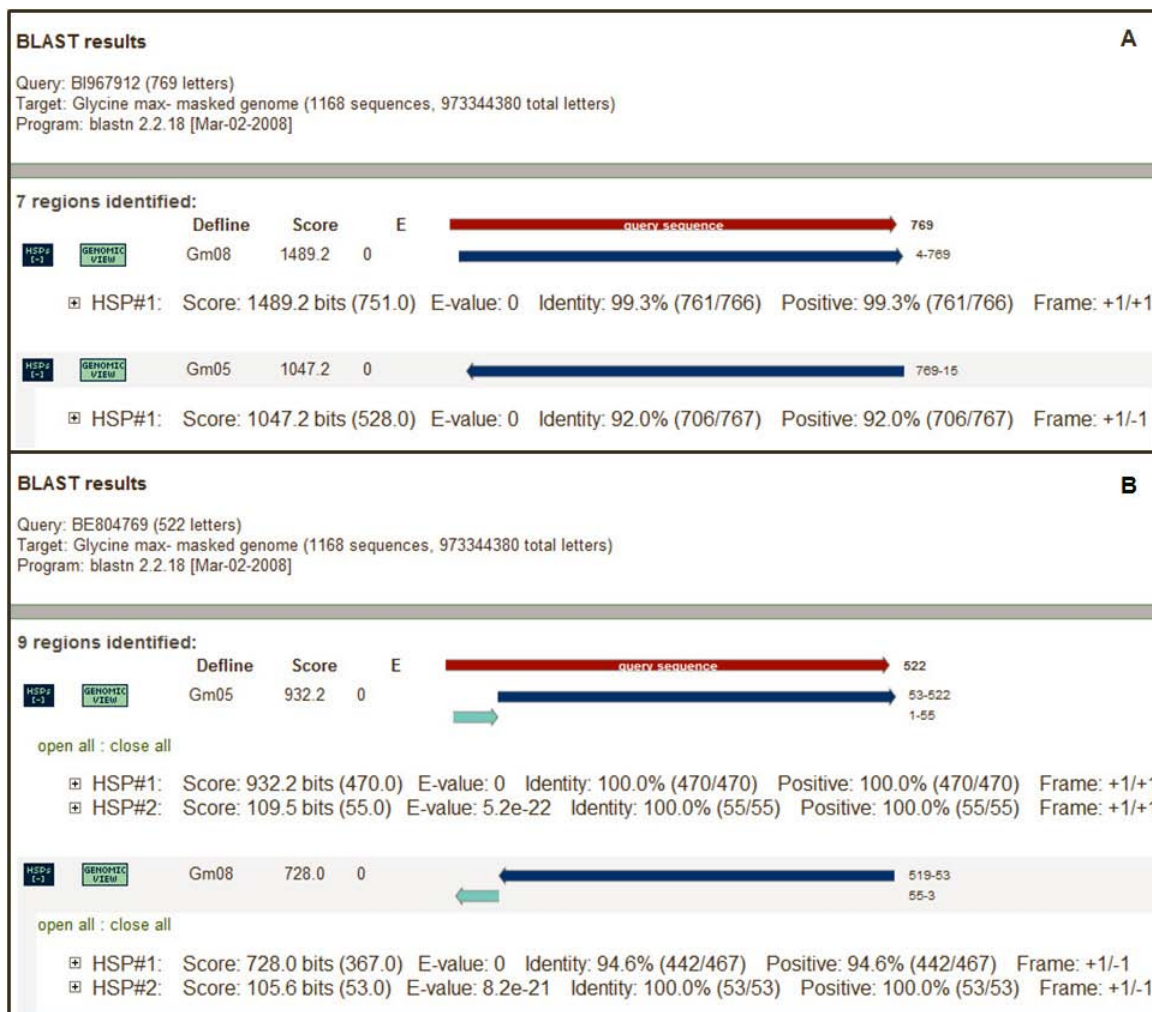
WRKY20      ITYRGKHTCSQGNNAVLPPKSPEKQEKPTHSHNIDIHRAQASQESLTKFRNILSVNTDNL
WRKY46      ITYRGNHTCSQGNNAVLPPKSPEKHEKPAHSHNIDIHHAQASQESLAKFRSILSVNTDNL
                      *****:*****:***:*****:*****:***.*****

WRKY20      NNGDMAYPFTFPSTSFSGCMKQDNHSLIP-----
WRKY46      DNGDMAYAFTFPSTSFSGCMKQDNHSLIPLALENDSFLSDLYQTHLLSEPTTPE SNYF PSPT
                      :*****.*****

WRKY20      -----
WRKY46      FQMNFEFDGIYNRSHSKSDINEIISTNTSATNSPIPDFNFSLDPVEIDPNFPFNTPGFLC

```

**Figura 8. Alinhamento das seqüências traduzidas das ESTs obtidas no GenBank, anotadas respectivamente como *GmWRKY20* e *GmWRKY46*. Asteriscos (\*) indicam resíduos idênticos, dois pontos (:) indicam resíduos conservados e pontos (.) indicam resíduos semi-conservados. O programa ClustalW v1.81 (<http://align.genome.jp/>) foi utilizado para o alinhamento.**



**Figura 9. Resultado das buscas realizadas no Phytozome (www.phytozome.org/soybean).** As seqüências genômicas correspondentes às ESTs obtidas no GenBank para *GmWRKY20* (A) e *GmWRKY46* (B) foram buscadas no Phytozome (ferramenta “blast genome”). Nota-se que para ambas as ESTs utilizadas, as mesmas duas regiões do genoma (Gm05 e Gm08) apresentaram correspondência, sendo que cada uma destas regiões apresentou maior ou menor grau de identidade com uma das duas seqüências.

Na tentativa de resolver esta contradição e confirmar a presença de dois genes codificando proteínas muito similares, “primers” foram desenhados visando o isolamento – tanto da seqüência genômica quanto da codificante – dos genes em questão. Os fragmentos obtidos foram clonados em vetor pGEM<sup>®</sup>-TEasy (Promega) e enviados para seqüenciamento (dados não mostrados). O alinhamento das seqüências completas de

aminoácidos de *GmWRKY20* e *GmWRKY46*, obtidas a partir de suas seqüências de nucleotídeos codificantes, pode ser visualizado na Figura 10. Nela estão ressaltados os resíduos de aminoácidos conservados presentes em cada gene. Sua análise possibilitou a classificação de ambos os genes no grupo III da superfamília de fatores de transcrição WRKY (Eulgem et al., 2000). A Figura 11 é uma representação gráfica da estrutura de éxons e íntrons de ambos os genes, ressaltando a presença uma pequena inserção/deleção de 12 nucleotídeos situada na primeira junção éxon/íntron, sendo esta a maior diferença estrutural apresentada entre as duas seqüências. Um resumo com os dados obtidos sobre os genes em estudo é apresentado na Tabela 6.

CLUSTAL W (1.81) multiple sequence alignment

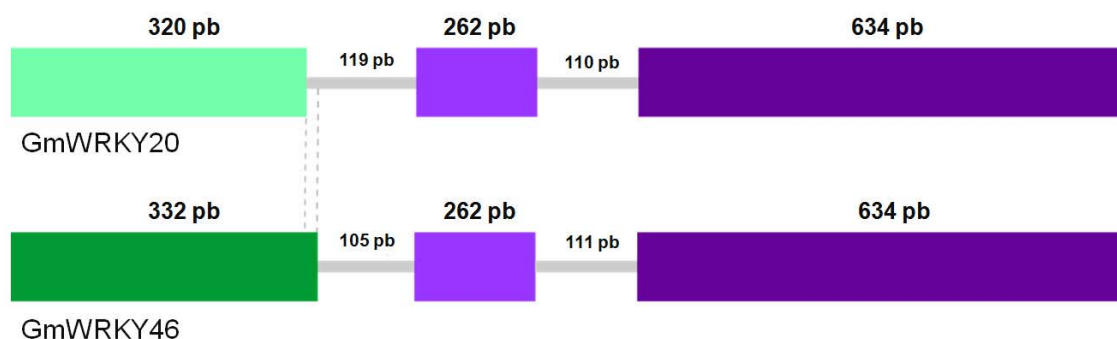
```

GmWRKY20      MESDLSWERNTLINELIQGMEVARKLKADLKLPSVDTRDLLLQRISSYEKALLILRCS
GmWRKY46      MESDLSWEQNTLINELIQGMEVARKLKADLRMPYSVDSRDLLVQRILSSYEKALLILRCN
*****.******.******.******.******.******.******.*
GmWRKY20      NASTSELQGMNQATPTLLPESPLSVHGSP LREDVHGAIKDHHSK-----REITPKWMDR
GmWRKY46      ASSTSELQAMSQATPTLLPESPLSVHGSP LREDVDGTIKDHQEVKHSKRRKATPKWMDH
.******.*.******.******.*.******.:.*:      *:*      *:*      *:*
GmWRKY20      VRVSCESGLEGPHE D GYN WRKYGQK D I LGAKY PRS Y YR C T FRSTQGCWATKQVQRS DEDP
GmWRKY46      VRVSCESGLEGPHE D SYN WRKYGQK D I LGAKY PRS Y YR C T FRNTQGCWATKQVQRS DEDP
*****.******.******.******.******.******.******.******.*
GmWRKY20      TMFDI TYR GNH T CSQGNNAVLPPKSP EKHEKPAHSHNIDIHHAQASQESLAKFRSILSVN
GmWRKY46      TVFDI TYR GNH T CSQGNNAVLPPKSP EKQEKPTHSHNIDIHRAQASQESLTKFRNILLSVN
*.******.*.******.******.*.*.*.******.******.******.******.*
GmWRKY20      TDNLDNGDMAYAF TFPSTSF GCMQDNHSLIPLALENDSFLSDLYQTHLLSPTTPESNYF
GmWRKY46      TDNLNNGDMAYP FTFPSTSF GCMQDNHSLIPWALENESFLSDLYQTHLLSTTIPESNYF
****.******.*.******.******.******.******.******.******.*
GmWRKY20      PSPTFQMNEFDGIYNRSHSKSDINEIISTNTSATNSPIPDFNFSLDPVEIDPNFPFNTPG
GmWRKY46      PSPTFQMNVFDGIYSKPHSESDINEIISTNTSATNSPIPDFNFSLDPVEIDPNFPFNTPG
*****.******.*.******.******.******.******.******.******.*
GmWRKY20      FLC
GmWRKY46      LFS
::.

```

**Figura 10. Alinhamento dos aminoácidos deduzidos a partir seqüências codificantes completas clonadas de *GmWRKY20* e *GmWRKY46* e identificação dos resíduos conservados entre WRKYs grupo III.** Asteriscos (\*) indicam resíduos idênticos, dois pontos (:) indicam resíduos conservados e pontos (.) indicam resíduos semi-conservados. O programa ClustalW v1.81 (<http://align.genome.jp/>) foi utilizado para o alinhamento. A identificação dos resíduos de aminoácidos conservados entre os fatores de

transcrição WRKY (em rosa) e do padrão C<sub>2</sub>-HC nos prováveis sítios de ligação a zinco (em verde), permitiu a classificação desses genes no grupo III da superfamília (de acordo com Eulgem et al., 2000).



**Figura 11. Representação gráfica comparativa da estrutura de éxons e íntrons de *GmWRKY20* e *GmWRKY46*.** Os éxons estão representados por blocos coloridos e os íntrons por linhas cinza. A diferença de tonalidade entre o primeiro éxon de cada gene (verde claro / verde escuro) representa a diferença de tamanho apresentada por eles: esta é causada por uma inserção / deleção de 12 nucleotídeos (linhas pontilhadas verticais) ao final do primeiro éxon, resultando na principal diferença estrutural existente entre os genes em estudo. Pb, pares de base.

**Tabela 6. Nomenclatura, características estruturais, classificação, localização no genoma e códigos de acesso dos genes selecionados para estudo.**

Gene	Cromossomo	Número de		cDNA (pb)	Proteína (aa)	Classificação na família*	Código de acesso		
		Éxons	Íntrons				Gene Chip* Soybean Genome Array (Affimetrix)	GenBank (NCBI)	Phytozome
<i>GmWRKY20</i>	8	3	2	1.080	360	Grupo III	Gma.1256.1.S1_at	BI967912 (EU019564.1)	Glyma08g02580
<i>GmWRKY46</i>	5	3	2	1.092	364	Grupo III	GmaAffx.51816.1.S1_at	BE804769 (EU019582.1)	Glyma05g36970

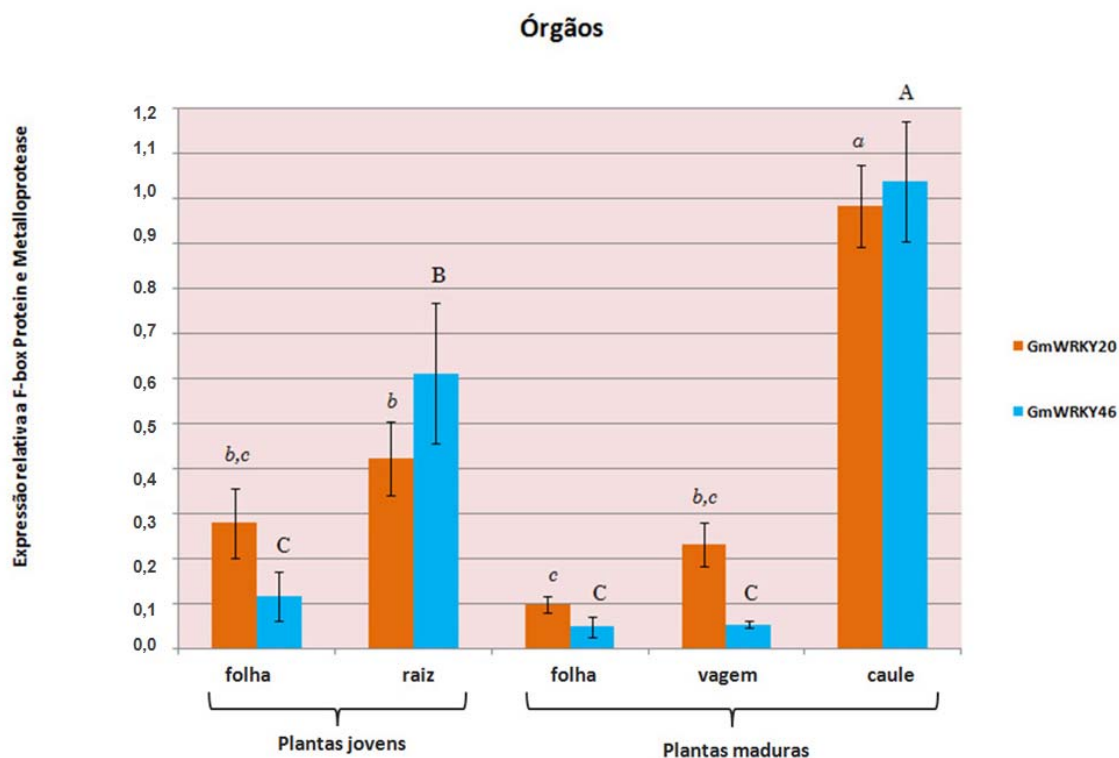
\* de acordo com Eulgem et al., 2000.



## **4.2. Análises da expressão gênica**

### **4.2.1. Diferentes órgãos e fases do desenvolvimento**

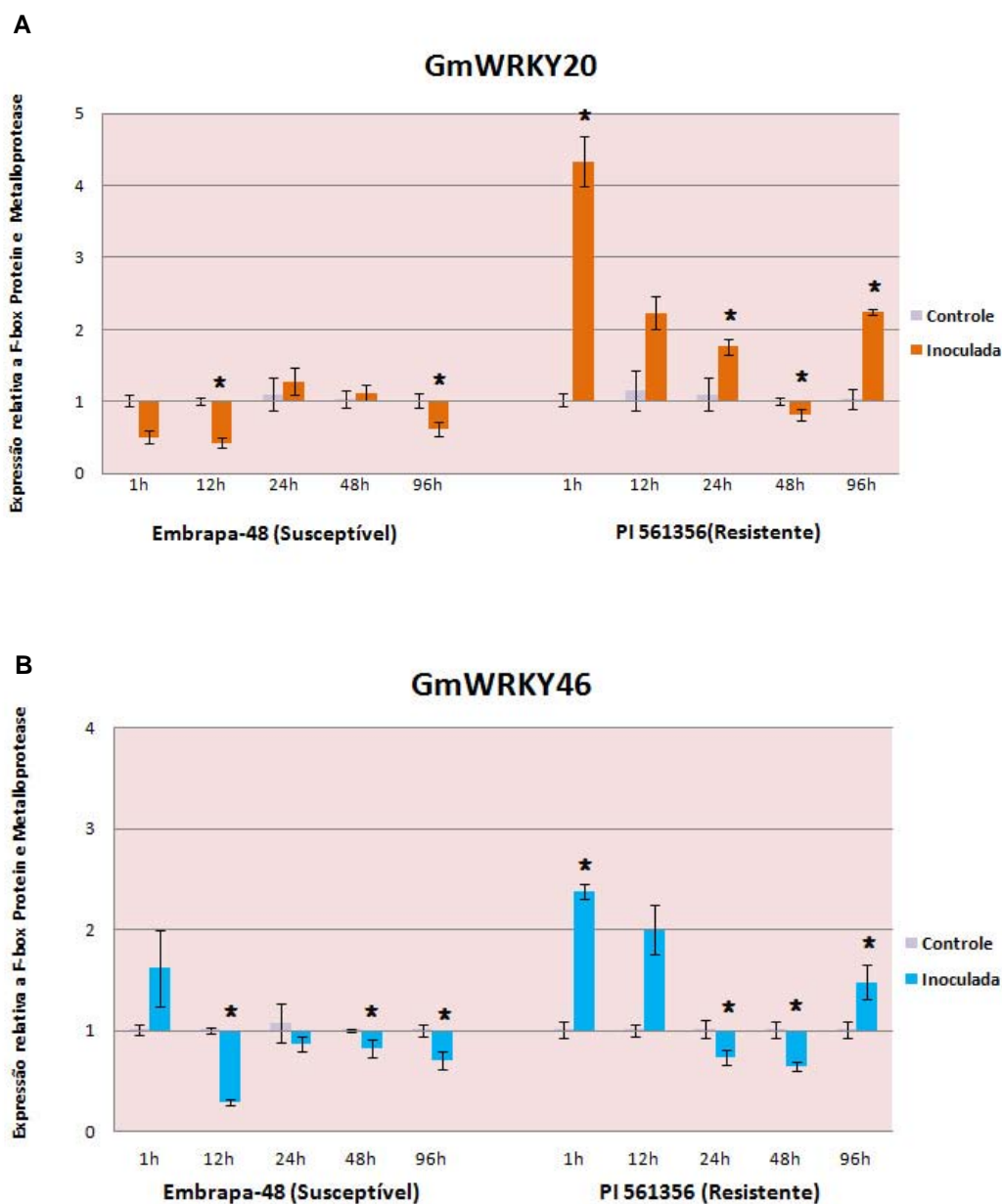
O nível de expressão de transcritos de *GmWRKY20* e *GmWRKY46* foi medido em diferentes órgãos e fases do desenvolvimento de plantas de soja (Figura 12). Nas condições em que o experimento foi realizado, a expressão de ambos os genes não foi detectada em nenhuma das amostras de flor nem de semente utilizadas. Na maioria dos órgãos analisados, os níveis de expressão entre *GmWRKY20* e *GmWRKY46* não diferiram estatisticamente, exceto na vagem. A expressão de *GmWRKY20* foi significativamente maior no caule de plantas maduras, sendo a folha dessas plantas o órgão que apresentou menor expressão. De forma semelhante, o maior nível de expressão de *GmWRKY46* foi encontrado no caule de plantas maduras. Entretanto, o nível de expressão apresentado em raízes de plantas jovens também foi significativamente maior, se comparado aos outros órgãos analisados.



**Figura 12. Expressão relativa de *GmWRKY20* e *GmWRKY46* em diferentes órgãos.** No gráfico é apresentada uma análise comparativa entre os níveis de expressão obtidos nos diversos órgãos utilizados. Os valores foram normalizados em relação ao nível de expressão obtido para *GmWRKY46* no caule (atribuído arbitrariamente como 1). Letras minúsculas diferentes indicam diferença significativa na expressão de *GmWRKY20* e letras maiúsculas diferentes indicam diferença significativa na expressão de *GmWRKY46* (ANOVA / Tukey,  $p < 0,05$ ). Com exceção da vagem, a comparação dos níveis de expressão entre os genes para cada órgão analisado não apresentou diferença significativa (teste-T de Student,  $p < 0,05$ ). Os genes “F-box protein” e “metalloprotease” foram utilizados como controles endógenos da reação.

#### **4.2.2. Infecção por *P. pachyrhizi***

O perfil de expressão de *GmWRKY20* e *GmWRKY46* obtido após a inoculação de plantas de soja com *P. pachyrhizi* é mostrado na Figura 13. Os resultados obtidos indicam que estes genes são modulados em resposta à inoculação com o patógeno (comparando-se com a situação controle) e que a modulação da expressão dos mesmos ocorre de forma diferencial entre o genótipo resistente (PI561356) e o genótipo suscetível (Embrapa-48), ao longo do período analisado.

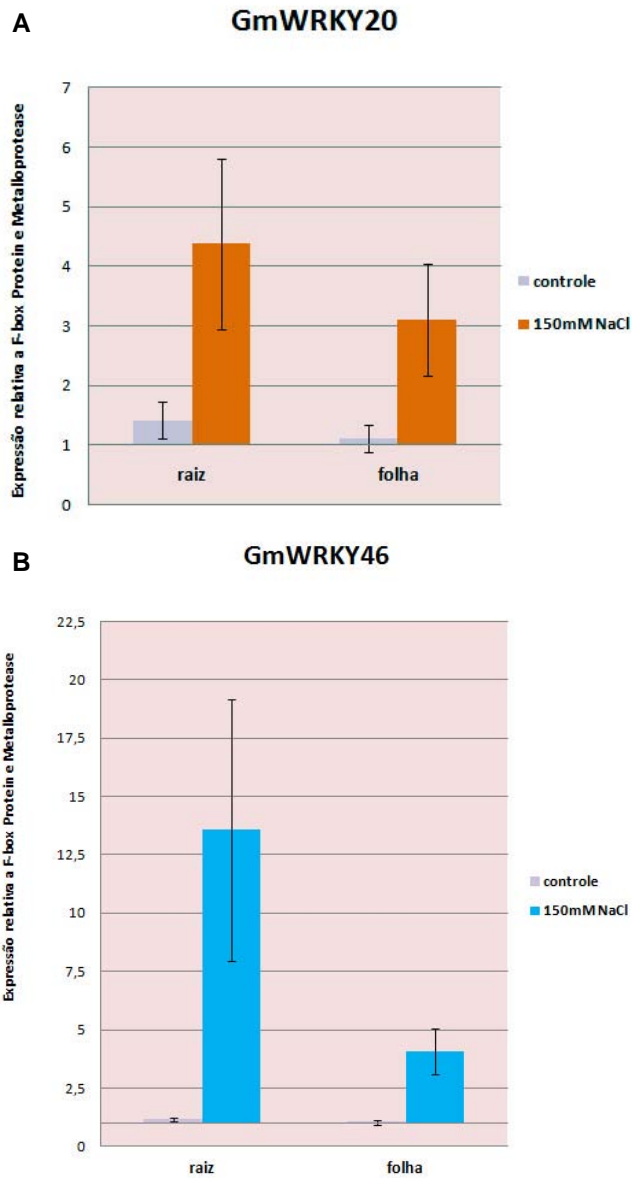


**Figura 13. Expressão relativa de *GmWRKY20* (A) e *GmWRKY46* (B) em resposta à infecção por *P. pachyrhizi*.** Nos gráficos são apresentados os níveis de expressão obtidos em folhas de plantas inoculadas com o patógeno, quando comparados a plantas controle (“mock-inoculated”). Diferenças significativas (teste-T de Student,  $p < 0,05$ ) são representadas por asteriscos (\*). Entre as plantas inoculadas, ao longo do período analisado, pode-se observar uma modulação diferencial da expressão destes genes nas plantas portando o genótipo resistente (PI 561356), se comparadas às portadoras do

genótipo suscetível (Embrapa-48). Os genes “F-box protein” e “metalloprotease” foram utilizados como controles endógenos da reação.

#### **4.2.3. Estresse salino**

Os resultados obtidos nas análises de expressão de *GmWRKY20* e *GmWRKY46* em resposta ao estresse salino (Figura 14) sugerem que estes genes possam ser induzidos na presença de 150 mM de NaCl (comparando-se com a situação controle), e que esta indução ocorre tanto em folhas quanto em raízes. Contudo, devido à alta variação nos níveis de expressão apresentada pelas amostras, os resultados obtidos não foram estatisticamente significantes.



**Figura 14. Expressão relativa de *GmWRKY20* (A) e *GmWRKY46* (B) em resposta ao estresse salino.** Nos gráficos são apresentados os níveis de expressão obtidos em raízes e folhas de soja, 6 h após serem submetidas ao estresse salino (150 mM de NaCl), comparados às plantas controle (“mock”). Devido ao grande desvio apresentado pelas amostras, a indução visualizada nos gráficos não é estatisticamente significativa (teste-T de Student,  $p > 0,05$ ).

### 4.3. Obtenção de vetores para a transformação de soja

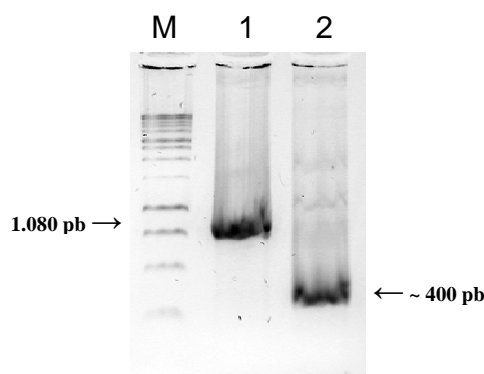
Tendo em vista a superexpressão e o silenciamento de plantas de soja, os fragmentos de DNA previamente isolados e clonados no vetor pGEM<sup>®</sup>-TEasy (Promega) foram transferidos para vetores de expressão em planta.

#### 4.3.1. Construção do vetor binário de superexpressão

Devido à alta similaridade entre as seqüências de aminoácidos de ambos os genes em estudo, apenas a seqüência codificante completa de *GmWRKY20*, contendo 1.080 pb, foi utilizada para a superexpressão. Através da tecnologia Gateway<sup>®</sup> (Invitrogen), esta foi primeiramente clonada no vetor de entrada pENTR Directional TOPO<sup>®</sup>, sendo a seguir recombinada no vetor de destino pH7WG2D,1 (Figura 15). O produto da recombinação foi inserido em *E. coli* Top 10 por choque térmico e selecionado pela adição de estreptomicina e espectinomicina ao meio de cultura bacteriano. Uma única colônia resistente à seleção foi obtida. Esta foi submetida à PCR, sendo a transformação confirmada pela amplificação do fragmento referente a *GmWRKY20* (1.080 pb) e do fragmento referente ao gene *hpt* (~400 pb) (Figura 16).



**Figura 15. Representação gráfica da construção para a superexpressão de *GmWRKY20*.** Na figura está representada apenas a região do T-DNA do vetor pH7WG2D,1, após a recombinação com a seqüência codificante completa de *GmWRKY20*. As setas horizontais (em cinza) indicam os “primers” (F, “forward” e R, “reverse”) utilizados para a confirmação da recombinação e para as subseqüentes análises moleculares em *A. tumefaciens* e nas plantas transformadas. RB, borda direita do T-DNA (“right border”); LB, borda esquerda do T-DNA (“left border”); EgfpER, proteína GFP contendo sinal de direcionamento do retículo endoplasmático (“enhanced green fluorescent protein linked to endoplasmatic reticulum targeting signal”); *hpt*, higromicina fosfotransferase.



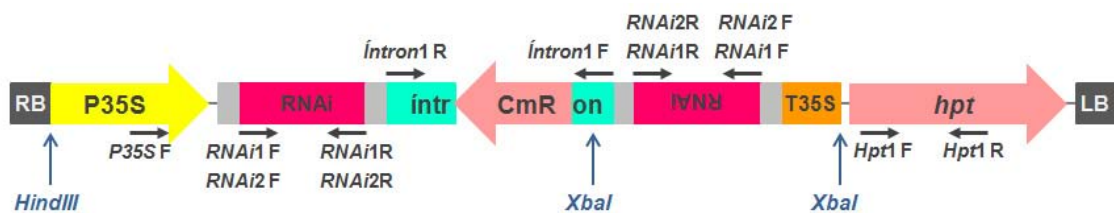
**Figura 16. Confirmação da transformação de *E. coli* com a construção pH7WG2D,1 + *GmWRKY20*.** Uma colônia isolada de *E. coli*, resistente à seleção pelos antibióticos estreptomicina e espectinomicina, foi submetida à PCR e os produtos desta reação foram analisados em gel de agarose 0,8%, após eletroforese. M, marcador de peso molecular 1 kb DNA Ladder (Promega); 1, amplificação com os “primers” *GmWRKY20* (1.080 pb); 2, amplificação com os “primers” *Hpt1* (~400pb).

#### 4.3.2. Construção dos vetores de silenciamento

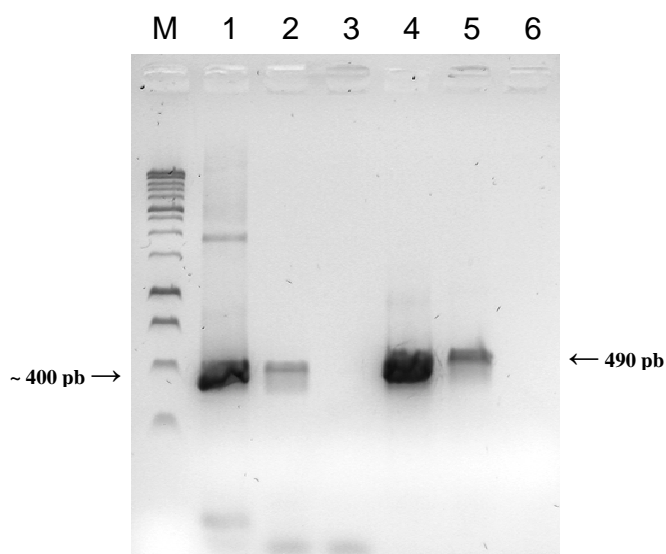
Dois fragmentos foram utilizados para as construções de silenciamento por RNAi: um maior, possuindo 490 pb e outro menor, com 249 pb. Ambos possuem seqüências homólogas exclusivas a *GmWRKY20* e *GmWRKY46* (Anexo 1), de forma a impedir o silenciamento de outros membros da família. Entretanto, o silenciamento de cada um dos genes em separado não deve ocorrer, devido à alta similaridade entre suas regiões codificantes. Estes foram primeiramente clonados no vetor de entrada pENTR Directional TOPO<sup>®</sup>, para em seguida serem recombinados no vetor de destino pH7GWIWG2(II),0 (Figura 17). Os produtos destas recombinações também foram inseridos em *E. coli* Top 10 por choque térmico e selecionados pela adição de estreptomicina, espectinomicina e cloranfenicol ao meio de cultura bacteriano. Colônias resistentes à seleção foram submetidas à PCR e as transformações foram confirmadas pela amplificação dos fragmentos *RNAi1* (490 pb), *hpt* (~400 pb) (Figura 18) e *RNAi2* (dados não mostrados). A inserção dos fragmentos de forma correta no vetor de destino foi confirmada de duas maneiras:



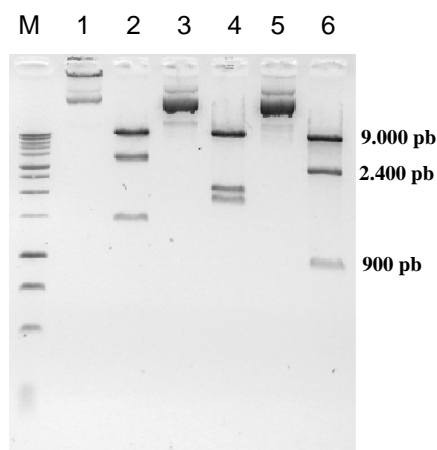
1. Pela utilização combinada (ver Figura 17) dos seguintes “primers” na reação de PCR: *RNAi1(2) F* x *Íntron F*, *RNAi1(2) F* x *Íntron R*, ocorrendo amplificação dessas bandas; além de *RNAi1(2) R* x *Íntron F* e *RNAi1(2) R* x *Íntron R*, as quais não geraram amplificação, como esperado (resultados não mostrados);
2. Além disso, realizou-se extração de plasmídeos (Miniprep) seguida pela clivagem com enzimas de restrição (*HindIII* e *XbaI*) dos vetores obtidos (Figura 19): como esperado, na clivagem do plasmídeo selvagem [pH7GWIWG2(II),0] foram liberadas bandas de tamanho aproximado de 9.000 pb, 3.400 pb e 1.900 pb. Já a clivagem do plasmídeo recombinante [pH7GWIWG2(II),0 + *RNAi1*] gerou fragmentos de aproximadamente 9.000 pb, 2.400 pb e 900 pb, como esperado. Uma colônia falso-positiva (resistente à seleção pelos antibióticos, mas que teve resultado negativo na PCR) também foi utilizada, para que seu padrão de clivagem fosse comparado.



**Figura 17. Representação gráfica da construção para o silenciamento de *GmWRKY20* e *GmWRKY46*.** Na figura está representada apenas a região do T-DNA do vetor pH7GWIWG2(II),0. Dois fragmentos de regiões comuns a *GmWRKY20* e *GmWRKY46* foram clonados, gerando duas construções (*RNAi1* e *RNAi2*), as quais são úteis para produzir o silenciamento simultâneo de ambos os genes. Os locais de inserção de ambos os fragmentos (de forma direta e inversa) estão representados em rosa. As setas horizontais (em cinza) indicam os “primers” utilizados para a confirmação da recombinação e das transformações subsequentes (de *A. tumefaciens* e também de planta). As setas verticais (em azul) indicam os sítios de restrição presentes na sequência e as respectivas enzimas utilizadas para as clivagens descritas no texto. RB, borda direita do T-DNA (“right border”); LB, borda esquerda do T-DNA (“left border”); CmR, gene de resistência ao cloranfenicol; *hpt*, higromicina fosfotransferase.



**Figura 18. Confirmação da transformação de *E. coli* com a construção pH7GWIWG2(II),0 + RNAi1.** Colônias isoladas de *E. coli*, resistentes à seleção pelos antibióticos estreptomicina, espectinomomicina e cloranfenicol, foram submetidas à PCR e os produtos da reação foram analisados em gel de agarose 1%, após eletroforese. M, marcador de peso molecular 1 kb DNA Ladder (Promega); 1-3, amplificação com os “primers” *Hpt1* (~400 pb), onde 1 é o controle positivo (amplificação de DNA plasmidial), 2 e 3 são colônias (positiva e falso-positiva, respectivamente); 4-6, amplificação com os “primers” *RNAi1* (490 pb), onde 4 é o controle positivo (amplificação de DNA plasmidial), 5 e 6 são colônias (positiva e falso-positiva, respectivamente).



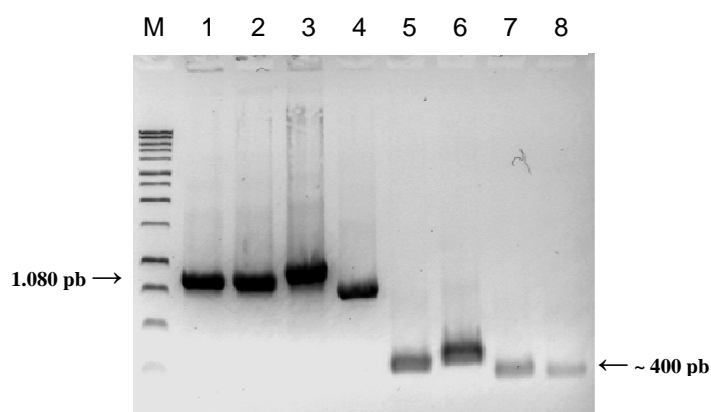
**Figura 19. Confirmação da transformação e da inserção adequada do fragmento RNAi1 no vetor de silenciamento pH7GWIWG2(II),0.** Após a PCR, os plasmídeos das colônias isoladas foram

extraídos por “Miniprep”, clivados com as enzimas *Hind*III e *Xba*I ( por 3 h cada) e submetidos a eletroforese em gel de agarose 1,5%. A confirmação da recombinação de forma adequada foi realizada pela análise do tamanho das bandas liberadas após a clivagem. M, marcador de peso molecular 1 kb DNA Ladder (Promega); 1, 3 e 5, plasmídeos não-clivados (1, selvagem; 3, falso-positivo e 5, positivo); 2,4 e 6, plasmídeos clivados (2, selvagem; 4, falso-positivo e 6, positivo).

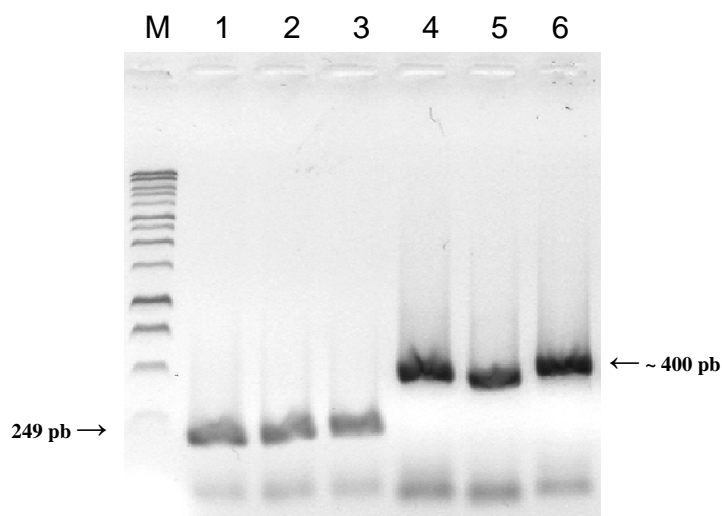
#### 4.4. Obtenção de plantas transgênicas

##### 4.4.1. Transformação de *Agrobacterium tumefaciens* com as construções de superexpressão e silenciamento

Células competentes da cepa desarmada LBA4404 de *Agrobacterium tumefaciens* foram transformadas com as construções obtidas por eletroporação. Colônias recombinantes foram selecionadas pela adição de rifampicina, estreptomicina e espectinomicina ao meio de cultura e submetidas à PCR para a confirmação da transformação: nas colônias contendo os vetores recombinantes foram amplificadas bandas referentes a *GmWRKY20* (1.080 pb) e *htp* (~400 pb) na construção de superexpressão (Figura 20); e nas construções de silenciamento, a *htp* (~400 pb) e a seus respectivos fragmentos RNAi (*RNAi* 1 – dados não mostrados; ou *RNAi* 2 – Figura 21).



**Figura 20. Confirmação da transformação de *A. tumefaciens* com a construção pH7WG2D,1 + *GmWRKY20*.** Colônias isoladas de *A. tumefaciens*, resistentes à seleção pelos antibióticos rifampicina, estreptomicina e espectinomicina foram submetidas à PCR e analisadas em gel de agarose 0,8 %, após eletroforese. M, marcador de peso molecular 1 kb DNA Ladder (Promega); 1-4, amplificação com os “primers” *GmWRKY20* (1.080 pb); 4-6, amplificação com os “primers” *Hpt1* (~ 400 pb).



**Figura 21. Confirmação da transformação de *A. tumefaciens* com a construção pH7GWIWG2(II),0 + RNAi2.** Colônias isoladas de *A. tumefaciens*, resistentes à seleção pelos antibióticos rifampicina, estreptomicina e espectinomicina foram submetidas à PCR e analisadas em gel de agarose 1%, após eletroforese. M, marcador de peso molecular 1 kb DNA Ladder (Promega); 1-3, amplificação com os “primers” *RNAi2* (249 pb); 4-6, amplificação com os “primers” *Hpt1* (~400pb).

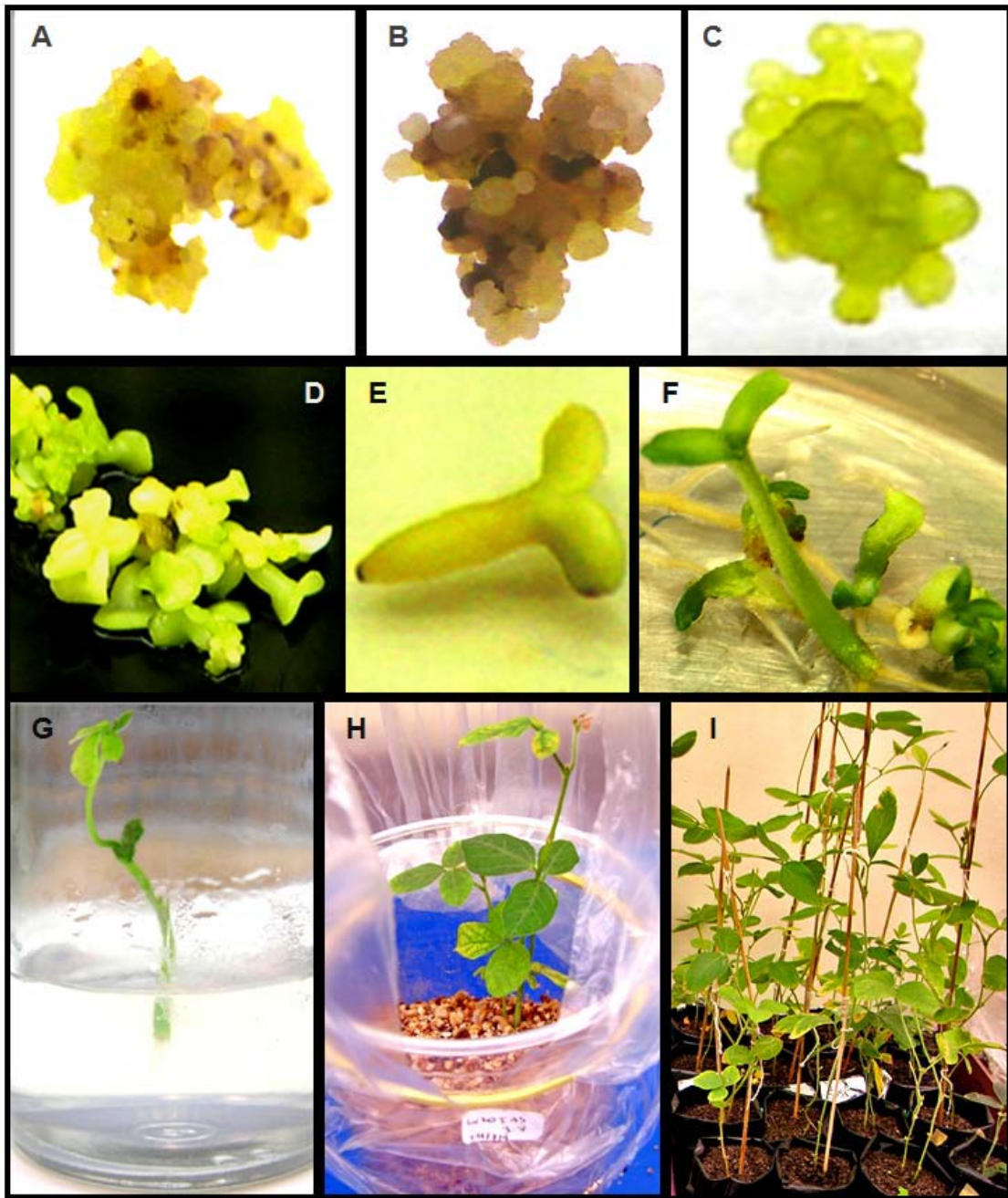
#### 4.4.2. Transformação de embriões somáticos de soja via sistema integrado “bombardeamento / *Agrobacterium*”

A transformação genética de soja foi realizada como descrito por Droste, Pasquali & Bodanese-Zanettini, em 2000. Conjuntos de embriões somáticos secundários em fase proliferativa (9 a 12 meses) sofreram bombardeamento por partículas de tungstênio livres de DNA em acelerador de partículas de baixa pressão (PIG; Finer et al., 1992); sendo a seguir, co-cultivados com suspensões de *A. tumefaciens* portando as três construções obtidas (um co-cultivo separado para cada construção). A eliminação da bactéria após o co-cultivo foi realizada como descrito por Wiebke et al., em 2006, adicionando-se doses adequadas de cefotaxima e vancomicina ao meio de cultivo. Além disso, para a construção de silenciamento pH7GWIWG2(II),0 + RNAi1, também foi realizado experimento de transformação através do bombardeamento de partículas de tungstênio carregadas com o

vetor recombinante. A seleção de embriões transformados foi realizada através da adição de doses crescentes de higromicina ao meio de cultivo. Conjuntos embriogênicos resistentes à higromicina (Tabela 7) passaram pelo processo previamente descrito (item 3.13.4) até a obtenção de plântulas regeneradas, sendo estas posteriormente transferidas para a vermiculita e finalmente para o solo — onde foram aclimatadas e mantidas em câmara de crescimento, visando à realização das análises moleculares subseqüentes. As diversas etapas transcorridas após a transformação dos embriões somáticos de soja podem ser visualizadas na Figura 22. Entre as plantas aclimatadas, 46 (representando ao menos 14 linhagens diferentes) foram transformadas via “bombardeamento / *A. tumefaciens*” com a construção de superexpressão, e 19 (representando ao menos 11 linhagens) foram transformadas via bombardeamento com a construção RNAi1. Plântulas portando as construções RNAi1 e RNAi2 , transformadas via “bombardeamento / *A. tumefaciens*”, ainda estão em fase de regeneração.

**Tabela 7. Conjuntos embriogênicos resistentes à seleção por higromicina.**

Cultivares / construções utilizadas	Experimento I (bombardeamento / <i>Agrobacterium</i> )		Experimento II (bombardeamento)		Experimento III (bombardeamento / <i>Agrobacterium</i> )		
	IAS5	Venc	IAS5	Venc	IAS5	Venc	Conq
pH7WG2D,1 + <i>GmWRKY20</i>	289	0	-	-	-	-	-
pH7GWIWG2(II),0 ++ RNAi1	-	-	144	7	17	3	0
pH7GWIWG2(II),0 + RNAi2	-	-	-	-	100	0	18



**Figura 22.** Etapas transcorridas após a transformação dos embriões somáticos de soja. A seleção dos transformantes teve início 10 dias após cada experimento de transformação, com a adição de  $12,5 \text{ mg.l}^{-1}$  de higromicina ao meio de cultivo vegetal. (A) Pontos de necrose (embriões não-resistentes) já podem ser observados algumas semanas após o início da seleção. (B) Passados 21 dias, os conjuntos embriogênicos foram transferidos para meio com dose aumentada de higromicina ( $25 \text{ mg.l}^{-1}$ ),

permanecendo neste até que o período de seleção completasse 3 meses. (C) Conjuntos de embriões resistentes à higromicina (pontos verdes) foram então destacados do tecido necrosado e proliferados por três meses, visando à multiplicação dos eventos de transformação obtidos. (D) Os conjuntos embriogênicos proliferados foram então submetidos aos processos de histodiferenciação e maturação, sendo transferidos para meio de cultivo contendo 6% de sacarose e 1% de carvão ativado por 30 dias; (E) por mais 30 dias, os embriões permaneceram em meio contendo 6% de sacarose, até que estivessem totalmente maduros. (F) Estes foram então submetidos à dessecação, na qual permaneceram por 6 h em placas sem meio de cultivo, sendo logo em seguida transferidos para meio de conversão em plântulas. (G) Ao surgimento das primeiras raízes e folhas nas plântulas obtidas, estas foram transferidas para frascos individuais contendo meio de cultivo adequado para sua regeneração. (H e I) Depois de regeneradas, as plântulas obtidas seguiram para aclimação, primeiramente em copos contendo vermiculita e cobertos com plástico (H), sendo então finalmente transferidas para solo orgânico (I), permitindo o completo desenvolvimento das plantas obtidas.

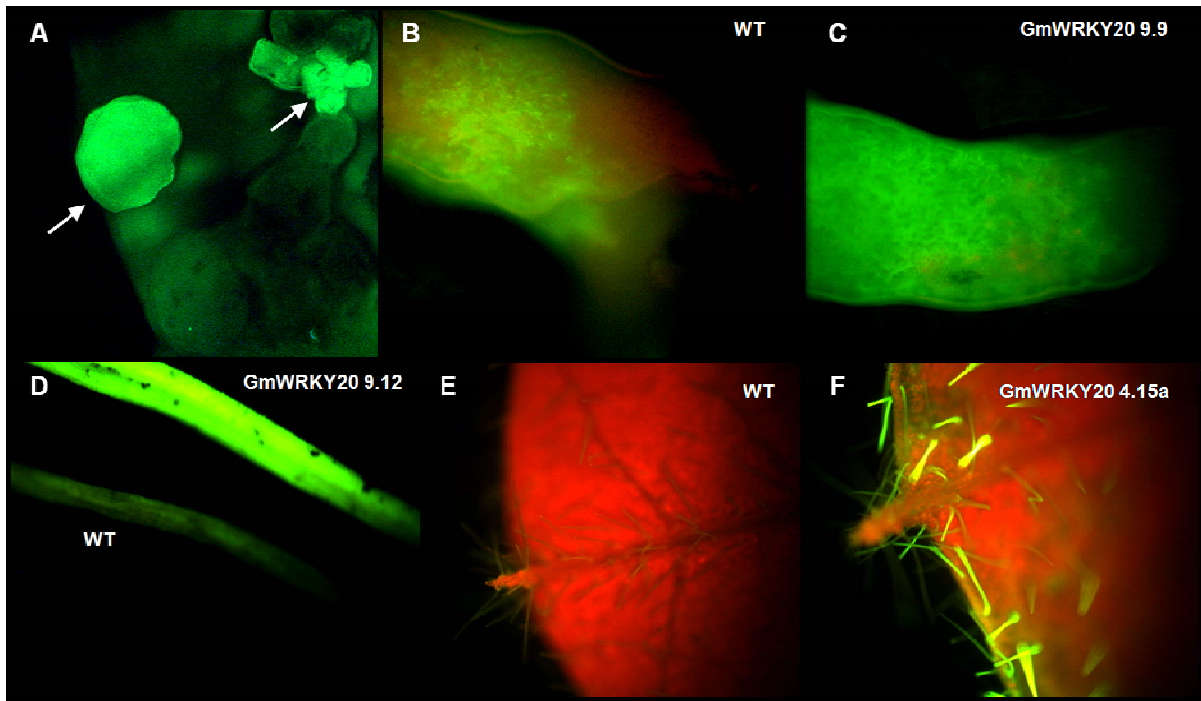
#### **4.5. Análise dos tecidos geneticamente modificados e das plantas transgênicas**

##### **4.5.1. Monitoramento da integração do transgene através da análise da expressão do gene repórter**

A integração do transgene ao genoma vegetal pôde ser monitorada através da visualização da expressão GFP – em microscopia de fluorescência, sob luz azul.

Na etapa de seleção por higromicina, um mês após a transformação, alguns embriões já apresentavam a expressão GFP (Figura 23). Infelizmente, nas etapas seguintes (proliferação e histodiferenciação), sua visualização se tornou dificultada pelo excesso de clorofila que é característico aos embriões nestas fases. A seguir, na etapa de maturação, a expressão GFP pôde ser mais uma vez observada; entretanto esta foi dificultada pelo alto nível de fluorescência basal apresentado por embriões selvagens (WT), aliado ao acúmulo de clorofila presente em algumas regiões dos embriões analisados. Desta maneira, evitou-se prosseguir com a análise em órgãos e tecidos clorofilados. Esta foi retomada à medida que as plântulas foram aclimatadas, quando foi possível observar níveis diversos de sua expressão em raízes – principalmente nas extremidades – e em tricomas foliares – principalmente em folhas com menor acúmulo de clorofila. Entretanto, à medida que a

análise foi sendo realizada nos tecidos de plantas mais maduras, a dificuldade de visualização da expressão GFP foi aumentada (dados não mostrados).



**Figura 23. Monitoramento da integração do transgene pela análise da expressão GFP.** (A) No período de seleção por higromicina, um mês após o processo de transformação genética, já era possível identificar a expressão do gene repórter em alguns embriões resistentes (setas), apesar da grande fluorescência basal apresentada pelo tecido utilizado. (B e C) A expressão GFP só pôde ser observada de novo ao final do estágio de maturação dos embriões, devido ao alto acúmulo de clorofila nas etapas anteriores; mesmo assim, a alta fluorescência basal apresentada pelo tecido não transformado (B), mais uma vez dificultou a comparação deste com o tecido transformado (C). Finalmente, após a aclimação das plântulas obtidas, a expressão GFP pôde ser visualizada com mais clareza em órgãos aclorofilados. Em D, pode-se observar a diferença de fluorescência apresentada entre uma raiz transformada e outra selvagem; já em E e F pode-se comparar o nível de fluorescência dos tricomas foliares de planta selvagem (E) e transformada (F). A visualização dos tecidos foi realizada em microscópio de fluorescência sob aumento de 40 vezes.



#### **4.5.2. Confirmação da integração do transgene por PCR**

As análises moleculares visando a confirmação da integração do transgene nas diversas plantas obtidas estão em andamento. Como as seqüências-alvo dos genes em estudo se tratam de fragmentos endógenos, pertencentes ao genoma da soja, sua amplificação não é indicativa do sucesso dos experimentos de transformação realizados. Dessa forma, a confirmação deve ser realizada pela amplificação de outros fragmentos presentes no T-DNA, como o marcador de seleção (*hpt*), o promotor 35S ou o íntron responsável pela formação do “hairpin” (no caso das construções de silenciamento). Entretanto, os “primers” que haviam sido previamente utilizados para a confirmação das transformações em bactéria não apresentaram resultados satisfatórios quando utilizados nas amostras vegetais, resultando na amplificação de diversos fragmentos inespecíficos. Assim, devido a estas dificuldades iniciais, alterações estão sendo feitas visando à obtenção de maior especificidade nas reações de PCR, além de maior pureza no processo de extração de DNA.

#### **4.5.3. Cálculo do número de cópias do transgene através de RT-qPCR**

A partir dos resultados obtidos no monitoramento da expressão do gene-repórter (*gfp*) nas plantas decorrentes da transformação para a superexpressão de *GmWRKY20*, a estimativa do número de cópias do transgene foi realizada através de RT-qPCR, pelo método de “quantificação relativa pela análise da curva padrão” (Shou et al., 2004; Anexo 2). A expressão “número de cópias”, neste caso se refere ao número de alelos que são detectados para cada gene no genoma diplóide. O gene *Le1* (lectina – 4 cópias) foi utilizado para eliminar os efeitos decorrentes da variação na quantidade de material apresentado inicialmente entre as amostras testadas. Plantas não-transformadas (WT) foram utilizadas para a calibração da reação. Como é esperado que seja encontrada uma cópia endógena de *GmWRKY20* no genoma da soja (2 alelos), no cálculo atribuiu-se aos calibradores o valor de 2 cópias. Até o momento, foram analisadas apenas as linhagens 4.15, 6.9, 7.4 e 9.9 (nas quais uma única inserção do transgene foi identificada), além da linhagem 9.12, onde duas

inserções foram detectadas. Os dados obtidos até o momento estão apresentados na Tabela 8.

**Tabela 8. Estimativa do número de cópias do transgene.**

<b>Linhagem</b>	<b>Quantidade relativa de DNA (<i>GmWRKY20/Le1</i>)<sup>†</sup></b>	<b>Número de cópias de <i>GmWRKY20</i><sup>+</sup></b>	<b>Número de inserções do transgene<sup>**</sup></b>
IAS5 WT	0,5	2,0	0
IAS5 <i>GmWRKY20</i> 4.15	0,8098	3,2394	1
IAS5 <i>GmWRKY20</i> 6.19	0,8061	3,2247	1
IAS5 <i>GmWRKY20</i> 7.4	0,7241	2,8966	1
IAS5 <i>GmWRKY20</i> 9.9	0,7733	3,0935	1
IAS-5 <i>GmWRKY20</i> 9.12	1,0130	4,0522	2

<sup>†</sup> A quantidade relativa de DNA representa a quantidade de *GmWRKY20* relativa à quantidade de *Le1*.

<sup>\*</sup> O número de cópias foi calculado a partir da quantidade relativa de DNA obtida para cada linhagem dividida pelo valor obtido para a planta WT, presumindo-se que o número de cópias de *GmWRKY20* nesta seja igual a 2.

<sup>\*\*</sup> O número de inserções do transgene é uma estimativa calculada a partir do número de cópias obtido para cada linhagem, presumindo-se que o valor referente às cópias endógenas de *GmWRKY20* seja igual a 2.

## 5. DISCUSSÃO

### 5.1. Análises *in silico* e isolamento dos genes em estudo

Em 2003, um grupo de pesquisa pertencente ao “Institute of Genetics and Developmental Biology” da “Chinese Academy of Sciences” (Tian et al. 2004) construiu uma biblioteca de ESTs de soja em resposta ao ácido salicílico. Através da análise dos clones desta biblioteca, juntamente com a análise de ESTs de soja previamente depositadas no GenBank, o grupo realizou a montagem em “contigs” (UniGene) destas seqüências e a anotação funcional dos dados obtidos de acordo com o Gene Ontology (<http://www.geneontology.org>), visando sua comparação com os genomas de *A.thaliana* e *Medicago truncatula*. Mais tarde, em 2007, procurando por genes envolvidos nas respostas da soja a estresses abióticos, o mesmo grupo realizou um “screening” (*in silico*) nos “contigs” formados, utilizando como sonda motivos conservados de fatores de transcrição WRKY. Como resultado, 64 genes WRKY de soja foram identificados e sua anotação foi depositada no NCBI (Tian, Zhang & Chen, 2007 – artigo não-publicado; Zhou et al. 2008). “Primers” foram desenhados visando a obtenção das seqüências codificantes (totais, ou em muitos casos, parciais) dos genes identificados e utilizando-se apenas as regiões contendo os motivos conservados – previamente utilizados como sondas – uma análise filogenética (“Neighbour-Joining”) desses genes foi realizada. Além disso, seus níveis de expressão em diversos órgãos da planta e diversas situações de estresses abióticos foram analisados por RT-PCR.

As seqüências codificantes parciais de *GmWRKY20* e *GmWRKY46* foram depositadas no GenBank através dos trabalhos citados acima. Como primeiro passo para a caracterização funcional destes genes, que é o objetivo maior do presente trabalho, tornaram-se necessárias tanto a obtenção de suas seqüências codificantes completas, quanto a análise *in silico* destas e de suas respectivas seqüências genômicas. Esta análise estrutural permitiu que se chegasse às seguintes conclusões:

- 1°. As seqüências parciais depositadas no GenBank foram anotadas de forma incorreta em 2007, uma vez que representam (com alguns erros de seqüenciamento), as porções amino- e carboxi-terminal da mesma proteína;
- 2°. Este erro é decorrente da montagem incorreta dos “contigs” referentes a esses genes, realizada em 2003, uma vez que ESTs de ambos foram agrupadas no mesmo “contig” (*GmWRKY46*); o qual provavelmente ocorreu devido à alta similaridade apresentada entre *GmWRKY20* e *GmWRKY46*;
- 3°. Assim, o grupo da “Chinese Academy of Sciences” desenhou primers para apenas um dos genes (*GmWRKY20*), conseqüentemente as duas seqüências codificantes parciais obtidas representam as porções 5’ e 3’ deste;
- 4°. Estas seqüências foram utilizadas tanto para as análises filogenéticas quanto para análises iniciais de expressão gênica realizadas pelo grupo, de forma que os resultados obtidos nessas análises devem ser, no mínimo, re-avaliados.

No presente trabalho, os genes selecionados para estudo foram isolados de forma bem sucedida. Suas seqüências codificantes completas foram obtidas (Figura 10), o que possibilitou a identificação das principais diferenças estruturais entre *GmWRKY20* e *GmWRKY46* (Figura 11); além da classificação destes no grupo III da superfamília (Tabela 6).

Além disso, os dados estruturais obtidos sobre os genes em estudo indicam que estes tenham sido alvo de um evento de duplicação recente, pois apesar de estarem situados em cromossomos diferentes, suas seqüências codificantes possuem alto grau de identidade. De acordo com análises recentemente publicadas, obtidas a partir do seqüenciamento do genoma da soja (Schmutz et al. 2010), foi inferido que este é do tipo paleopoliplóide, tendo sofrido dois grandes eventos de duplicação: estima-se que o primeiro evento tenha ocorrido há 59 Myr, estando relacionado à origem da linhagem *Papilionoideae*; já a segunda duplicação, denominada “*Glycine*-específica”, ocorreu há provavelmente 13 Myr. Ressalta-se ainda, que a retenção de blocos homólogos nesta espécie é excepcionalmente alta,

podendo envolver não apenas dois, mas vários cromossomos. Além do mais, eventos de duplicação gênica são inerentes à própria história evolutiva dos fatores de transcrição WRKY e em especial, à sua expansão de forma gradual dentro do Reino Vegetal: um maior número de membros é encontrado entre as Fanerógamas, provavelmente refletindo a maior adaptabilidade destas ao ambiente (Zhang et al. 2005).

## 5.2. Análises da expressão gênica

### 5.2.1. Genes-referência

A quantificação da expressão de *GmWRKY20* e *GmWRKY46* em diversas situações foi realizada por RT-qPCR. Este tipo de análise requer a normalização dos valores obtidos contra genes-referência que são utilizados como controles internos da reação, partindo-se do pressuposto de que a expressão destes não varia ao longo do experimento analisado. Devido a sua função em processos celulares básicos, genes como *actina*, *tubulina* e *fator de alongação 1- $\alpha$*  (*eF1- $\alpha$* ) são geralmente escolhidos como genes-referência, muitas vezes sem ter seu nível de expressão previamente avaliados nas condições analisadas, o que acaba por interferir na clareza e na precisão dos resultados obtidos para os genes em estudo.

Em 2008, Libault et al. buscaram por genes que pudessem ser utilizados como referência em análises de expressão gênica de soja em resposta a condições de estresse biótico: primeiramente através de busca por genes com expressão constitutiva em análises de microarranjo disponíveis; seguida da quantificação (RT-qPCR) dos níveis de expressão do conjunto de genes pré-selecionados em novos experimentos envolvendo estresse biótico. Os dados obtidos foram avaliados no geNorm (<http://medgen.ugent.be/~jvdesomp/genorm/>), que é um “software” desenvolvido para indicar os genes mais adequados a serem utilizados como referência em cada experimento, através do cálculo da *variação média entre pares* ( $M$ ) de um gene em relação aos outros (maior valor  $M$  representa menor estabilidade); além de gerar um ranqueamento da *estabilidade de expressão* dos genes testados e calcular o número necessário de genes a serem utilizados em cada experimento, para uma normalização acurada da expressão

gênica. Como resultado deste trabalho, quatro novos genes referência foram indicados para a normalização das análises de expressão gênica em soja: “*ATP-binding cassette (ABC) transporter*”, “*F-box protein Family*”, “*metalloprotease*” e “*CDPK-related protein kinase*” (Libault et al. 2008). Os três últimos foram escolhidos para serem testados nas análises de expressão de *GmWRKY20* e *GmWRKY46*, juntamente com os genes *ciclofilina (CYP2)*, *actina (ACT II)* e *fator de alongação eucariótico 1-α (ELF1A)* (Jian et al. 2008). Seus níveis de expressão foram medidos nas condições experimentais realizadas e os valores obtidos foram analisados no GeNorm. A utilização de dois genes-referência foi indicada, sendo “*F-box protein Family*” e “*metalloprotease*” os genes que apresentaram maior estabilidade de expressão nessas condições (dados não mostrados).

#### **5.2.2. Expressão em diferentes órgãos, fases do desenvolvimento e em resposta ao estresse salino**

Como já mencionado, Zhou et al. publicaram em 2008, um artigo com análises referentes aos 64 genes WRKY de soja identificados a partir de uma biblioteca de ESTs. Entre as análises realizadas, inclui-se a quantificação da expressão dos genes WRKY identificados – através de RT-PCR (não-quantitativo) – em caules, cotilédones, folhas e raízes de plantas de soja com duas semanas de idade; além de plantas da mesma idade submetidas a estresse salino (200 mM de NaCl), estresse por seca e por frio (4°C). Os resultados obtidos demonstraram expressão relativamente baixa de *GmWRKY20* e *GmWRKY46* (quando comparados aos outros WRKYs) em todos os órgãos analisados, havendo leve diferença entre o padrão de expressão de *GmWRKY20* e *GmWRKY46* – o que deve refletir a inacurácia da técnica utilizada, uma vez que os “primers” desenhados para a detecção de ambos os genes são referentes a duas regiões distintas de *GmWRKY20* apenas. Em nossas análises, tanto o padrão quanto o nível de expressão entre ambos os genes se mostrou similar, sendo que o maior nível de expressão encontrado foi no caule de plantas maduras. Mais amostras devem ser analisadas (como amostras de caule de plantas jovens, por exemplo) para que um perfil de expressão mais amplo destes genes seja obtido, possibilitando relacioná-lo aos processos biológicos nos quais podem estar envolvidos.

Além disso, na análise de estresse salino realizada por Zhou et al. em 2008, nenhuma mudança significativa foi observada para os genes em questão. No presente trabalho, condições similares foram utilizadas no experimento (150 mM de NaCl, plantas com duas semanas de idade) e o resultado obtido sugere que não há indução significativa de ambos os genes na referida condições. Entretanto, tanto nas amostras de raiz como de folhas utilizadas, pode-se observar uma tendência de maior expressão destes genes nas plantas tratadas com NaCl. Desta forma, novos experimentos nestas condições devem ser realizados visando minimizar as variações nos níveis de expressão apresentadas entre as replicatas biológicas, permitindo assim, determinar se as diferenças obtidas entre plantas tratadas e não-tratadas são de fato significativas.

### **5.2.3. Infecção por *P. pachyrhizi***

Os dados obtidos no microarranjo realizado por van De Mortel et al. em 2007, permitiram aos autores inferir que um padrão bifásico de expressão é apresentado por grande parte dos genes envolvidos na defesa contra o patógeno causador da ferrugem asiática. Nossos resultados confirmam este perfil bifásico de expressão para os genes *GmWRKY20* e *GmWRKY46* no genótipo resistente (*Rpp2*, PI561356) em resposta à infecção por *P. pachyrhizi*; apesar de este não refletir o perfil exato obtido para cada um dos dois genes no microarranjo. Essas diferenças apresentadas podem ser decorrentes das diferenças de sensibilidade entre as técnicas utilizadas – em geral RT-qPCR possui maior acurácia para medir a expressão gênica individual, enquanto que o microarranjo, como ferramenta de análise em larga escala, é extremamente útil para retratar mudanças globais no transcriptoma.

Em teoria, o perfil bifásico apresentado no microarranjo seria decorrente da formação de dois “picos” de respostas de defesa após a inoculação com o patógeno: o primeiro seria reflexo da ativação de mecanismos basais de defesa da planta, ocorrendo em geral, tanto no genótipo resistente quanto no suscetível, até 24 hpi. Após este período, entre 24 e 48 hpi (que corresponde ao período de desenvolvimento de hifas secundárias e formação de haustórios), a maioria dos genes volta a níveis basais de expressão,

possivelmente como resultado da liberação pelo patógeno de moléculas efetoras. No genótipo resistente, foi observado um segundo pico de expressão de genes de defesa em 72 hpi, o qual só ocorreu no genótipo suscetível de 1 a 2 dias depois. Essa segunda indução pode ser decorrente da ativação de *genes R*, resultando na expressão de proteínas R responsáveis pelo reconhecimento dos efetores liberados pelo patógeno (*Avr*, interação incompatível) e conseqüente inativação dos mesmos, o que resulta na restrição do crescimento do patógeno e na prevenção do desenvolvimento da infecção.

Em 2008, Choi et al., após análise de microarranjo comparando as respostas obtidas na reação imune (governada pelo gene *Rpp1*) em relação à reação suscetível, também observaram a indução inicial (de 6 a 12 hpi) de fatores de transcrição WRKY, seguida pela sua repressão (de 24 a 48 hpi), corroborando com os resultados aqui apresentados. Além disso, estes autores complementam a hipótese inferida por van De Mortel et al., 2007, sugerindo que a resistência à ferrugem é mais uma questão de tempo e intensidade de ativação de vias da imunidade inata, do que o resultado da indução de alguns poucos genes específicos. Desta forma, a imunidade ativada por efetores (ETI) seria uma re-ativação da imunidade ativada por PAMPs (PTI), apenas de uma maneira acelerada e potencializada. Neste contexto, a reprogramação transcricional mediada por fatores de transcrição da família WRKY – retratada tanto nas análises de microarranjo citadas, quanto neste trabalho – se mostra crucial na modulação dos mecanismos de defesa contra *P. pachyrhizi*.

#### **5.2.4. Estrutura, expressão e função gênica**

Como a regulação da expressão de um gene em condições diversas é decorrente do reconhecimento de motivos conservados (*cis*-elementos) presentes em sua região promotora, e de sua conseqüente interação com outros genes através destes *cis*-elementos, um alto grau de homologia entre as seqüências codificantes de dois genes não necessariamente quer dizer que estes possuam o mesmo perfil de expressão, ou participem das mesmas vias regulatórias. É fato que funções redundantes entre fatores de transcrição WRKY são comumente encontradas: através de técnicas de genômica funcional, diversos grupos de pesquisa relatam a co-regulação destes e a contribuição de vários genes WRKY

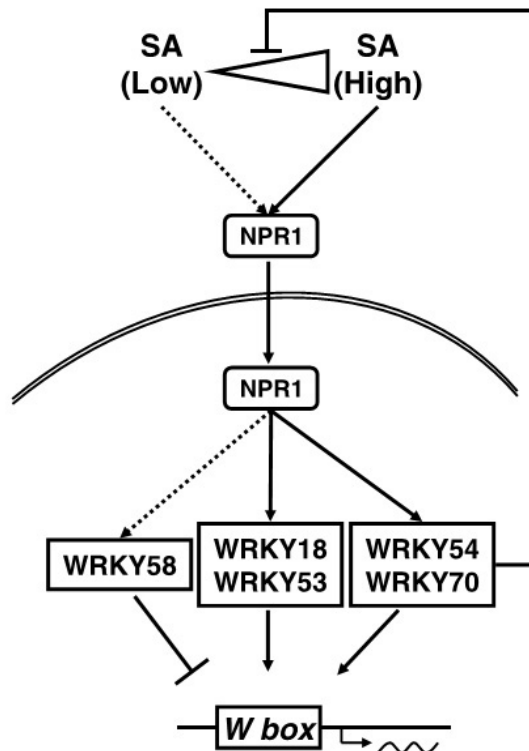


na composição de um fenótipo. Por outro lado, genes WRKY altamente homólogos já foram descritos como possuindo funções diversas em determinada via regulatória, como reportado por Tao et al., em 2009. Este grupo demonstrou que o par alélico de genes WRKY de arroz (*Oryza sativa*) *OsWRKY45-1* e *OsWRKY45-2*, cujas seqüências protéicas diferem em apenas 10 aminoácidos, possuem papéis distintos nos mecanismos de resistência a patógenos: tanto a superexpressão de *OsWRKY45-1*, quanto o silenciamento de *OsWRKY45-2* resultaram em maior susceptibilidade de plantas de arroz a *Xanthomonas oryzae*, e de forma inversa, o silenciamento de *OsWRKY45-1* e a superexpressão de *OsWRKY45-2* resultaram na resistência contra este patógeno. Além disso, diferenças no balanço de fitormônios envolvidos nestas respostas também foram observadas, levando os autores à análise da região promotora destes alelos a fim de obterem respostas sobre sua modulação diferencial.

Mesmo que uma análise aprofundada dos motivos que compõem a região regulatória de *GmWRKY20* e *GmWRKY46* ainda não tenha sido realizada, os resultados obtidos nas diversas análises de expressão gênica realizadas demonstraram que, ao menos nas condições testadas, ambos os genes apresentam perfis de expressão extremamente similares. A partir destes resultados, pode-se pressupor que os genes em estudo estejam provavelmente sob ação de forte pressão seletiva, uma vez que eventos de duplicação gênica geralmente levam a processos de perda de função da cópia extra; formação de pseudogenes; silenciamento epigenético ou ainda à sub-funcionalização entre os homólogos, com conseqüente formação de genes parálogos. A retenção de dois genes WRKY altamente homólogos foi descrita por Petitot et al., em 2008: *CaWRKY1a* e *CaWRKY1b* foram isolados de uma biblioteca de ESTs construída na resposta do café (*Coffea arabica*) ao fungo *Hemileia vastatrix*, patógeno responsável pela ferrugem da folha nesta espécie. Os níveis de expressão destes genes foram quantificados (RT-qPCR) em diversas situações, como em folhas senescentes, no tratamento com SA e em resposta ao fungo *H. vastatrix*, além do nematódeo *Meloidogyne exiguus*; sendo que ambos os genes apresentaram perfis de expressão similares em todos os experimentos realizados. Uma explicação para a recorrente retenção da redundância funcional entre genes WRKY é que

esta sirva para proteger o organismo de efeitos deletérios eventualmente causados pela mutação ou deleção gênica (Zhang et al. 2005).

Utilizando-se a ferramenta GOST – GreenPhyl Orthologs Search Tool (<http://greenphyl.cirad.fr/cgi-bin/gost.cgi>), foi possível identificar os genes *AtWRKY53* e *AtWRKY30* como sendo os prováveis ortólogos em *A. thaliana* dos genes de soja em estudo. Nesta espécie, foi demonstrado que *AtWRKY53* – mas não *AtWRKY30* – é rapidamente induzido após o tratamento de plantas com SA (Kalde et al., 2003); nesta via, relacionada à defesa contra patógenos biotróficos, *AtWRKY53* é alvo direto de NPR1 e atua ativando genes de defesa no núcleo, no desenvolvimento da SAR (Figura 24) (Wang et al. 2006); o gene também é induzido em plantas inoculadas com *Blumeria graminis* (patógeno de gramíneas), estando aparentemente envolvido na resistência não-hospedeira. Além do mais, em experimentos com o patógeno *Ralstonia solanacearum* foi demonstrado que *AtWRKY53* também está envolvido em processos relacionados à senescência e à sinalização mediada por ABA, havendo aparente sobreposição entre os genes envolvidos nestas vias (Hu et al. 2008). Já *AtWRKY70* – outro membro do grupo III – além de atuar como ativador da defesa sinalizada por SA (possuindo função redundante a *AtWRKY53* na indução da SAR), também previne o acúmulo excessivo da molécula-sinal nesta via, de maneira a agir na regulação entre as vias mediadas por SA (biotróficos) e JA (necrotróficos) (Wang, Amornsiripanitch & Dong, 2006). Em contraste, também fazem parte do grupo III em *A. thaliana*, os fatores de transcrição *AtWRKY38* e *AtWRKY62*, que possuem papel na repressão da defesa basal mediada por SA em resposta a *P. syringae*: estes também são alvos de NPR1, mas atuam como ativadores transcricionais de genes repressores da defesa. Por outro lado, o reconhecimento do patógeno pela planta pode induzir a ação de HDA19, uma histona-desacetilase que resulta na formação de um complexo de co-repressão, tornando regiões eucromáticas inacessíveis a estes fatores de transcrição, conseqüentemente permitindo a ativação de genes de defesa contra o patógeno em questão (Kim et al., 2008).



**Figura 24. Representação esquemática da regulação da SAR por fatores de transcrição WRKY induzidos por NPR1, em *A. thaliana*.** O acúmulo de SA ativa o transporte de NPR1 para o núcleo, resultando na indução de diversos genes *WRKY*: quando os níveis de SA estão baixos, *WRKY58* funciona na prevenção da ativação espúria da SAR; já quando estes estão altos, a sinalização através de fatores de transcrição *WRKY* ativadores (como *WRKY18*, 53, 54 e 70) sobrepõe o efeito repressor de *WRKY58*, ativando a transcrição de genes-alvo. Além disso, *WRKY70* e *WRKY54* também atuam na prevenção do acúmulo excessivo de SA na célula (Modificado de Wang, Amornsiripanitch & Dong, 2006)

Recentemente, Yang et al. (2009) isolaram 46 fatores de transcrição *WRKY* de canola (*Brassica napus*), uma espécie economicamente importante e com filogenia próxima à *A. thaliana* – ambas pertencem à família *Brassicaceae* – cuja seqüência completa do genoma ainda não está disponível ao público. Os genes isolados foram anotados de acordo com seus ortólogos na espécie-modelo e tiveram seus níveis de expressão analisados na infecção com os patógenos *Sclerotinia sclerotiorum* e *Alternaria brassicae*, e em diversos tratamentos com fitormônios. Curiosamente, os genes *BnWRKY53* e *BnWRKY70* tiveram

seu nível de expressão aumentados em resposta a ambos os patógenos necrotróficos, além de *BnWRKY53* ser fortemente reprimido no tratamento com ABA; pelos trabalhos já realizados com seus respectivos ortólogos em *A. thaliana*, um perfil oposto seria esperado para estes.

### **5.3. Obtenção e análises de plantas transgênicas para o estudo funcional de *GmWRKY20* e *GmWRKY46***

Partindo-se do pressuposto de que a superexpressão de ambos os genes em estudo, sob o controle do mesmo promotor constitutivo, resultaria no desenvolvimento do mesmo fenótipo em plantas de soja (devido a sua alta similaridade de seqüência), apenas *GmWRKY20* foi utilizado nos processos de transformação visando a superexpressão desse gene em soja. Já nas construções de silenciamento utilizadas, os fragmentos clonados de *GmWRKY20* possuem alto nível de identidade com a *GmWRKY46*, desta forma é esperado que os níveis de expressão de ambos os genes sejam reduzidos nas plantas obtidas para estas construções, o que pode ser muito útil para revelar funções gênicas redundantes. Ressalta-se ainda que a principal diferença entre as construções RNAi1 e RNAi2 é que o fragmento clonado da primeira construção é composto, além da seqüência codificante, por parte de um íntron; assim é esperado que esta construção gere um silenciamento mais inespecífico entre genes WRKY do que a segunda, possivelmente sendo capaz de diminuir a expressão de *GmWRKY12*, que possui similaridade de seqüência com *GmWRKY20* e *GmWRKY46*.

#### **5.3.1. Expressão GFP**

Em 1994, Chalfie et al. demonstraram pela primeira vez, que uma proteína fluorescente verde (GFP) derivada de *Aequorea victoria* (popularmente conhecida como “água-viva”) ao ser expressa – tanto em células procarióticas quanto eucarióticas – é capaz de emitir forte fluorescência quando excitada por luz azul ou UV. Desde então, esta descoberta têm sido amplamente utilizada como ferramenta para vários tipos de análises, inclusive como uma poderosa ferramenta de monitoramento da expressão gênica (“gene -

repórter”) – sendo extremamente vantajosa em relação a outras técnicas disponíveis para este fim por não ser invasiva e não requerer a utilização de substratos ou co-fatores exógenos. Mais recentemente, a descoberta rendeu aos autores o Prêmio Nobel de Química em 2008.

Dois tipos de filtros podem ser utilizados para a visualização da expressão GFP: um deles é o “long pass filter”, que permite a passagem de comprimentos de onda maiores do que 500 nm, incluindo as fluorescências verde, vermelha e amarela. Na tentativa de eliminar a passagem destas outras fluorescências (vermelha e amarela) o “narrow-band pass filter” ( $525 \pm 25$  nm) é utilizado; entretanto, o resultado geralmente observado é um efeito aditivo destas à fluorescência verde, de forma a dificultar a discriminação entre a fluorescência verde basal e a expressão GFP.

Neste trabalho, o vetor binário de superexpressão utilizado possui – além do cassete de expressão do gene de interesse – um cassete de expressão da GFP, cuja finalidade é acompanhar o desenvolvimento das plantas transformadas, distinguindo-as das plantas não transformadas. Para a visualização da fluorescência o “long pass filter” foi utilizado, sendo o espectro de emissão analisado de 505 a 530 nm. Em decorrência disso, ressalta-se a grande dificuldade na detecção da expressão GFP nos tecidos de soja analisados, uma vez que os pigmentos fotossintéticos presentes na célula vegetal refletem essas fluorescências interferentes. Chama-se especial atenção à clorofila, que em altos níveis resulta na visualização de forte fluorescência vermelha. A dificuldade de visualização da expressão GFP por interferência da clorofila já foi reportada em algumas espécies, como no arroz (*O. sativa*), na leguminosa *M. truncatula* e em *A. thaliana* (apesar de nesta, ocorrer em menor grau) (Zhou et al., 2005); sendo mais recentemente, reportada em soja (Wu et al., 2008). Neste último, os autores propõem o tratamento de embriões somáticos de soja com o herbicida “Isoxaflutole (IFT; 5-cyclo-propyl-1,2-isoxazol-4-yl  $\alpha\alpha\alpha$ -trifluoro-2-mesyl p-tolyl ketone)”, que resulta na interrupção da via de síntese de carotenóides e na inibição do desenvolvimento cloroplastídico; tornando os tecidos vegetais “esbranquiçados” (“bleaching effect”). Além de eficiente em seu objetivo de facilitar a visualização da

expressão GFP em tecidos com altos níveis de clorofila, este processo se mostrou reversível: duas semanas após o final do tratamento, embriões somáticos re-cultivados em meio D20 já apresentavam coloração verde e não tiveram seu desenvolvimento comprometido. Apesar de não ter sido utilizada no presente trabalho, atualmente esta prática vem sendo testada em nosso laboratório (onde também se mostrou eficiente) e provavelmente será adotada nos próximos experimentos nos quais o monitoramento da expressão GFP é requerido.

Outro fator importante que parece ter interferido nas análises descritas (item 4.5.1), foi a virtual perda de expressão GFP em tecidos mais maduros de plantas cuja expressão já havia sido visualizada em etapas anteriores do desenvolvimento. O mesmo problema foi relatado por Zhou et al., em 2005, ao analisar plantas de *O. sativa*, *M. truncatula* e *A. thaliana*. Neste trabalho os autores demonstram que o aumento gradual da dificuldade de detecção ao longo do desenvolvimento das plantas está também relacionado à presença de clorofila (neste caso de forma quantitativa) já que foi observado que o aumento nos níveis de clorofila é acompanhado pela diminuição relativa da proteína GFP (em relação à composição protéica total da planta). Assim, os autores justificaram que o menor efeito observado em *A. thaliana* ocorre porque, nesta espécie, os níveis de clorofila são levemente diminuídos ao longo do desenvolvimento; além da proporção da proteína GFP permanecer relativamente constante ao longo deste, diferentemente do que foi observado em *O. sativa* e *M. truncatula*. Duas estratégias para a redução dos níveis de clorofila visando facilitar a visualização de GFP foram utilizadas: a “estiolação” de folhas (durante 6 dias) e o tratamento com etanol 95% (por no máximo 8 h). Ressalta-se ainda que, nessas análises, assim como foi observado nos tecidos de soja analisados no presente trabalho, níveis basais (mas detectáveis) de fluorescência foram encontrados em tecidos selvagens (não-transformados) de arroz, sendo denominados “autofluorescência”, que é o terceiro fator contribuinte para a dificuldade de detecção da expressão GFP nos diversos tecidos de soja aqui analisados.

O quarto fator envolvido nas variações observadas nos níveis de expressão GFP é o próprio promotor responsável pela regulação de sua expressão no cassete transferido para a planta: no vetor pH7WG2D,1, o promotor *rolD* está fusionado à seqüência codificante de *EgfpER* (“enhanced green fluorescent protein linked to endoplasmatic reticulum targeting signal”). Esta seqüência promotora (*rolD*, “rooting loci gene”) é proveniente do plasmídeo *Ri* (“root-inducing”) de *Agrobacterium rizhogenes* e sua capacidade de promover forte expressão, tanto em raiz como em partes aéreas de plantas transgênicas, já foi reportada em diversos trabalhos através de sua fusão a genes marcadores, como o *uidA* – que codifica a enzima  $\beta$ -glucuronidase (GUS) (Elmayan & Tepfer, 1995; Kamo & Blowers, 1999). Mais recentemente, em experimento de transformação transiente de soja via *A. rizhogenes* (“hairy-root transformation”), forte expressão GFP sob controle desse promotor foi demonstrada nos primórdios radiculares formados (Klink et al. 2008). Entretanto, em 2008, Wally, Jayaraj & Punja, ao comparar os níveis de expressão de GUS sob controle de diversos promotores em plantas transgênicas de cenoura (*Daucus carota*), observaram que a utilização do promotor *rolD*, assim como de *mas2* (“mannopine synthase”, de *A. tumefaciens*), resultou em níveis baixos de expressão GUS, em relação aos outros promotores testados: 35S de CaMV, duplo 35S (2x35S) e ubiquitina de *A. thaliana* (*UBQ3*). Além disso, os autores demonstraram ainda, que a expressão governada por *rolD* apresentou variações, tanto de forma espacial (os níveis obtidos em raízes foram muito mais altos que em folhas, onde a expressão pôde ser apenas observada nos feixes vasculares) quanto temporal (raízes de plantas maduras também apresentaram níveis muito baixos de expressão, se comparados ao tecido jovem). Este último padrão de expressão também parece ter ocorrido em nossas análises, uma vez que a expressão GFP fortemente observada em raízes de plantas transformadas também foi reduzida, à medida que estas plantas tornaram-se maduras.

Além disso, com relação à expressão GFP visualizada nos tricomas foliares, esta é provavelmente decorrente da falta de pigmentos apresentada por esses órgãos (fator já discutido anteriormente), uma vez que um padrão de expressão “tricoma-específico” não parece ser característico do promotor *rolD*; apesar de já ter sido relatado em alguns

trabalhos que utilizaram a sequência codificante de *gfp* fusionada a sequências diversas, como ferramenta na busca por ESTs ou *cis*-elementos envolvidos especificamente no desenvolvimento de tricomas e na resposta contra patógenos (Wu & Liu, 2005; Souza, 2006).

Por fim, vale lembrar que a superexpressão do gene de interesse (*GmWRKY20*) nas plantas transformadas obtidas neste trabalho é governada pelo promotor 35S, não possuindo desta forma, nenhuma relação com o padrão de expressão obtido para o gene-repórter analisado, sob regulação do promotor *rolD*.

### **5.3.2. Estimativa do número de cópias do transgene**

Por muitos anos, a detecção do número de cópias do transgene inseridas no genoma de um organismo geneticamente modificado era realizada exclusivamente através de técnicas baseadas em hibridização, como o “Southern blot”. A demanda por grandes quantidades de DNA para que a análise de uma única planta pudesse ser processada, aliada à demora (em média três dias) na obtenção do resultado e à falta de sensibilidade resultante de eventuais hibridizações inespecíficas estão entre os fatores responsáveis pela busca de técnicas novas e aprimoradas para a quantificação do DNA-alvo. Técnicas baseadas em PCR têm sido utilizadas; dentre elas, métodos de análise desenvolvidos a partir de PCR em tempo real quantitativo (RT-qPCR) estão entre os mais promissores. Entre as principais vantagens destes, citam-se fatores como a rapidez da técnica, as facilidades de automação e de escalonamento das análises (diversas amostras podem ser processadas em apenas uma placa); sem contar a alta sensibilidade e especificidade alcançadas pelo RT-qPCR – poucos nanogramas de DNA são necessários para o processamento de uma amostra, a detecção do gene-alvo é acurada mesmo na análise de genes pertencentes a grandes famílias homólogas e é possível distinguir diferenças pequenas na quantidade do gene-alvo (como uma “cópia” do transgene ou diferenças entre homo- e heterozigotos, por exemplo) (Schmidt & Parrott, 2001; Bubner & Baldwin, 2004; Koyama, et al., 2008).



O método utilizado em nossas análises se baseia na “Quantificação relativa pela análise da curva padrão” (Shou et al., 2004). Este foi escolhido por sua praticidade, já que a simples comparação entre as curvas padrão relativas obtidas para o gene-alvo e para o controle endógeno (cujo número de cópias é conhecido) são suficientes para se determinar quantas cópias do gene-alvo são detectadas na amostra utilizada (Bubner et al. 2004). Se por um lado não é necessária a quantificação absoluta da quantidade “física” do transgene (através de diluições seriadas do plasmídeo e quantificação exata da concentração e do peso molecular referentes a cada cópia); a escolha do calibrador (ou ao menos a determinação precisa do número de cópias do gene-alvo presentes neste) é fator crucial, especialmente quando a quantificação é feita levando-se em conta o número de cópias endógenas do gene em estudo, como é o nosso caso: a simples presunção do número de cópias apresentados pela planta utilizada como calibrador, sem que este valor seja de fato comprovado, pode resultar na inconsistência dos dados obtidos. Além do mais, os dados obtidos até o momento para o número de inserções de *GmWRKY20* são apenas uma estimativa, pois partem do pressuposto de que cada linhagem é homocigota (possui dois alelos) para o gene endógeno e cada cópia extra detectada é tratada como uma inserção do transgene. Uma idéia interessante talvez seja a detecção do número de inserções através da amplificação de um fragmento do cassete que não faça parte da região codificante de *GmWRKY20*, ou seja, que não esteja presente de forma endógena no genoma da soja. Esta idéia certamente será adotada como um complemento para a detecção do número de inserções do transgene nas plantas obtidas.

Sob essa perspectiva, pode-se concluir que a quantificação através de RT-qPCR é muito útil como estimativa do número de cópias do transgene, ainda mais quando diferentes genes endógenos e diferentes fragmentos do cassete de expressão são utilizados, garantindo alto grau de precisão dos resultados obtidos. Ainda assim, os resultados obtidos podem ser complementados através da realização de “Southern blot” e análises de segregação.

### 5.3.3. Revelação do fenótipo

Quando se têm por objetivo elucidar a função de fatores de transcrição WRKY em determinada via regulatória, através da utilização de técnicas que envolvam sua superexpressão e seu silenciamento gênico, ou ainda pela utilização de mutantes que apresentem perda de função, duas características inerentes aos membros desta superfamília devem ser levadas em consideração na hora de se avaliar o fenótipo obtido: a interação estrutural (formação de homo- ou heterocomplexos) e funcional (papéis redundantes ou antagônicos) entre os fatores de transcrição WRKY.

Um estudo bem sucedido demonstrando esta interação foi realizado por Xu et al., em 2006: *AtWRKY18*, *40* e *60* são induzidos por patógenos e codificam três proteínas estruturalmente similares. A interação física (estrutural) entre *AtWRKY18*, *40* e *60* foi demonstrada através de ensaios de duplo híbrido (“two-hybrid system”) e de co-immunoprecipitação e provavelmente está relacionada à obtenção de especificidade e seletividade necessárias para a ligação *in vivo* destes fatores de transcrição a seqüências-alvo envolvidas em vias distintas. Linhagens de *A. thaliana* apresentando expressão constitutiva, silenciamento ou perda de função de um destes genes (ou de combinações entre eles), foram avaliadas em resposta aos patógenos *P. syringae* e *Botrytis cinerea*, responsáveis pela ativação de vias distintas de sinalização na célula vegetal, sendo o primeiro considerado hemibiotrófico (envolvido na via mediada por SA) e o último necrotrófico (envolvido na via mediada por JA e ET). Nos ensaios com *P. syringae*, apesar de mutantes simples não terem apresentado diferença significativa de fenótipo, o nocaute dos três genes de forma combinada resultou em reduzido crescimento bacteriano (comparando-se a plantas WT), revelando assim sua redundância funcional nesta via. De forma intrigante, a superexpressão de *AtWRKY18* também resultou em maior resistência ao patógeno; ao contrário, a co-expressão deste com *AtWRKY40* ou *AtWRKY60* levou à maior suscetibilidade a *P. syringae*. Este resultado foi obtido possivelmente em decorrência de efeitos pleiotrópicos gerados pela superexpressão de *AtWRKY18* levando a ativação ou repressão de genes que não são geralmente regulados por este. A expressão combinada com

*AtWRKY40* ou *AtWRKY60* pode ter aumentado a especificidade de ligação de *AtWRKY18* a DNAs-alvos e conseqüentemente reduzido seus efeitos inespecíficos na expressão gênica; desta forma foi possível desvendar a real função destes como reguladores negativos da via mediada por SA.

Os resultados obtidos pelos autores nos ensaios com *B. cinerea* foram ainda mais intrigantes: tanto mutantes quanto linhagens superexpressando estes genes (de forma combinada) resultaram em maior suscetibilidade ao patógeno necrotrófico, indicando que o papel de genes WRKY nas vias de sinalização contra este tipo de patógeno seja mais complexo. Aparentemente, a resistência a *B. cinerea* é associada a determinados níveis de expressão dos três genes; dessa maneira, apesar de os mecanismos de sinalização ativados na perda de função de *AtWRKY18*, *40* e *60* e na sua superexpressão serem distintos, estes convergiram para o aumento da morte celular e conseqüente promoção da virulência do patógeno necrotrófico.

A partir deste e de outros exemplos encontrados na literatura, e levando-se em consideração os resultados obtidos nas análises de expressão de *GmWRKY20* e *GmWRKY46*; é possível que a superexpressão de *GmWRKY20* apenas, apesar de útil para a indicação de seu modo de ação na infecção com *P. pachyrhizi*, não seja suficiente no sentido de revelar sua real função nesta situação, sendo provavelmente necessária a análise conjunta com genes que apresentem potencial redundância funcional a este (como pode ser o caso de *GmWRKY46*). Se bem sucedidas, as linhagens obtidas com as construções RNAi podem ser úteis para este fim, uma vez que devem possibilitar o estudo dos efeitos gerados pelo silenciamento de ambos (e possivelmente de um terceiro gene WRKY, como já mencionado) em diversas situações, nas quais provavelmente a falta de um deles apenas não resultaria em alterações observáveis de fenótipo.

Com relação a alterações morfológicas apresentadas por plantas apresentando níveis alterados de expressão de genes WRKY, estas não são de ocorrência comum, salvo exceções (Wang, Amornsiripanitch & Dong, 2006), apesar desses genes estarem envolvidos em diversos processos biológicos, além das respostas a estresses. Algumas

variações morfológicas foram encontradas entre as linhagens de superexpressão obtidas neste trabalho (dados não mostrados); contudo, uma vez que as alterações não são recorrentes entre as diferentes linhagens, essas parecem ser decorrentes ou da variação somaclonal – freqüentemente observada em plântulas regeneradas a partir da cultura de tecidos (Kaeppler et al. 2000) – ou de efeito posicional (Matzke et al. 1998), uma vez que o local de inserção do transgene no genoma pode afetar a expressão de algum outro gene responsável por alguma alteração morfológica. Ressalta-se ainda, que superexpressão e o silenciamento dos genes em estudo nas plantas obtidas ainda não foram confirmados, sendo este mais um motivo pelo qual não é possível inferir que alguma variação morfológica obtida seja decorrente da alteração nos níveis de expressão de *GmWRKY20* e *GmWRKY46*.

## 6. CONCLUSÕES

Os dados obtidos neste trabalho nos permitiram chegar às seguintes conclusões:

- 1°. Os genes *GmWRKY20* e *GmWRKY46* codificam duas proteínas distintas, mas com alta similaridade estrutural, sendo provavelmente decorrentes de um evento recente de duplicação gênica;
- 2°. Devido a suas características estruturais, ambos pertencem ao grupo III da superfamília de fatores de transcrição WRKY;
- 3°. O perfil de expressão desses dois genes dá indícios de que eles estejam envolvidos na complexa rede de sinalização da soja em resposta ao ataque por patógenos, inclusive durante a infecção pela ferrugem asiática, além de participarem em outras vias de sinalização na resposta desta espécie contra estresses abióticos;
- 4°. A similaridade estrutural, aliada ao perfil de expressão obtido nas situações avaliadas, indicam que *GmWRKY20* e *GmWRKY46* possuam redundância funcional.

## 7. PERSPECTIVAS

- 1°. Repetições dos experimentos utilizados nas análises de expressão devem ser realizadas, com a adição de um maior número de indivíduos nestas análises;
- 2°. Novas análises de expressão devem ser realizadas, tanto de forma específica, nos processos de resposta a patógenos; quanto de forma mais ampla, nas vias de sinalização contra estresses abióticos e em outros processos fisiológicos da planta (como a senescência), para que um padrão de expressão mais amplo de *GmWRKY20* e *GmWRKY46* seja obtido;
- 3°. Análises moleculares que possibilitem a confirmação das plantas transgênicas obtidas devem ser feitas, assim como análises que confirmem que os genes em estudo possuam expressão modificada (superexpressão ou silenciamento) nestas;
- 4°. Plantas de soja superexpressando *GmWRKY20* e silenciadas para *GmWRKY20* e *GmWRKY46* poderão ser utilizadas em estudos funcionais e submetidas a ensaios biológicos com patógenos;
- 5°. Análise minuciosa da região regulatória de ambos os genes deve ser feita, além da realização de experimentos que permitam a identificação dos possíveis alvos dos fatores de transcrição em estudo, como ensaios de “pull-down” ou experimentos de microarranjo envolvendo plantas que estejam superexpressando *GmWRKY20*, por exemplo.

Combinadas, estas análises certamente serão úteis para se desvendar os papéis destes dois fatores de transcrição nas complexas vias de sinalização celular, especialmente nos mecanismos moleculares de defesa da soja contra patógenos. Além disso, a caracterização funcional destes genes na infecção por *P. pachyrhizi* poderá servir para o incremento dos programas de melhoramento genético de leguminosas, podendo eventualmente contribuir para a construção de um “arsenal” de fontes de resistência contra a ferrugem asiática no germoplasma comercial da soja.

## 8. REFERÊNCIAS

Bailey, M., Boerma, H., and Parrott, W. (1993). Genotype effects on proliferative embryogenesis and plant regeneration of soybean. *In Vitro Cell Dev Biol Plant* 29, 102–108.

Bent, A. F., and Mackey, D. (2007). Elicitors, effectors, and R genes: the new paradigm and a lifetime supply of questions. *Annual Review of Phytopathology* 45, 399–436. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17506648>.

Berri, S. et al. (2009). Characterization of WRKY co-regulatory networks in rice and Arabidopsis. *BMC Plant Biology* 9, 120. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19772648>.

Biodiesel Programa Nacional de Produção e Uso de Biodiesel. Available at: <http://www.biodiesel.gov.br/>.

Bubner, B., and Baldwin, I. T. (2004). Use of real-time PCR for determining copy number and zygosity in transgenic plants. *Plant Cell Reports* 23, 263–71.

Buchheim, J., Colburn, S., and Ranch, J. (1989). Maturation of soybean somatic embryos and the transition to plantlet growth. *Plant Physiology* 89, 768–775.

Calvo, É. S., Kiihl, R. A., Garcia, A., Harada, A., and Hiromoto, D. M. (2008). Two major recessive soybean genes conferring soybean rust resistance. *Crop Science* 48, 1350–1354.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science* 263, 802–805.

Choi, J., Alkharouf, N., Schneider, K., Matthews, B., and Frederick, R. (2008). Expression patterns in soybean resistant to *Phakopsora pachyrhizi* reveal the importance of peroxidases and lipoxygenases. *Functional and Integrative Genomics* 8, 341–359.

Clough, S. J., and Bent, A. F. (1999). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* 16, 735–743.

Doyle, J., and Doyle, J. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissues. *Phytochem Bull* 19, 11–15.

Droste, A., Pasquali, G., and Bodanese-Zanettini, M. H. (2000). Integrated bombardment and *Agrobacterium* transformation system: an alternative method for soybean transformation. *Plant Molecular Biology Reporter* 18, 51-59.

EMBRAPA Soja (2005). Doenças e medidas de controle. In *Sistemas de produção - região central do Brasil* (Londrina, PR: Embrapa Soja: Embrapa Cerrados: Embrapa Agropecuária Oeste: Fundação Meridional), p. 239.

Elmayan, T., and Tepfer, M. (1995). Evaluation in tobacco of the organ specificity and strength of the *thero1D* promoter, domain A of the 35S promoter and the 35S2 promoter. *Transgenic Research* 4, 388-396.

Eulgem, T., Rushton, P. J., Robatzek, S., and Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. *Trends in Plant Science* 5, 199-206.

Eulgem, T., and Somssich, I. E. (2007). Networks of WRKY transcription factors in defense signaling. *Current Opinion in Plant Biology* 10, 366-371.

Finer, J., Vain, P., Jones, M., and McMullen, M. (1992). Development of the particle inflow gun for DNA delivery to plant cells. *Plant Cell Rep* 11, 323-328.

Finer, J., and McMullen, M. (1991). Transformation of soybean via particle bombardment of embryogenic suspension culture tissue. *In Vitro Cell Dev Biol Plant* 27, 175-182.

Gamborg, O., Miller, R., and Ojima, K. (1968). Nutrient requirements of suspension cultures of soybean root cells. *Exp Cell Res* 50, 151-158.

Gepts, P., Beavis, W. D., Brummer, E. C., Shoemaker, R. C., Stalker, H. T., Weeden, N. F., and Young, N. D. (2005). Legumes as a model plant family. Genomics for food and feed report of the Cross-Legume Advances through Genomics conference. *Plant Physiology* 137, 1228-1235.

Gutterson, N., and Zhang, J. Z. (2004). Genomics applications to biotech traits: a revolution in progress? *Current Opinion in Plant Biology* 7, 226-230.

Göhre, V., and Robatzek, S. (2008). Breaking the barriers: microbial effector molecules subvert plant immunity. *Annual Review of Phytopathology* 46, 189-215.

Hartman, G. L., Miles, M. R., and Frederick, R. D. (2005). Breeding for resistance to soybean rust. *Plant Disease* 89, 664-666.



Hu, C., and Bodanese-Zanettini, M. H. (1995). Embryo culture and embryo rescue for wide cross hybrids. In *Plant Cell, Tissue and Organ Cultures: Fundamental Methods* O. Gamborg (Berlin: Springer), pp. 129-141.

Hu, J., Barlet, X., Deslandes, L., Hirsch, J., Feng, D. X., Somssich, I., and Marco, Y. (2008). Transcriptional responses of *Arabidopsis thaliana* during wilt disease caused by the soil-borne phytopathogenic bacterium, *Ralstonia solanacearum*. *PloS one* 3, e2589.

Hückelhoven, R. (2007). Cell wall-associated mechanisms of disease resistance and susceptibility. *Annual review of phytopathology* 45, 101-27.

Jian, B., Liu, B., Bi, Y., Hou, W., Wu, C., and Han, T. (2008). Validation of internal control for gene expression study in soybean by quantitative real-time PCR. *BMC molecular biology* 9, 59.

Kaeppler, S. M., Kaeppler, H. F., and Rhee, Y. (2000). Epigenetic aspects of somaclonal variation in plants. *Plant molecular biology* 43, 179-88.

Kalde, M., Barth, M., Somssich, I. E., and Lippok, B. (2003). Members of the *Arabidopsis* WRKY group III transcription factors are part of different plant defense signaling pathways. *Molecular Plant-Microbe Interactions* 16, 295-305.

Kamo, K., and Blowers, A. (1999). Tissue specificity and expression level of *gusA* under *rolD*, mannopine synthase and translation elongation factor 1 subunit promoters in transgenic *Gladiolus* plants. *Plant Cell Reports* 18, 809-815.

Karimi, M., Inzé, D., and Depicker, A. (2002). GATEWAY™ vectors for *Agrobacterium*-mediated plant transformation. *Trends in Plant Science* 7, 193-195.

Kim, K., Lai, Z., Fan, B., and Chen, Z. (2008). *Arabidopsis* WRKY38 and WRKY62 transcription factors interact with histone deacetylase 19 in basal defense. *The Plant Cell* 20, 2357-2371.

Klink, V. P., MacDonald, M. H., Martins, V. E., Park, S., Kim, K., Baek, S., and Matthews, B. F. (2008). MiniMax, a new diminutive *Glycine max* genotype with a rapid life cycle, embryogenic potential and transformation capabilities. *Plant Cell, Tissue and Organ Culture* 92, 183-195.

Koyama, H., Kihara, T., Takita, E., and Honda, M. (2008). Quantitative PCR in the analysis of transgenic plants. In *Handbook of new technologies for genetic improvement of legumes* P. Kirti (Boca Raton, Florida: CRC Press), pp. 459-466.

Kuang, H., Padmanabhan, C., and Li, F. (2009). Identification of miniature inverted-repeat transposable elements (MITEs) and biogenesis of their siRNAs in the Solanaceae: New functional implications for MITEs. *Genome Research* 19, 42-56.

Libault, M., Thibivilliers, S., Bilgin, D. D., Radwan, O., Benitez, M., Clough, S. J., and Stacey, G. (2008). Identification of four soybean reference genes for gene expression normalization. *The Plant Genome* 1, 44-54.

Liu, J., and Ekramoddoullah, A. (2009). Identification and characterization of the WRKY transcription factor family in *Pinus monticola*. *Genome* 52, 77-88.

Livak, K., and Schmittgen, T. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the  $2(-\Delta\Delta C(T))$ . *Methods*, 402-408.

Matzke, A. J., and Matzke, M. A. (1998). Position effects and epigenetic silencing of plant transgenes. *Current opinion in plant biology* 1, 142-1488.

Mazarei, M., Elling, A. A., Maier, T. R., Puthoff, D. P., and Baum, T. J. (2007). GmEREBP1 is a transcription factor activating defense genes in soybean and *Arabidopsis*. *Molecular Plant-Microbe Interactions* 20, 107-119.

Meyer, J. D. et al. (2009). Identification and analyses of candidate genes for Rpp4-mediated resistance to Asian Soybean Rust in soybean. *Plant Physiology* 150, 295-307.

Miles, M. R., Frederick, R. D., and Hartman, G. L. (2006). Evaluation of soybean germplasm for resistance to *Phakopsora pachyrhizi*. *Plant Health Progress*.

Miles, M., Hartman, G., and Fredrick, R. (2005). Update on sources of host resistance. USDA-Agricultural Research Service Department of Crop Sciences.

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol Plant* 15, 473-497.

Navarro, L., Dunoyer, P., Jay, F., Arnold, B., Dharmasiri, N., Estelle, M., Voinnet, O., and Jones, J. D. (2006). A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science* 312, 436-4399.

Navarro, L., Jay, F., Nomura, K., He, S. Y., and Voinnet, O. (2008). Suppression of the MicroRNA Pathway by Bacterial Effector Proteins. *Science* 321, 964 - 967.

Pandey, S. P., and Somssich, I. E. (2009). The role of WRKY transcription factors in plant immunity. *Plant physiology* 150, 1648-1655. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19420325>.

Panthee, D. R., Marois, J. J., Wright, D. L., Narváez, D., Yuan, J. S., and Stewart, C. N. (2009). Differential expression of genes in soybean in response to the causal agent of Asian soybean rust (*Phakopsora pachyrhizi* Sydow) is soybean growth stage-specific. *TAG. Theoretical and Applied Genetics* 118, 359-370. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18853130>.

Panthee, D. R., Yuan, J. S., Wright, D. L., Marois, J. J., Mailhot, D., and Stewart, C. N. (2007). 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* 7, 291-301.

Petitot, A., Lecouls, A., and Fernandez, D. (2008). Sub-genomic origin and regulation patterns of a duplicated WRKY gene in the allotetraploid species *Coffea arabica*. *Tree Genetics & Genomes* 4, 379-390.

Pugh, S. Asian Soybean Rust remains a concern in the 2006 crop year. USDA/Risk Management Agency. Available at: <http://www.rma.usda.gov/news/2006/04/soybeanrust.html>.

Schmidt, M., and Parrott, W. (2001). Quantitative detection of transgenes in soybean [*Glycine max* (L.) Merrill] and peanut (*Arachis hypogaea* L.) by real-time polymerase chain reaction. *Plant Cell Reports* 20, 422-428.

Schmutz, J. et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463.

Shou, H., Frame, B. R., Whitham, S. A., and Wang, K. (2004). Assessment of transgenic maize events produced by particle bombardment or *Agrobacterium*-mediated transformation. *Molecular Breeding* 13, 201-208.

Silva, D. C. et al. (2008). Molecular mapping of two loci that confer resistance to Asian rust in soybean. *Theoretical and Applied Genetics* 117, 57-63.

Somers, D. A., Samac, D. A., and Olhoft, P. M. (2003). Recent advances in legume transformation. *Plant Physiology* 131, 892-899.

Souza, R. G. (2006). Desenvolvimento de uma estratégia de fusão gênica visando à localização de proteínas em plantas. Dissertação. UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL, PROGRAMA DE PÓS-GRADUAÇÃO EM FITOTECNIA, 69.

Tao, Z., Liu, H., Qiu, D., Zhou, Y., Li, X., Xu, C., and Wang, S. (2009). A pair of allelic WRKY genes play opposite roles in rice-bacteria interactions. *Plant physiology* 151, 936-48.

Tian, A. et al. (2004). Characterization of soybean genomic features by analysis of its expressed sequence tags. *Theoretical and Applied Genetics* 108, 903-913.

Tsukahara, R. Y., Hikishima, M., and Canteri, M. G. (2008). Relações entre o clima e o progresso da ferrugem asiática (*Phakopsora pachyrhizi*) em duas micro-regiões do Estado do Paraná. *Semina: Ciências Agrárias* 29, 47-52.

Vain, P., McMullen, M., and Finer, J. (1993). Osmotic treatment enhances particle bombardment mediated transient and stable transformation of maize. *Plant Cell Rep* 12, 84-88.

Wally, O., Jayaraj, J., and Punja, Z. K. (2008). Comparative expression of beta-glucuronidase with five different promoters in transgenic carrot (*Daucus carota* L.) root and leaf tissues. *Plant cell reports* 27, 279-287.

Wang, D., Amornsiripanitch, N., and Dong, X. (2006). A genomic approach to identify regulatory nodes in the transcriptional network of systemic acquired resistance in plants. *PLoS pathogens* 2, e123.

Waterhouse, P. M., and Helliwell, C. A. (2003). Exploring plant genomes by RNA-induced gene silencing. *Nature reviews. Genetics* 4, 29-38.

Wiebke, B., Ferreira, F., Pasquali, G., Bodanese-Zanettini, M. H., and Droste, A. (2006). Influence of antibiotics on embryogenic tissue and *Agrobacterium tumefaciens* suppression in soybean genetic transformation. *Bragantia* 65, 543-551.

Wright, M., Launis, K., Novitzky, R., Duesiing, J., and Harms, C. (1991). A simple method for the recovery of multiple fertile plants from individual somatic embryos of soybean [*Glycine max* (L.) Merrill]. *In Vitro Cell Dev Biol Plant* 27, 153-157.

Wu, A., and Liu, J. (2005). Analysis of the Cotton E6 Promoter. *Tsinghua Science and Technology* 10, 409-413.

Wu, C., Chiera, J. M., Ling, P. P., and Finer, J. J. (2008). Isoxaflutole treatment leads to reversible tissue bleaching and allows for more effective detection of GFP in transgenic soybean tissues. *In Vitro Cell.Dev.Biol.-Plant* 44, 540-547.

Xu, X., Chen, C., Fan, B., and Chen, Z. (2006). Physical and functional interactions between pathogen-induced Arabidopsis WRKY18, WRKY40, and WRKY60 transcription factors. *Plant Cell* 18, 1310-1326.

Yang, B., Jiang, Y., Rahman, M. H., Deyholos, M. K., and Kav, N. N. (2009). Identification and expression analysis of WRKY transcription factor genes in canola

(*Brassica napus* L.) in response to fungal pathogens and hormone treatments. *BMC Plant Biology* 9, 68.

Yorinori, J. T., Paiva, W. M., Frederick, R. D., Costamilan, L. M., Bertagnolli, P. F., Hartman, G. E., Godoy, C. V., and Nunes, J. (2005). Epidemics of soybean rust (*Phakopsora pachyrhizi*) in Brazil and Paraguay from 2001 to 2003. *Plant Disease*, 675-677.

Zhang, L., Wang, X., Bi, Y., Zhang, C., Fan, Y., and Wang, L. (2008). Isolation and functional analysis of transcription factor GmWRKY57B from soybean. *Chinese Science Bulletin* 53, 3538-3545. Available at: <http://www.springerlink.com/index/10.1007/s11434-008-0483-2>.

Zhang, X., Dai, Y., Xiong, Y., Defraia, C., Li, J., Dong, X., and Mou, Z. (2007). Overexpression of Arabidopsis MAP kinase kinase 7 leads to activation of plant basal and systemic acquired resistance. *The Plant Journal* 52, 1066-1079.

Zhang, Y., and Wang, L. (2005). The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evolutionary Biology* 5, 1.

Zhang, Z., Wei, L., Zou, X., Tao, Y., Liu, Z., and Zheng, Y. (2008). Submergence-responsive microRNAs are potentially involved in the regulation of morphological and metabolic adaptations in maize root cells. *Annals of Botany* 102, 509 -519.

Zhou, Q. et al. (2008). Soybean WRKY-type transcription factor genes, GmWRKY13, GmWRKY21, and GmWRKY54, confer differential tolerance to abiotic stresses in transgenic Arabidopsis plants. *Plant Biotechnology Journal* 6, 486-503.

Zhou, X., Carranco, R., Vitha, S., and Hall, T. C. (2005). The dark side of green fluorescent protein. *New Phytologist* 168, 313-322.

van De Mortel, M., Recknor, J. C., Graham, M. A., Nettleton, D., Dittman, J. D., Nelson, R. T., Abdelnoor, R. V., Baum, T. J., and Whitham, S. A. (2007). Distinct biphasic mRNA changes in response to Asian Soybean Rust infection. *Molecular Plant-Microbe Interactions* 20, 887-899.

Ülker, B., and Somssich, I. E. (2004). WRKY transcription factors: From DNA binding towards biological function. *Current Opinion in Plant Biology* 7, 491-498.

### **ANEXO 1:**

**Seqüências genômicas e codificantes de *GmWRKY20* e *GmWRKY46* e alinhamento dos “primers” utilizados nas construções RNAi e para as análises de expressão por RT-qPCR.**

## Alinhamento das seqüências genômicas:

```

GmWRKY20      ATGGAGAGTGACTTGAGCTGGGAGCGGAACACACTCATCAATGAGCTAATTCAGGGGATG
GmWRKY46      ATGGAGAGTGACTTGAGCTGGGAGCAGAACACGCTCATAAATGAGCTAATTCAGGGGATG
*****

GmWRKY20      GAGGTAGCAAGGAAGTTGAAGGCTGACTTGAAGTTGCCGTATTCAGTTGACACAAGAGAT
GmWRKY46      GAGGTAGCAAGGAAGTTGAAGGCAGACTTGAGGATGCCGTATTCAGTTGACTCAAGAGAT
*****

GmWRKY20      CTGCTTCTGCAGAGGATATTAATCTTCTTACGAAAAGGCTCTACTGATTCTAAGATGCAGC
GmWRKY46      TTGCTTGTGCAGAGGATATTAATCTTCTTATGAAAAGGCTCTACTGATTCTAAGATGCAAT
*****

GmWRKY20      AATGCATCAACTTCCGAGTTGCAGGCCATGAATCAAGCAACACCAACTTTGCTACCCGAG
GmWRKY46      GCATCATCAACTTCCGAGTTGCAGGCCATGAGTCAAGCAACACCAACTTTGCTACCCGAG
*****

GmWRKY20      TCCCCATTATCTGTCCATGGAAGTCCTCTGCGCGAGGACGTTTCATGGGGCCATCAAGGAT
GmWRKY46      TCCCCATTATCTGTCCATGGAAGTCCTCTGCGCGAGGACGTTGATGGGACCATCAAGGAT
*****

GmWRKY20      CACCA-----TAATCAAAGAAAAGGTGAATATTACTGAACAGACTTGGTAGC
GmWRKY46      CACCAAGAAGTTAAACATGATTCAAAGAAAAGGTGAATTTTACTATACAGACTTGGTAGC
*****

GmWRKY20      TGTTCCTGGCACATAAACATCAGCTATGTCTGGTAAGTGGTTATACTATAGCTCCTAAC
GmWRKY46      AGTTTCTGCCACATAAAGATCAGCTATGTTGGTT-----CTACGGCTCCTAAC
*****

GmWRKY20      TTGTATTGTGTTTCTTTTTTAATTGCTTTCAGAAAGATAACGCCCAAATGGATGGATCG
GmWRKY46      T--ATTTGTGTTTCTTTTTT-AATTGCTTTCAGAAAGGCAACGCCCAAATGGATGGATCA
*

GmWRKY20      TGTAAGAGTGAGCTGTGAGAGTGGTCTTGAAGGACCACATGAAGATGGTTACAACGGAG
GmWRKY46      TGTAAGAGTGAGCTGTGAAAGTGGTCTTGAAGGACCACATGAAGATAGCTACAACGGAG
*****

GmWRKY20      AAAATATGGTCAGAAAGACATCCTTGGAGCCAAATATCCAGGTGAGAGAAGTTCCTCTG
GmWRKY46      AAAATATGGTCAGAAAGACATCCTTGGAGCCAAATATCCAGGTGAGAGAAGTTCCTCTG
*****

GmWRKY20      GTCTGCTTTA-AAAAGAATTACAAGTTTT-CTTAAGGATAAAATGCCTTCACCTTGGC
GmWRKY46      ATCCTGCTTTAGAAAAGAATCACAAGTTTTTCTTAAGGATAAAATGCCTTCACCTTGGC
*****

GmWRKY20      CAGCTTCTAATTAACGACTCATTGTGACAATTTGCAGAAGTTACTATAGGTGCACCTTCC
GmWRKY46      CAGCTTCTAATAAACTCATTTGTGACAATT-GCAGAAGTTACTATAGGTGCACCTTCC
*****

GmWRKY20      GCAGCACTCAGGGCTGTTGGGCAACAAAGCAAGTGCAGAGATCAGATGAAGATCCACAA
GmWRKY46      GCAACACTCAGGGCTGTTGGGCAACAAAGCAAGTGCAGAGATCAGATGAAGATCCACAG
***

GmWRKY20      TGTTGACATAACTTACAGAGGAAATCATACTGTTTCAGGGAAACAATGCCGTTCTGC
GmWRKY46      TATTTGACATAACTTACAGAGGCAAGCATACTGTTTCAGGGAAACAATGCCGTTCTGC
*

```

GmWRKY20 CACCCAAGTCACCAGAAAAACATGAGAAACCAGCACACAGTCATAACATTGACATTACC  
GmWRKY46 CACCTAAGTCACCAGAAAAACAAGAAAAACCAACACACAGTCATAACATTGATATTACC  
\*\*\*\*\*

GmWRKY20 ATGCACAGGCATCACAAGAAAGCCTTGCAAAGTTCAGAAGTATCTTGTCTGTCAACACAG  
GmWRKY46 GTGCACAGGCATCTCAAGAAAGCCTTACAAAGTTCAGAAACATCTTGTCTGTCAACACAG  
\*\*\*\*\*

GmWRKY20 ATAATCTGGACAATGGAGATATGGCATATGCTTTCACTTCCCTTCCACTTCATTGGAT  
GmWRKY46 ATAATCTGAACAATGGAGATATGGCATATCCTTTCACTTCCCTTCAACTTCATTGGAT  
\*\*\*\*\*

GmWRKY20 GCATGAAACAAGACAACCACAGCTTGATTCTTTGGCCCTGGAGAATGACTCCTTCTTAA  
GmWRKY46 GCATGAAACAAGACAACCACAGCTTGATTCTTTGGCCCTGGAGAACGAGTCTTCTTAA  
\*\*\*\*\*

GmWRKY20 GCGACCTATACCAAACACACCTGTTATCCCAACAACACCAGAATCAAATTATTTCCCAT  
GmWRKY46 GCGACCTATACCAAACACACCTTTTATCCACAACAATACCAGAATCAAATTTCCCAT  
\*\*\*\*\*

GmWRKY20 CTCCAACCTTTCCAGATGAATGAGTTTGATGGGATCTACAACAGGTCACATTCGAAATCCG  
GmWRKY46 CTCCAACCTTTCCAGATGAATGTGTTTGATGGGATCTACAGCAAGCCACATTCCGAATCCG  
\*\*\*\*\*

GmWRKY20 ATATCAATGAGATCATTTCCACCAACACATCAGCAACCAATTCTCCAATTCTGATTTCA  
GmWRKY46 ATATCAATGAGATCATTTCCACCAACACATCAGCAACCAATTCTCCAATTCTGATTTCA  
\*\*\*\*\*

GmWRKY20 ATTTCTCACTTGATCCGGTGGAAATTGATCCAAATTTCCCTTTCAATACCCCAGGATTTT  
GmWRKY46 ATTTCTCACTTGATCCAGTGGAAATTGACCCAAATTTCCCTTTCAATACCCCAGGACTTT  
\*\*\*\*\*

GmWRKY20 TGTGCTGA  
GmWRKY46 TCTCCTGA  
\* \* \*\*\*\*

Obs.: Os “primers” utilizados nas construções de silenciamento estão marcados em cores, segundo a legenda abaixo. As seqüências homólogas entre os dois genes, alvos das construções RNAi utilizadas estão sublinhadas.

Legenda:

RNAi1 For

RNAi2 For

RNAi1/2 Rev



### Alinhamento das seqüências codificantes:

```

GmWRKY20      ATGGAGAGTGACTTGAGCTGGGAGCGGAACACACTCATCAATGAGCTAATTCAGGGGATG
GmWRKY46      ATGGAGAGTGACTTGAGCTGGGAGCAGAACACGCTCATAAATGAGCTAATTCAGGGGATG
*****
GmWRKY20      GAGGTAGCAAGGAAGTTGAAGGCTGACTTGAAGTTGCCGTATTCAGTTGACACAAGAGAT
GmWRKY46      GAGGTAGCAAGGAAGTTGAAGGCAGACTTGAGGATGCCGTATTCAGTTGACTCAAGAGAT
*****
GmWRKY20      CTGCTTCTGCAGAGGATATTATCTTCTTACGAAAAGGCTCTACTGATTCTAAGATGCAGC
GmWRKY46      TTGCTTGTGCAGAGGATATTATCTTCTTATGAAAAGGCTCTACTGATTCTAAGATGCAAT
*****
GmWRKY20      AATGCATCAACTTCCGAGTTGCAGGGCATGAATCAAGCAACACCAACTTTGCTACCCGAG
GmWRKY46      GCATCATCAACTCCGAGTTGCAGGCCATGAGTCAAGCAACACCAACTTTGCTACCCGAG
*****
GmWRKY20      TCCCCATTATCTGTCCATGGAAGTCCTCTGCGCGAGGACGTTTATGGGGCCATCAAGGAT
GmWRKY46      TCCCCATTATCTGTCCATGGAAGTCCTCTGCGCGAGGACGTTGATGGGACCATCAAGGAT
*****
GmWRKY20      CACCA-----TAATTCAAAGAAAAGAGAGATAACGCCCAAATGGATGGATCGT
GmWRKY46      CACCAAGAAGTTAAACATGATTCAAAGAAAAGAAAGCAACGCCCAAATGGATGGATCAT
*****
GmWRKY20      GTAAGAGTGAGCTGTGAGAGTGGTCTTGAAGGACCACATGAAGATGGTTACAACCTGGAGA
GmWRKY46      GTAAGA GTGAGCTGTGAAAGTGGTCTTGAAGGACCACATGAAGATAGCTACAACCTGGAGA
*****
GmWRKY20      AAATATGGTCAGAAAAGACATCCTTGGAGCCAAATATCCCAGAAGTTACTATAGGTGCACC
GmWRKY46      AAATATGGTCAGAAAAGACATCCTTGGAGCCAAATATCCCAGAAGTTACTATAGGTGCACC
*****
GmWRKY20      TTCCGCAGCACTCAGGGCTGTGGGCAACAAAGCAAGTGCAGAGATCAGATGAAGATCCC
GmWRKY46      TTCCGCAACACTCAGGGCTGTGGGCAACAAAGCAAGTGCAGAGATCAGATGAAGATCCC
*****
GmWRKY20      ACAATGTTTGACATAAATTACAGAGGAAATCATACTGTCTCAGGGAAACAATGCCGTT
GmWRKY46      ACAGTATTTGACATAAATTACAGAGGCAAGCATACTGTCTCAGGGAAACAATGCCGTT
*** *
GmWRKY20      CTGCCACCAAGTCACCAGAAAACATGAGAAAACCAGCACACAGTCATAACATTGACATT
GmWRKY46      CTGCCACCTAAGTCACCAGAAAACAAGAAAACCAACACACAGTCATAACATTGATATT
*****
GmWRKY20      CACCATGCACAGGCATCACAAGAAAGCC TGC AAAAGTTCAGAAGTACTTGTCTGTCAAC
GmWRKY46      CACCGTGCACAGGCATCTCAAGAAAGCCTTACAAAAGTTCAGAAACATCTTGTCTGTCAAC
*** *
GmWRKY20      ACAGATAATCTGGACAATGGAGATATGGCATATGCTTTCACTTTCCCTTCCACTTCATT
GmWRKY46      ACAGATAATCTGAACAATGGAGATATGGCATATCCTTTCACTTTCCCTTCAACTTCATT
*****

```

GmWRKY20 GGATGCATGAAACAAGACAACCACAGCTTGATTTCCTTTGGCCCTGGAGAATGACTCCTTC  
GmWRKY46 GGATGCATGAAACAAGACAACCACAGCTTGATTTCCTTTGGCCCTGGAGAACGAGTCCTTC  
\*\*\*\*\*

GmWRKY20 TTAAGCGACCTATACCAAACACACCTGTTATCCCCAACAAACACCAGAATCAAATTATTTTC  
GmWRKY46 TTAAGCGACCTATACCAAACACACCTTTTATCCACAACAATACCAGAATCAAACCTATTTTC  
\*\*\*\*\*

GmWRKY20 CCATCTCCAACCTTCCAGATGAATGAGTTT **GATGGGATCTACAACAGGTCAC**ATTTCGAAA  
GmWRKY46 CCATCTCCAACCTTCCAGATGAATGTGTTTGATGGGATCTACAGCAAGCCACATTTCGGAA  
\*\*\*\*\*

GmWRKY20 TCCGATATCAATGAGATCATTTCACCAACACATCAGCAACCAATTCTCCAATTCCTGAT  
GmWRKY46 TCCGATATCAATGAGATCATTTCACCAACACATCAGCAACCAATTCTCCAATTCCTGAT  
\*\*\*\*\*

GmWRKY20 TTCAATTTCTCACTTGATCCGGTGGAAATTGATCCAAATTTCCCTTTCAATACCCAGGA  
GmWRKY46 TTCAATTTCTCACTTGATCCAGTGGAAATTGACCCAAATTTCCCTTTCAATACCCAGGA  
\*\*\*\*\*

GmWRKY20 TTTTGTGCTGA  
GmWRKY46 CTTTCTCCTGA  
\*\*\*\* \* \*\*\*\*

Obs.: Os “primers” utilizados nas reações de RT-qPCR estão marcados em cores, segundo a legenda abaixo. Os fragmentos amplificados para cada gene estão sublinhados.

Legenda:

**GmWRKY46 RT For**

**GmWRKY46 RT Rev**

**GmWRKY20 RT For**

**GmWRKY20 RT Rev**

## **ANEXO 2:**

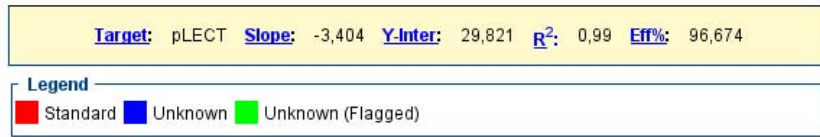
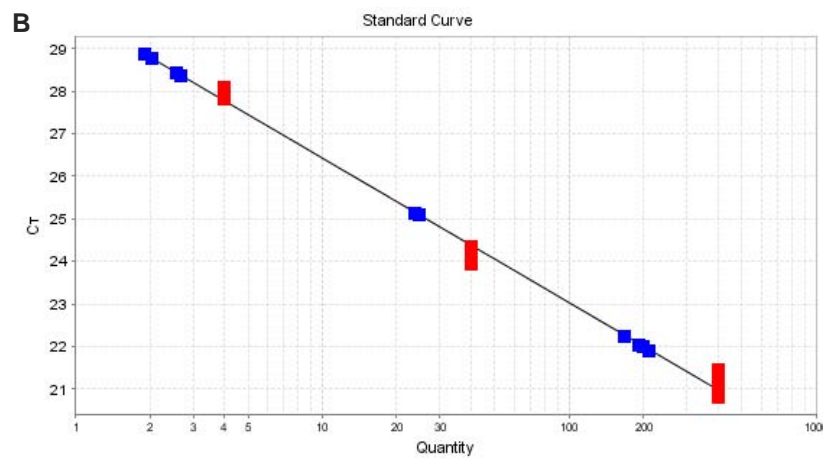
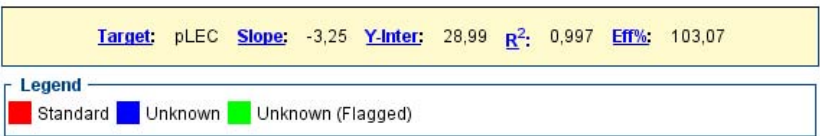
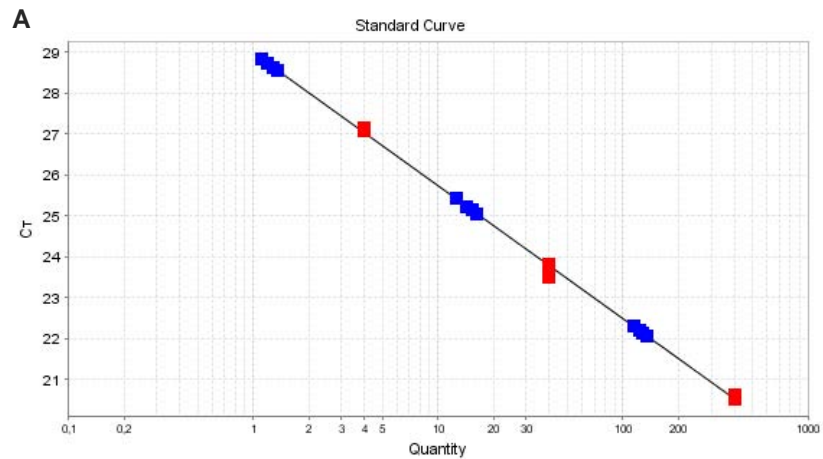
**Cálculo do número de cópias do transgene através da quantificação relativa pela análise da curva padrão.**

A seguir são descritos os passos realizados para o cálculo do número de cópias do transgene<sup>2</sup>:

- 1°. As amostras de DNA de plantas transgênicas, assim como de plantas WT, foram diluídas 100, 1.000 e 10.000 vezes e submetidas à RT-qPCR para a quantificação, tanto do gene endógeno (*Le1*) quanto do gene-alvo *GmWRKY20*;
- 2°. Os valores de Ct obtidos para ambos os genes foram comparados e, a partir destes, uma curva padrão relativa para cada amostra de DNA foi obtida (Figura 25);
- 3°. Foi feita uma média dos valores plotados na curva padrão para *GmWRKY20* (em azul), após estes serem multiplicados pelos respectivos valores de diluição. Esta foi utilizada para o cálculo do número de cópias: primeiramente, as diferenças de quantidade de “input” entre as amostras (WT e transgênica) foram normalizadas pela multiplicação do valor obtido na planta transgênica pelo “fator de correção”, o qual foi obtido pela relação entre o número de cópias de *Le1* inferidos na planta transgênica e na planta WT (Tabela 9);
- 4°. Uma segunda correção (regra de três simples) foi feita para minimizar as diferenças de eficiência apresentadas entre os “primers” utilizados (Tabela 9);
- 5°. Os dados corrigidos obtidos foram então comparados (Tabela 9): a amostra de uma planta WT foi utilizada para a calibração; o número de cópias foi determinado dividindo-se o valor obtido para a planta transgênica pelo valor obtido para o calibrador, sendo o resultado multiplicado por dois (presumindo-se que duas cópias do gene endógeno são esperadas no genoma diplóide; Tabela 8 – Resultados).

---

<sup>2</sup> Os dados aqui apresentados servem apenas para a exemplificação da técnica utilizada, sendo referentes a uma linhagem WT e uma linhagem transgênica (IAS5 *GmWRKY20* 4.15) apenas.



**Figura 25.** Curvas padrão relativas, geradas a partir dos valores de Ct e das quantidades de DNA obtidas para o gene endógeno *Le1* (em vermelho) e para o gene-alvo *GmWRKY20* (em azul). (A) planta WT; (B) planta transgênica. O software StepOne Plus™ Real Time PCR System (Applied Biosystems™) foi utilizado.

**Tabela 9. Cálculos realizados para a determinação do número de cópias do transgene.**

	Quantidade <i>Le1</i> 4.15		Quantidade de <i>GmWRKY20</i> WT	Quantidade de <i>GmWRKY20</i> 4.15
	251		114	145
	339		134	183
	303		122	175
	389		126	164
	318		152	186
	356		162	194
	413		125	187
	403		142	132
	274		128	179
	216		118	187
	247		111	141
<b>Média</b>	319	<b>Média</b>	131,82	170,27
<b>Quantidade <i>Le1</i> WT</b>	400			
<b>Quantidade relativa <i>Le1</i> 4.15/ <i>Le1</i> WT</b>	0,7975	<b>Correção (quantidade)</b>	131,82	258,34
<b>Fator de correção (quantidade)</b>	<b>1,2539</b>	<b>Correção (eficiência do primer)</b>	<b>200</b>	<b>323,94</b>