



UNIVERSIDADE
ESTADUAL DE LONDRINA

LUCIANO NOBUHIRO AOYAGI

**CARACTERIZAÇÃO ESTRUTURAL E TRANSCRICIONAL
DOS FATORES DE TRANSCRIÇÃO R2R3-MYB NO
GENOMA DA SOJA (*GLYCINE MAX*) EM RESPOSTA A
DOENÇAS**

LUCIANO NOBUHIRO AOYAGI

**CARACTERIZAÇÃO ESTRUTURAL E TRANSCRICIONAL
DOS FATORES DE TRANSCRIÇÃO R2R3-MYB NO
GENOMA DA SOJA (*GLYCINE MAX*) EM RESPOSTA A
DOENÇAS**

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

Orientadora: *Dra.* Francismar Correa
Guimarães Marcelino

Co-orientadora: Prof. *Dra.* Mayra C.C. Gallo de
Carvalho

Londrina
2013

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

Dados Internacionais de Catalogação-na-Publicação (CIP)

A638c Aoyagi, Luciano Nobuhiro.
Caracterização estrutural e transcricional dos fatores de transcrição R2R3-
MYB no genoma da soja (*Glycine max*) em resposta a doenças / Luciano
Nobuhiro Aoyagi. – Londrina, 2013.
167 f. : il.

Orientador: Francismar Corrêa Marcelino-Guimarães.
Coorientador: Mayra Costa da Cruz Gallo De Carvalho.
Dissertação (Mestrado em Biotecnologia) – Universidade Estadual de
Londrina, Centro de Ciências Exatas, Programa de Pós-Graduação em
Biotecnologia, 2013.
Inclui bibliografia.

1. Biotecnologia vegetal – Teses. 2. Soja – Resistência a doenças e pragas –
Teses. 3. *Phakopsora pachyrhizi* – Teses. 4. Genética vegetal – Teses. 5.
Ferrugem da soja (Doença) – Teses. I. Marcelino-Guimarães, Francismar Corrêa.
II. De Carvalho, Mayra Costa da Cruz Gallo. III. Universidade Estadual de
Londrina. Centro de Ciências Exatas. Programa de Pós-Graduação em
Biotecnologia. IV. Título.

CDU 663.1

LUCIANO NOBUHIRO AOYAGI

**CARACTERIZAÇÃO ESTRUTURAL E TRANSCRICIONAL DOS
FATORES DE TRANSCRIÇÃO R2R3-MYB NO GENOMA DA SOJA
(*GLYCINE MAX*) EM RESPOSTA A DOENÇAS**

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

BANCA EXAMINADORA

Dra. Francismar Correa Marcelino Guimarães
EMBRAPA SOJA/UDEL – Londrina – PR

Prof. Dr. Luiz Filipe Protasio Pereira
Embrapa café/IAPAR – Londrina – PR

Dra. Liliane Marcia Mertz Henning
EMBRAPA SOJA – Londrina – PR

Londrina, 23 de julho de 2013

Dedico este trabalho a meus pais, Shintaro Aoyagi e Hiroko Tamura Aoyagi, principais incentivadores e exemplos de dedicação, esforço e vida.

AGRADECIMENTOS

Agradeço a Dra. Francismar C. Marcelino Guimarães e a Dra. Mayra Gallo de Carvalho por aceitarem me orientar, assim como por todo o incentivo, dicas e apoio durante os momentos de dificuldade e dúvida, por toda paciência nos ensinamentos e pela grande contribuição para a realização deste trabalho, bem como de minha formação.

Aos membros da Banca examinadora Dra. Liliane M. Mertz Henning e Dr. Luiz Filipe Protasio Pereira, por disponibilizarem seu precioso tempo na avaliação e melhora do presente trabalho.

Aos analistas e técnicos do Laboratório de Biotecnologia Vegetal, César Silveira, Silvana Rockenback, Vera Pierotti, Jairo, Allan M. Flausino, assim como os demais funcionários da Embrapa-Soja que contribuíram imensamente para a concretização deste trabalho, pela total disposição em fornecer suporte e novos conhecimentos. Em especial a analista Dra. Márcia K. Kuwahara que me acompanhou do início ao fim em minhas atividades de bancada, sempre disposta a ajudar e apoiar, fornecendo muitos conselhos e conhecimento, sem os quais não seria possível concluir este trabalho.

À Pós-doutoranda Adriana Maria Polizel-Podanosqui, bem como a Doutoranda Luana Darben e a Mestre Valéria Lopes-Caitar, por toda paciência com as quais me ensinaram muitas das coisas mais importantes que aprendi nestes dois anos, por toda ajuda nos experimentos e bancada, e acima de tudo, pela amizade e pelo bom convívio.

A todo o pessoal do Laboratório de Biotecnologia Vegetal da Embrapa-Soja Adriana P. Podanosqui, Adriano Camolese, Aguida Morales, Alan Pereira, Alessandra Yokoyama, Amanda Paiva, André Passianotto, Andressa Nakagawa, Anelise Cruz, Bruna Bley, César Silveira, Carolina Rocha, Cynara Romero, Danielle Gregório, Idenize Orsine, Jairo, João Vitor, Juliana Marcolino, Juliana Marinho, Kátia Kuma, Larissa, Letícia Botelho, Lívia Nogueira, Lizandra Catelli, Luana Darben, Luís Fernando, Maíra Maia, Márcia Kuwahara, Mayla, Michele Rincão, Patrícia Honna, Paulo Renan, Renata Fuganti, Salvador Lima, Silvana Rockenback, Tatiana, Valéria Lopes-Caitar, Vera Pieotti, pela amizade e pelos bons momentos nestes dois anos de mestrado.

Aos meus pais Shintaro Aoyagi e Hiroko Tamura Aoyagi, por acreditar, incentivar e sempre apoiar meus estudos e por serem meus exemplos de dedicação, esforço, caráter e vida.

À minha namorada Yoshie, por todo amor, companhia, incentivo apoio e compreensão que me ajudaram a superar mais esta etapa.

A todos os meus amigos de Pós-Graduação em Biotecnologia, pela amizade e por todos os momentos, ajuda e conhecimentos compartilhados durante o convívio no curso.

À Embrapa-Soja e a CAPES, pelo apoio financeiro, concessão de bolsa, disponibilização da estrutura e aperfeiçoamento técnico e profissional, para a realização e desenvolvimento de meu trabalho.

À Universidade Estadual de Londrina e ao Programa de Pós-Graduação em Biotecnologia pela realização deste curso. Agradeço também a coordenação e secretarias pela constante atenção e dedicação fornecidas.

Por fim, a todos que contribuíram de alguma forma para a realização deste trabalho,

AOYAGI, Luciano Nobuhiro. **Caracterização estrutural e funcional dos fatores de transcrição R2R3-MYB no genoma da soja (*Glycine max*) em resposta a estresse biótico**. 2013. 167 f. Dissertação (Mestrado em Biotecnologia) – Universidade Estadual de Londrina, Londrina, 2013.

RESUMO

A soja representa uma das mais importantes culturas para a economia, uma vez que é o principal produto do setor agropecuário e de exportação do Brasil. A Ferrugem asiática da soja (FAS), causada por *Phakopsora pachyrhizi*, representa além de um fator limitante para a expansão da cultura, uma ameaça devido ao seu potencial em causar perdas na produção de grãos e aumento nos custos do cultivo. O controle baseado em produtos químicos, além de ser dispendioso, causa impactos indesejáveis ao homem e ao ambiente, criando a necessidade de se buscar formas alternativas e igualmente eficazes de controle. A compreensão dos mecanismos de defesa e resistência de plantas de soja, bem como a busca de genes ligados a esses processos, representam um caminho para encontrar formas de se transpor os problemas causados por este e outros patógenos. Estudos da interação (transcritoma) de soja com FAS indicam que fatores de transcrição (FT) da família MYB são expressos com frequência e diferencialmente em plantas resistentes. Em *Arabidopsis*, MYBs são amplamente estudados e muitos apresentam funções bem determinadas, incluindo o controle sobre as respostas de defesa. Apesar das evidências apontarem a possível participação destes FTs no controle de importantes processos em soja, incluindo a resposta de defesa, existe uma grande escassez de conhecimento, em comparação. Atualmente MYB é a maior família de FT de *G. max*, assim como a maior entre as plantas, indicando grande potencial de conter candidatos que atuem no controle de importantes processos, assim como é observado em *Arabidopsis*. Objetivando aumentar o conhecimento sobre estes FT, bem como seu real papel e funções para soja, o presente estudo, analisou a família de fatores de transcrição MYB. GmMYBs da classe R2R3 foram identificados a partir do genoma da soja por meio de ferramentas de bioinformática (SMART, Pfam e MEME), e a função putativa foi determinada com base na construção de árvore filogenética e classificação dos GmMYBs em subfamílias, utilizando guias (AtMYB) com funções conhecidas. Todas as subfamílias MYB de *Arabidopsis thaliana* e algumas de soja observadas em trabalhos prévios foram formadas, assim como novo subgrupos. A expressão e o perfil transcricional dos GmMYBs R2R3 em ensaios com *P. pachyrhizi*, *Aphis glycines*, *Heterodera glycines*, *Pratylenchus brachyurus* e *Phytophthora sojae*, através de análises em bancos de transcritomas (Soybase, LGE-Genosojá, Genevestigator), foi realizada e associada a determinação das funções putativas em busca de potenciais genes envolvidos no controle de mecanismos de defesa. Após a seleção dos modelos gênicos alvos, a indução pela infecção com *P. pachyrhizi* foi avaliada, por meio de PCR quantitativo em tempo real, utilizando três genótipos, dois resistentes, PI230970 e Shiranui que possuem o gene Rpp2 e Rpp5, respectivamente, e Williams 82 como modelo suscetível. Entre os genes avaliados, houve destaque para aqueles ligados a síntese de lignina, com três GmMYBs induzidos em plantas resistentes em maior nível e mais rapidamente, um GmMYB com função desconhecida e outro possivelmente ligado a modulação de genes responsivos a quitina, ambos induzidos por um período mais prolongado na planta resistente. Análises funcionais futuras irão ser conduzidas para confirmar o envolvimento desses genes na resposta a infecção por patógenos em soja.

Palavras-chave: *Phakopsora pachyrhizi*. Filogenia. GmMYB.

AOYAGI, Luciano Nobuhiro. **Identification and functional characterization of transcription factors R2R3-MYB in genome of soybean (*Glycine max*)**. 2013. 169 p. Dissertation (Masters in Biotechnology) – State University of Londrina, Londrina, 2013.

ABSTRACT

Soybean is one of the most important crops for the economy, since it is the main product of the agricultural sector and export of Brazil. Asian soybean rust caused by *Phakopsora pachyrhizi*, is a limiting factor for the crop expansion and a threat because of its potential yield losses and increase in cultivation costs. The control carried out with chemicals is expensive and cause undesirable impacts to humans and the environment, creating the need to seek for alternatives and equally effective control. A more profound understanding of the mechanisms of defense and resistance of soybean plants, as well as the search for genes associated with these processes represent a path to find ways to overcome the problems caused by FAS and other pathogens. Studies of the interaction of soybeans with FAS indicate that transcription factors (TF) MYB family are differentially expressed in resistant plants. In Arabidopsis, MYBs are extensively studied and have well-defined functions, including control over the defense of the plant. Despite evidence suggest a possible role of these TFs in controlling important processes in soybean, including defense response, there is a lack of knowledge, compared with Arabidopsis. MYB is currently the largest family of FT of *G. max*, and also in plants, indicating great potential to contain candidates that act to control important processes. Aiming to increase knowledge about these mechanisms, in this work we evaluated the family of MYB transcription factors. GmMYBs class R2R3 were identified from the soybean genome using bioinformatics (SMART, Pfam and MEME), and putative function was determined on the basis of phylogenetic tree construction and classification of subfamilies in GmMYBs using guides (AtMYB) with functions known. All MYB subfamilies of Arabidopsis thaliana and some soybean seen in previous works were formed, as well as new subgroups. The expression and transcriptional profile of GmMYBs R2R3 assays with FAS through analysis in banks of transcritomas (Soybase, LGE-Genosoja, Genevestigator) was performed and associated with determining the putative functions for search potential genes involved in the control mechanism defense. After the selection of target genes, induction by infection with FAS was assessed by quantitative PCR in real time at different times of inoculum (12, 24, 48, 72 and 96 hai) into two treatments (with or without inoculum of spores of *P. pachyrhizi*). Assays were performed in three genotypes, two resistant PI230970 (resistant) having Rpp2 allele and the allele that has Shiranui Rpp5, and showing lesions of the type RB in response to FAS and Williams 82 (susceptible) having TAN type lesions. Among the genes evaluated, there was especially those related to lignin synthesis with three GmMYBs induced in plant resistance in higher levels and faster, one GmMYB with unknown function and other possibly connected to modulation of responsive genes chitin, both induced by a longer period in the plant resistant

Key words: *Phakopsora pachyrhizi*. Phylogeny. GmMYB.

LISTA DE ILUSTRAÇÕES

- Figura 1** – Estrutura dos domínios de receptores imunes de planta.....26
- Figura 2** – Estrutura tridimensional de R2R3-MYB. A e B – Reconhecimento do sulco maior da molécula de DNA pela 2º e 3º alfa hélice de R2R3.43
- Figura 3** – Classes de fatores de transcrição MYB em plantas. H, hélice; T, volta; W, triptofano; X, aminoácido. (Fonte: Dubos et al., 2010).....44
- Figura 4** – Distribuição das classe R2R3 MYB em diferentes espécies de plantas45
- Figura 5** – A - Árvore filogenética com 86 AtMYBs e 35 R2R3-MYBs obtido por KRANZ et al., 1998; B – Árvore filogenética composta por 135 AtMYBs obtido por DUBOS et al., 201047
- Figura 6** – Esquema das diferentes etapas conduzidas para caracterização *in silico* e validação *in vivo* de genes Myb R2R3 em soja envolvidos na resposta a doenças.....55
- Figura 7** – Filogenia, designação de subgrupos e motivos da região carboxi-terminal em proteínas MYB de soja e *Arabidopsis*. A árvore filogenética (*Neighbor-joining*) representa a relação entre sequências R2R3-MYB de *Arabidopsis* (89 sequências) e *Glycine max* (264 sequências.). As sequências proteicas foram agrupados em 41 subgrupos (triângulos), identificados pelos números e letras (“S” para subgrupos identificados a partir de *Arabidopsis* conforme trabalhos de STRACKE et al., 2001 e DUBOS et al., 2010 e “C” para subgrupos identificados previamente em soja). Dois subgrupos (C5 e C8) foram formados unicamente por R2R3-MYB de soja65

- Figura 8** – Perfil fenotípico de genótipos de soja após infecção por *P. pachyrhizi*. A. WH82 com intensa esporulação característica no fenótipo de plantas suscetíveis (lesão do tipo Tan); B. PI587880A com fenótipo misto de plantas resistentes (lesão do tipo Rb) e suscetíveis (lesão do tipo Tan); C. PI230970 com lesões do tipo Rb, indicando resistência à população de isolados utilizados; D. Genótipo Shiranui com lesões do tipo Rb, indicando resistência do genótipo à infecção por *P. pachyrhizi*.....98
- Figura 9** – Níveis de expressão gênica dos GmMYBs no genótipo Williams 82 (WH82) nos cinco tempos de inóculo de *P. pachyrhizi* (12, 24, 48, 72 e 96 hai). As barras de erro representam o desvio padrão para três experimentos independentes, com três repetições técnicas 100
- Figura 10** – Níveis de expressão gênica dos GmMYBs no genótipo resistente PI230970 (Rpp2) nos cinco tempos de inóculo de *P. pachyrhizi* (12, 24, 48, 72 e 96 hai). As barras de erro representam o desvio padrão para três experimentos independentes, com três repetições técnicas 100
- Figura 11** – Níveis de expressão gênica dos GmMYBs no genótipo resistente Shiranui (Rpp5) nos cinco tempos de inóculo de *P. pachyrhizi* (12, 24, 48, 72 e 96 hai). As barras de erro representam o desvio padrão para três experimentos independentes, com três repetições técnicas 101

LISTA DE TABELAS

Tabela 1 – Classificação da família R2R3-MYB de <i>A. thaliana</i>	45
Tabela 2 – Subgrupos de GmMyb identificados no genoma de soja e funções putativas, atribuídas a partir do agrupamento com FT de <i>Arabidopsis thaliana</i> :	68
Tabela 3 – Sequências consensos dos motivos da região C-terminal de sequências completas de GmMYBs identificados através de análise na ferramenta MEME	70
Tabela 4 – Caracterização transcricional de GmMYB R2R2 em resposta a doenças em soja.....	75
Tabela 5 – <i>cis</i> -motivos conservados presentes a montante a região promotora de genes de MYB de soja e Arabidopsis.....	87
Tabela 6 – GmMYBs com expressão diferencial agrupados nos subgrupos juntamente de guias AtMYB	92
Tabela 7 – Expressão geral dos <i>GmMYB</i> com indução nos três genótipos de soja após a infecção com <i>P. pachyrhizi</i>	102

SUMÁRIO

1	INTRODUÇÃO	13
2	REVISÃO BIBLIOGRÁFICA	16
2.1	SOJA – <i>GLYCINE MAX</i> (L.) MERRILL.....	16
2.2	IMPORTÂNCIA ECONÔMICA DA SOJA.....	17
2.3	FERRUGEM ASIÁTICA DA SOJA (FAS)	19
2.4	RESPOSTA DE RESISTÊNCIA A DOENÇAS EM PLANTAS	22
2.5	RESPOSTA DE RESISTÊNCIA A FAS.....	32
2.6	FATORES DE TRANSCRIÇÃO MYB ENVOLVIDOS NA RESPOSTA DE RESISTÊNCIA A DOENÇAS EM PLANTAS	40
2.6.1	FATORES DE TRANSCRIÇÃO MYB	42
3	OBJETIVO GERAL	49
3.1	OBJETIVOS ESPECÍFICOS	49
4	MATERIAL E MÉTODOS	50
4.1	BUSCA E ANOTAÇÃO DE GENES MYBs R2R3 DE <i>GLYCINE MAX</i> A PARTIR DE BANCO DE DADOS	50
4.2	CARACTERIZAÇÃO TRANSCRICIONAL DE MYBs DE <i>GLYCINE MAX</i> EM BANCOS DE TRANSCRITOMAS	50
4.3	CONSTRUÇÃO DA FILOGENIA E CLASSIFICAÇÃO EM SUBGRUPOS DOS MYBs	54
4.4	VALIDAÇÃO <i>IN VIVO</i> DE FATORES DE TRANSCRIÇÃO MYB R2R3 NA RESPOSTA A INFECÇÃO POR <i>P. PACHYRHIZI</i>	56
4.4.1	Material Biológico	56
4.4.2	Inoculação de <i>P. pachyrhizi</i>	56
4.4.3	Avaliação da Expressão dos Genes MYBs	57
4.4.3.1	Coleta do material	57
4.4.3.2	Extração de RNA total	57
4.4.3.3	Síntese do cDNA	58
4.4.3.4	PCR quantitativo em tempo real (RT- qPCR)	59

5	RESULTADOS	61
5.1	IDENTIFICAÇÃO DE MYBs R2R3 EM SOJA	61
5.2	CARACTERIZAÇÃO FILOGENÉTICA E DISTRIBUIÇÃO EM SUBGRUPOS	63
5.3	PERFIL TRANSCRICIONAL DE GMMYB R2R3 EM REPOSTA A DOENÇAS EM SOJA	73
5.4	SELEÇÃO DE GMMYB PARA VALIDAÇÃO <i>IN VIVO</i> E ANÁLISE FUNCIONAL EM RESPOSTA A FAZ	87
5.5	VALIDAÇÃO <i>IN VIVO</i> DO PERFIL TRANSCRICIONAL DE GMMYB ENVOLVIDOS NA RESPOSTA A FAS	97
6	CONCLUSÕES	111
	REFERÊNCIAS BIBLIOGRÁFICAS	113
	ANEXOS	122
	ANEXO A – GmMYBs com domínio R2R3, identificados através de análise das sequências proteicas nas ferramentas SMART, PFAM e MEME	123
	ANEXO B – GmMYBs induzidos após a infecção por <i>P. pachyrhizi</i> em soja	131
	ANEXO C – GmMYBs induzidos em soja durante ataque por <i>Aphis glycines</i> , <i>Heterodera glycines</i> e <i>Phytophthora sojae</i>	141
	ANEXO D – Sequências anotadas como MYBs excluídas	152
	ANEXO E – Comparação dos subgrupos MYB de soja com os identificados por Du et al., (2012)	159
	ANEXO F – Versão completa da Figura 8: Relações Filogenéticas, classificação em subgrupos e motivos da região C-terminal de proteínas GmMYB e AtMYB	162

1 INTRODUÇÃO

A sojicultura, que gera um dos principais produtos agropecuários, é um dos setores que mais contribuem significativamente no saldo positivo da balança comercial, sustentando, desta maneira, a economia nacional. A soja é atualmente o maior produto de exportação do país, movimentando anualmente cerca de 19 bilhões de dólares (CONAB, 2012), e o Brasil é o segundo maior produtor mundial deste grão.

O consumo da soja cresce anualmente, devido a sua importância como fonte de alimento humano e animal, assim como de outros produtos. Apesar da demanda ser crescente e o país apresentar um grande potencial para aumentar a produção deste grão, alguns fatores limitam o desenvolvimento da cultura. Entre os principais fatores estão os estresses bióticos, causados por patógenos (fungos, bactérias, nematóides e vírus) que são responsáveis por consideráveis perdas anuais de produção.

A Ferrugem Asiática da soja (FAS), causada pelo fungo biotrófico obrigatório *Phakopsora pachyrhizi* (Sydow & P.Syd), é atualmente a principal doença que afeta a produção de soja, com danos variando de 10% a 90%, de acordo com o clima da região. Desde a sua identificação em território nacional, estima-se que ela já tenha causado prejuízos, que incluem perda da produtividade, gastos com o controle da doença através de produtos químicos, contabilizados em US\$ 19,7 bilhões, (Instituto Mato-Grossense de Economia Agropecuária - IMEA, 2012).

Os problemas apresentados por esse fitopatógeno são contornados em uma grande maioria dos casos através da adoção de formas de manejo que visam reduzir a quantidade de inóculos nos cultivos (vazio sanitário) e por meio do controle químico, método mais eficiente. Contudo, manejos baseados em aplicações de fungicidas aumentam consideravelmente o custo da produção, são fontes constantes de contaminação ambiental e nem sempre representam uma forma eficiente de controle do patógeno, principalmente na época da floração da soja, uma vez que, neste período, o dossel da planta é muito denso e representa uma barreira eficaz à penetração de fungicidas (TREMBLAY et al., 2010). Soma-se a este fato a preocupação do desenvolvimento de resistência à ação dos fungicidas.

Estes fatos reforçam a significância do investimento em pesquisas e desenvolvimento de novas estratégias visando reduzir os prejuízos causados pela

ferrugem asiática da soja. Além do desenvolvimento das práticas previamente citadas, estudos vêm sendo focados na obtenção de cultivares resistentes e na elucidação dos mecanismos de defesa que contribuem para a resistência ao ataque por este fitopatógeno. Por consequência, a identificação de genes, bem como de processos e mecanismos de resposta de defesa nos quais eles atuam, representam importantes alvos de estudo.

O controle sobre a expressão gênica é responsável pela regulação de um grande número de processos biológicos, que incluem mecanismos de resposta aos mais variados estímulos provindos do ambiente, entre eles, os estresses bióticos causados pelo ataque de patógenos. A regulação gênica ao nível transcricional se destaca devido ao controle absoluto sobre a expressão do gene, assim como do padrão de expressão temporal e espacial. A importância deste processo reflete na composição no genoma dos organismos onde, em média, 4% dos genes (eucariotos), codificam fatores de transcrição (FTs) (RIECHMANN et al., 2001). Apesar de serem encontrados em todos os eucariotos conhecidos, apresentam maior número e diversidade funcional em plantas. Na soja, estima-se que eles representem 12% dos genes codificadores (5.671 genes), um número superior ao observado na planta modelo *Arabidopsis*, onde 6% dos genes codificadores são de FT (FELLER et al., 2010).

Entre os FTs conhecidos, a família MYB, composta por sequências com um a quatro repetições do tipo HTH (*Helix-Turn-Helix*), é uma das mais numerosas e diversificadas funcionalmente. Notadamente, *G. max* apresenta uma das maiores famílias MYB conhecidas do reino vegetal (14% dos seus FT) (WANG et al., 2010). Contudo, em *G. max* pouca informação sobre as funções dos mesmos está disponível. Todo conhecimento sobre esta família provem de estudos com *Arabidopsis*, onde são conhecidos 86 membros com funções determinadas (DUBOS et al., 2010), os quais incluem AtMYBs (R2R3-MYB), que atuam principalmente no controle de respostas de defesa (18,25% dos MYBs). Alguns AtMYBs compõem importantes mecanismos de defesa, tais como morte celular, vias de sinalização e síntese de fenilpropanóides, sendo expressos em resposta a ataques de patógenos e elicitores, e contribuindo com a resistência da planta.

Nos últimos anos, estudos do perfil transcricional da interação da soja com FAS, tem revelado a constante presença de FTs, que incluem WRKY, NAC e MYBs (MORTEL et al., 2007; CHOI et al., 2008; PANTHEE et al., 2009; PANDEY

et al., 2011), entre os genes expressos. Contudo, apesar do destaque dos MYBs, não existem estudos que relacionem estes FTs expressos e seu real papel na defesa e resistência da soja, bem como a sua real função, assim como já determinado na planta modelo *Arabidopsis*.

Estes estudos fornecem evidências de que a seleção de genes, com base nas análises de transcriptomas, permite a identificação de candidatos com potencial efeito na resposta ao estresse em estudo. No caso de FAS, a análise funcional de fatores de transcrição, especialmente da família MYB, fortemente ativados após a infecção, poderá contribuir para a elucidação da complexa regulação das vias de defesa de soja mediada por estes genes frente à ferrugem asiática, bem como contribuir para reduzir a lacuna de conhecimentos existentes quanto a esta família de FTs.

O presente trabalho objetivou a identificação e classificação de genes preditos da família MYB (R2R3) no genoma de soja, bem como identificar e analisar aqueles que estão envolvidos na resposta à FAS e a diferentes patógenos.

2 REVISÃO BIBLIOGRÁFICA

2.1 SOJA – *GLYCINE MAX* (L.) MERRILL

Compreendendo cerca de 727 gêneros e 19.325 espécies, a família Fabacea (ou Leguminosas) é a terceira maior família dentre as angiospermas e também uma das principais do ponto de vista econômico (LEWIS et al., 2005). A característica típica dessa família é a presença de frutos do tipo legume, também conhecido como vagem, que possuem duas linhas de deiscência (abertura longitudinal) e englobam desde espécies arbóreas às herbáceas anuais, muitas de grande importância econômica e principalmente alimentar (soja, feijão, ervilha, amendoim, grão-de-bico). Essa família é dividida em três subfamílias: Caesalpinioideae, Faboideae e Mimosoideae (EMBRAPA, 2010). Dentro destes grupos, a subfamília Faboidea (Papilionoidea) é a maior, com 476 gêneros e aproximadamente 14.000 espécies (LEWIS et al., 2005), e é considerada a de maior importância econômica, visto que nesta subfamília está presente o maior número de espécies importantes para o ser humano, tais como o feijão (*Phaseolus vulgaris*) e a soja [*Glicine Max* (L.) Merrill] (EMBRAPA, 2010).

Glicine max (L.) Merrill é uma planta dicotiledônea herbácea, que pertence à família Fabaceae, subfamília Faboideae e gênero *Glycine* L., compreendendo 15 espécies. As plantas de soja possuem folhas alternadas, pecioladas, compostas de três folíolos ovalados, sendo que na maioria das variedades as folhas amarelam à medida que os frutos amadurecem e caem quando as vagens estão totalmente maduras. O caule é ereto e atinge de 60 cm a 1,5m de altura e apresenta dois estádios de desenvolvimento, o vegetativo (estabelecimento e desenvolvimento) e reprodutivo (florescimento e maturação). Apresenta como frutos vagens achatadas e curtas que possuem de duas a cinco sementes elipsóides de coloração variada, de acordo com a variedade. O sistema radicular é composto por uma raiz pivotante com ramificações e nódulos que abrigam bactérias fixadoras de nitrogênio atmosférico, mais comum dentre os integrantes de Faboideae (ASSUMPÇÃO, 2008).

Atribuí-se a origem dos ancestrais de [*Glicine max* (L.) Merrill] ao continente asiático, crescendo na forma de plantas rasteiras principalmente na costa leste da Ásia nas proximidades do rio Yangtse (também conhecido como Rio Azul,

maior rio em extensão da Ásia – 6300 km), na China. As primeiras citações datam no período entre 2838 a 2883 A.C em livros chineses, nos quais se atribuem o status de grãos sagrados, juntamente com arroz, trigo, cevada e milho, visto que era a base alimentar deste povo. Apesar de ser conhecida, apreciada e consumida a milhares de anos pelos povos orientais, a mesma permaneceu restrita aos países asiáticos até o século XV. A sua disseminação no continente Europeu ocorreu apenas ao final do século XV, com o período das navegações e a abertura de novas rotas de comércio (EMBRAPA, 2011).

O primeiro relato da cultura da soja no continente americano (EUA) data de 1804, sendo inicialmente utilizada como planta forrageira e em menor escala, como produtora de grãos. Seu cultivo e utilização foram expandidos até 1930, quando apresentou produção quase que totalmente voltada para grãos. A introdução da soja no Brasil se deu em 1882 no estado da Bahia, sendo levada mais tarde por imigrantes japoneses para o estado de São Paulo. Em 1892 foi cultivada pela primeira vez pelo Instituto Agrônomo de Campinas (IAC), dando início às pesquisas com essa cultura. Em 1914, a soja foi introduzida no estado do Rio Grande do Sul, local onde as variedades provindas dos Estados Unidos tiveram sucesso, devido às características edafoclimáticas favoráveis encontradas neste estado. Na década de 40, iniciaram-se no Brasil as produções comerciais da soja na forma de grãos, não sendo mais voltada para o forrageio, com gradativos aumentos nas áreas de produção e produtividade na região sul. Por volta das décadas de 70 e 80, a cultura foi expandida para as regiões do Brasil Central, graças ao desenvolvimento de cultivares adaptados à região, bem como pelo desenvolvimento das tecnologias de manejos de solos e implantação do manejo integrado das pragas. Desde então a soja produzida vêm passando por constantes programas de melhoramento genético que visam obter plantas mais adaptadas as diferentes regiões, mais produtivas e resistentes a patógenos e pragas, o que viabiliza seu cultivo em quase todos os estados brasileiros (ASSUMPÇÃO, 2008; FREITAS, 2011).

2.2 IMPORTÂNCIA ECONÔMICA DA SOJA

Dentro das leguminosas, a soja é a espécie de maior importância no mercado internacional, visto que seus grãos possuem alto teor proteico e lipídico; é

fonte de alimentação humana e animal há milênios e, é utilizada na produção de óleo vegetal comestível (principal fonte mundial) e como fonte alternativa de produção de combustível. A semente é um dos alimentos vegetais mais ricos, contendo cerca de 40% de massa proteica, 20% de óleo e possui todos os nove aminoácidos essenciais (histidina, isoleucina, leucina, lisina, metionina, fenilalanina, treonina, triptofano e valina) (ASSUMPÇÃO, 2008).

A soja é um grão que dá origem a diversos produtos e subprodutos muito utilizados pela agroindústria, indústria química e de alimentos. Na alimentação humana, seu uso mais conhecido é como óleo refinado, mas também participa na composição de vários produtos embutidos, como salsichas, chocolates, temperos para saladas, maioneses, entre outros. Sua proteína é à base de ingredientes como massas, produtos de carne, cereais, misturas preparadas, bebidas, alimentação para bebês e alimentos dietéticos. Também são muito usadas pela indústria de adesivos e nutrientes, alimentação animal, adubos, formulador de espumas, fabricação de fibra, revestimento, papel e emulsão de água para tintas. Recentemente, sua utilização vem crescendo também como fonte alternativa de combustível. Da produção atual de biodiesel, que já atinge quase 2,5 milhões de litros ao ano no Brasil, estima-se que o óleo de soja representa mais de 80% da matéria-prima seguidos por gordura bovina (12,4%) e óleo de algodão com 2,1% (ANP, 2011).

A soja é considerada, no Brasil, a cultura de maior expressão comercial uma vez que ela sustenta o agronegócio brasileiro, e representa o maior produto nacional de exportação, com 48,95 milhões de toneladas de janeiro a dezembro de 2012, correspondendo a 26,11 bilhões de dólares (cerca de 10,76% das exportações brasileiras) ou 27,25% do total de exportações de produtos agropecuários, de acordo com o último balanço comercial do agronegócio da CONAB (CONAB, 2013). Ainda, de acordo com dados da CONAB, o saldo do agronegócio é responsável pelo saldo positivo final da balança comercial do país, sendo o setor do agronegócio responsável por garantir o superávit comercial (CONAB, 2013).

Nas safras anteriores a soja representou 46,26% (2010/2011) e 40,01% (66 milhões de toneladas) (2011/2012) do volume de grãos produzidos no país, se mantendo como principal grão produzido. Para a safra atual (2012/2013), estima-se uma safra recorde, de acordo com o 5º levantamento de safras, com produção de 83,42 milhões de toneladas (CONAB, 2013). A área agrícola destinada

ao plantio desta cultura também figura entre as maiores entre os grãos, ocupando cerca de 48,44% (24,18 dos 49,91 milhões de hectares) na safra 2010/2011 e 49,24% (25,04 dos 50,85 milhões de hectares) na safra 2011/2012, sendo estimada em um aumento desta área para 27,65 milhões de hectares para a atual safra (2012/2013).

O maior estado produtor brasileiro de soja é o Mato grosso, com 32,9% da produção brasileira desta cultura (21,84 milhões de toneladas) seguido do estado do Paraná (16,48 % ou 10,9 milhões de toneladas) (CONAB, 2012).

De acordo com o USDA (2013), a produção mundial da safra (2012/2013) foi de 267,606 milhões de toneladas, com produção de 82,055 milhões de toneladas para os Estados Unidos, 82 milhões de toneladas para o Brasil e 51 milhões de toneladas para a Argentina, o que torna o Brasil o segundo maior produtor mundial. O consumo mundial de soja na safra 2011/2012 fechou em 254,91 milhões de toneladas, o que representa um aumento de 1,77% em relação a 251,47 milhões de toneladas atingidas na safra anterior (USDA, 2013).

Dentre os fatores que contribuem para o aumento progressivo no consumo mundial de soja está principalmente o crescimento do poder aquisitivo da população nos países em desenvolvimento, o que promove uma mudança no hábito alimentar. Neste cenário se observa uma progressiva troca do consumo de cereais por carne bovina, suína e de frango, fato que resulta numa maior demanda de soja, principal ingrediente (compõe até 70%) da ração para esses animais (FREITAS, 2011).

Para a safra 2013/2014, a USDA (Junho, 2013), prevê um aumento mundial da produção de soja de 6,6% (17,7 milhões de toneladas). Estima-se ainda que EUA siga como líder na produção (92,2 milhões de toneladas), seguido do Brasil (85 milhões de toneladas).

2.3 FERRUGEM ASIÁTICA DA SOJA

Apesar de todo cenário mundial favorável para a produção e comercialização da soja nos últimos anos, em especial o Brasil com capacidade de expandir a área (ALVES, 2007), alguns fatores que limitam o crescimento e a obtenção de altos rendimentos da cultura, dentre estes se destacam a ocorrência de

estresses bióticos, como doenças causadas por fitopatógenos, além dos abióticos como a seca.

Aproximadamente 40 doenças na soja causadas por fungos, bactérias, nematóides e vírus já foram identificadas no Brasil, número que cresce proporcionalmente com a expansão desta cultura para novas áreas e como consequência da prática da monocultura. A importância econômica de cada doença varia de ano para ano e de região para região, dependendo das condições climáticas de cada safra. As perdas anuais de produção por doenças são estimadas em cerca de 15% a 20%, entretanto, algumas doenças podem ocasionar perdas de quase 100% se constituindo, deste modo, uma ameaça à produtividade e competitividade da soja nacional (EMBRAPA, 2008).

Algumas doenças fúngicas são foco de muitos estudos devido ao prejuízo que têm causado à produção de soja, e entre elas, a Ferrugem Asiática é atualmente a principal doença que acomete a produção de soja no Brasil e uma das mais importantes do mundo.

A ferrugem asiática da soja (FAS), causada pelo fungo biotrófico obrigatório *Phakopsora pachyrhizi* Sydow & P.Syd., é uma das mais severas doenças que incide na cultura da soja, com danos variando de 10% a 90% nas diversas regiões geográficas onde foi relatada (GODOY, 2009). Foi descrita pela primeira vez em 1903 no Japão (HENNING, 1903), sendo constatada pela primeira vez, no Continente Americano no Paraguai, em 5 de março, e no Estado do Paraná em 26 de maio de 2001 (safra 2001/2002), e deste então, se espalhou para todas as regiões produtoras no Brasil, com exceção do Estado de Roraima (EMBRAPA, 2008).

As ferrugens são parasitas obrigatórios (biotróficos), isto é, não matam seus hospedeiros, mas desenvolvem estratégias efetivas para explorar células vivas como fontes de nutrientes.

O ciclo de infecção por *P. pachyrhizi* é típico de outras ferrugens. Esporos de *P. pachyrhizi*, chamados de urediniosporos, são transportados por correntes de ar e podem ser disseminados a quilômetros de distância da planta infectada de origem, em poucos dias. Uma vez que ela atinge a superfície foliar na face superior ou inferior, sob condições adequadas: alta umidade e temperatura entre 15°C e 25°C, a germinação do esporo começa (entre 1 e 2 horas após a infecção - hai) formando um tubo germinativo que cresce por toda a superfície foliar

até atingir uma superfície adequada quando um apressório (estrutura de penetração) é formado. Esta etapa de penetração ocorre entre 7 a 12 hai. Em uma infecção por *P. pachyrhizi*, a penetração ocorre diretamente no centro da célula epidérmica intacta ou sobre as paredes anticlinais, e diferentemente de outras ferrugens, raramente penetra pelo estômato (ZAMBOLIN, 2006; TSUKAHARA et al., 2008; TREMBLAY et al., 2010). Quando esta hifa de penetração passa a cutícula e atinge o espaço intercelular abaixo da epiderme, um septo é formado produzindo hifas primárias (entre 15 e 20 hpi). Entre 24 a 48 hai, a hifa de infecção cresce entre as células paliçádicas para atingir as células do mesófilo esponjoso, onde forma uma célula mãe do haustório, a qual realiza a penetração da célula hospedeira e a diferenciação de um haustório. Este é uma estrutura altamente ramificada, delimitada pela membrana celular da célula hospedeira, denominada de membrana extra-haustorial, uma matriz extra-haustorial e a parede celular do fungo. O haustório proporciona uma ampla superfície de contato com a célula hospedeira e através dele são adquiridas quantidades massivas de açúcares e aminoácidos por mecanismos de simporte a favor de gradiente de prótons (DA SILVA, 2008).

Além da obtenção de nutrientes, através dos haustórios o fungo também promove a secreção de proteínas efetoras importantes para o processo de infecção. Em 96 horas após o esporo chegar às folhas da planta, o fungo segue colonizando o mesófilo esponjoso, através da formação de hifas secundárias e haustórios adicionais. Aproximadamente em 144 horas após a infecção, hifas se agregam e uma urédia (estrutura de reprodução) surge entre as células do mesófilo esponjoso. Cerca de sete a nove dias (168 a 216 horas) após a infecção, novos urediniósporos são formados, podendo ser disseminados por até quatro semanas seguintes. Os primeiros esporos formados emergem na superfície inferior (abaxial) da folha (MORTEL et al., 2007; TREMBLAY et al., 2010).

Três reações do hospedeiro em resposta ao patógeno são possíveis, de acordo com genes que condicionam resistência a FAS (*Rpp*, *Resistance to Phakopsora pachyrhizi genes*). A reação imune governada por *Rpp1*, identificado no genótipo PI200492 (McLEAN; BYTH, 1980), na qual a planta não apresenta sintomas visíveis; a reação de resistência governada por *Rpp2*, identificado em PI230970 (BROMFIELD; HARTWING, 1980); *Rpp3* em PI462312 (HARTWING; BROMFIELD, 1983); *Rpp4* em PI459025 (HARTWING, 1986); *Rpp5* em PI200456 (GARCIA et al., 2008) e *Rpp6* em PI567102B (LI et al., 2012), 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* (MORTELL et al., 2007).

A infecção por *P. pachyrhizi* causa rápido amarelecimento ou bronzeamento e queda prematura das folhas, fato que afeta diretamente o tamanho final do grão e, portanto, o rendimento e qualidade da produção. Observa-se ainda o aborto e a queda das vagens em casos mais graves, em sua totalidade, quando o fungo ataca a cultura na fase de formação das vagens. Elevadas perdas de rendimento têm sido registradas na Austrália (80%), na Índia (90%) e em Taiwan (70%-80%). No Brasil, reduções de produtividade de até 80% têm sido observadas, quando se comparam áreas tratadas e não tratadas com fungicidas (EMBRAPA, 2008).

A doença é favorecida por chuvas bem distribuídas e longos períodos de molhamento, condições ótimas nas quais as perdas na produtividade podem chegar a 90% (EMBRAPA, 2008). No país, o fungo *Phakopsora pachyrhizi* encontrou então condições climáticas muito favoráveis ao seu desenvolvimento, explicando a sua alta agressividade em níveis não antes relatados em outros países produtores (GODOY; HENNING, 2006).

A ferrugem asiática da soja tem causado impactos negativos à economia, devido à elevação dos custos finais da produção e a redução da produtividade. Desde a sua ocorrência na safra 2001/2002 até a safra de 2011/2012, estima-se que a ferrugem asiática da soja tenha onerado o Brasil em torno de US\$ 19 bilhões (IMEA, 2012).

2.4 RESPOSTA DE RESISTÊNCIA À DOENÇAS EM PLANTAS

A resistência de um hospedeiro, dentro do contexto da fisiologia do parasitismo, pode ser definida como a capacidade da planta em atrasar ou evitar a entrada e/ou a subsequente atividade de um patógeno em seus tecidos. Embora as plantas, na natureza, estejam normalmente expostas a um número incalculável de patógenos, a resistência das mesmas mostra-se como regra, enquanto que a suscetibilidade mostra-se como exceção (PASCHOLATI; LEITE, 1995).

De modo geral, existem três razões possíveis que explicam a falha do estabelecimento do patógeno na planta: 1) A planta não possui os requerimentos fisiológicos de um patógeno potencial; 2) a planta possui barreiras estruturais pré-formadas ou compostos tóxicos que limitam as infecções bem sucedidas para espécies de patógenos especializados; 3) ocorre o reconhecimento do ataque do patógeno, e os mecanismos de defesa são ativados e a invasão permanece localizada (HAMMOND-KOSACK; JONES, 1996).

As plantas são imunes a maior parte dos micro-organismos patogênicos, devido à presença de mecanismos de defesa pré-formados ou passivos, os quais proporcionam proteção não específica contra uma ampla gama de patógenos.

Para serem patogênicos, os micro-organismos devem conseguir acessar o interior da planta, por meio da penetração direta através da superfície de folhas e raízes ou entrando através de aberturas como estômatos ou ferimentos. Para isso os invasores devem superar defesas primárias das plantas, barreiras físicas pré-existentes (passivo, constitutivos) (como cutículas, camadas de cera, tricomas, parede celular, estômatos) ou induzidas (deposição de calose; fechamento de estômatos; agregação citoplasmática; formação de halos e papilas; lignificação), além de barreiras bioquímicas (liberação de quitinas, beta-glucanases, fenóis, alcaloides, lactonas insaturadas, glicosídeos fenólicos, glicosídeos cianogênicos, fototoxinas, inibidores proteicos) (PASCHOLATI & LEITE, 1995).

Antes de alcançar o interior da planta, os micro-organismos enfrentam uma rígida parede celular de celulose, antes de alcançar a membrana celular, palco das principais interações ocorrentes entre o patógeno e o hospedeiro (CHISHOLM et al., 2006). A essa etapa onde o organismo invasor se relaciona com as estruturas da defesa primária, dá-se o nome de interações inespecíficas ou basais de resistência, visto que não ocorre a interação direta entre ambos os organismos. Quando um patógeno consegue ultrapassar todas as barreiras iniciais impostas pela planta, ocorre interação direta com o hospedeiro. Nesse caso, a reação pode tomar dois caminhos diferentes. Caso a planta hospedeira seja resistente ao patógeno invasor, uma rápida resposta de defesa é induzida e ocorre a prevenção do desenvolvimento da doença (interação incompatível). Por outro lado, caso a planta hospedeira seja incapaz de reconhecer ou responde muito lentamente

ao invasor (suscetível), a doença desenvolve-se e espalha-se pela planta (interação compatível) (GARCIA-BRUGGER, 2006).

As plantas desenvolveram duas estratégias para detectar patógenos que conseguem sobrepor a estas barreiras iniciais da defesa pré-invasiva. Na face externa da célula hospedeira, elicitores microbianos conservados chamados padrões moleculares associados a patógenos (PAMPs) são reconhecidos por receptores proteicos chamados receptores de reconhecimento de PAMPs (PRRs). As estruturas ou moléculas conservadas pertencentes aos microrganismos que servem como alvos do reconhecimento de receptores eucarióticos são chamados, de maneira geral, de elicitores. PAMPs são tipicamente componentes essenciais de grupos inteiros de patógenos, tais como a flagelina bacteriana, lipopolissacarídeos e quitina fúngica, e seu reconhecimento por PRRs leva a ativação da imunidade basal desencadeada por PAMPs (PTI). A segunda classe de percepção envolve o reconhecimento por receptores intracelulares de moléculas de virulência (proteínas), patógeno-específico, chamados de efetores, onde este reconhecimento induz a imunidade desencadeada por efetores (ETI). Efetores são moléculas do patógeno que manipulam a estrutura e funções celulares do hospedeiro, facilitando a infecção e/ou incitando respostas de defesa e, diferentemente dos PAMPs, não desempenham funções fisiológicas e estruturais essenciais, estando sujeitas a fortes evoluções adaptativas. Efetores podem ser elicitores e/ou toxinas que atuam interceptando os sinais disparados pelo reconhecimento de PAMPs. 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) à resistência sistêmica adquirida (SAR), sendo estas respostas de longa duração (DODDS; RATHJEN, 2010). Estes dois sistemas fazem parte de um processo de coevolução popularmente denominado “queda de braço evolucionária” (“evolutionary arms race”), descrito mais adiante.

Assim, para fornecer a segunda linha de defesa contra os patógenos que tenham superado a primeira linha de defesa baseada na ativação PRRs, as plantas possuem receptores exclusivos ou os produtos dos genes de resistência, que interagem direta ou indiretamente com os efetores dos patógenos. Em contraste ao PRR que reconhece PAMPs, os genes de resistência interagem com as proteínas efetoras de patógenos específicos (VAKHRUSHEVA; NEDOSPASOV, 2010). Esse

reconhecimento envolve a interação gene-a-gene (teoria de FLOR). De acordo com esse modelo, um efector específico (produto potencial de um gene *Avr*) é reconhecido por um receptor (produto potencial de um gene *R*) na planta resistente. Os genes que dão origem às proteínas efetoras 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. O reconhecimento da proteína *Avr* leva a uma cascata de eventos de sinalização que irá ativar genes da célula vegetal hospedeira, resultando em uma rápida resposta de defesa.

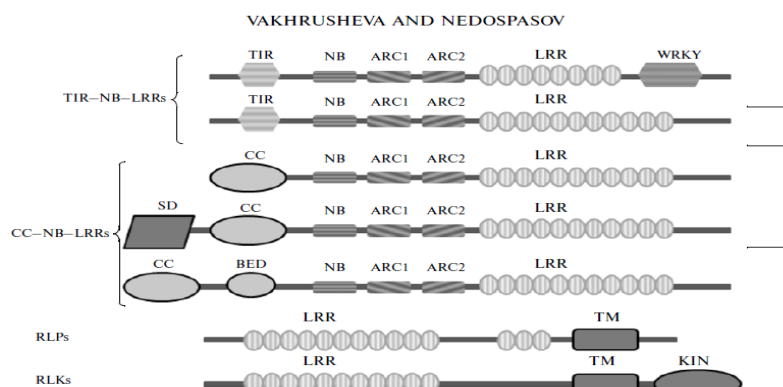
Os receptores PRRs de plantas, que medeiam a resposta inata (PTI), bem como daqueles que medeiam a imunidade específica (secundária ou ETI), são proteínas extracelulares com um domínio transmembrana e uma porção extracelular (presentes na superfície celular), normalmente contendo repetições LRRs (regiões ricas em leucinas) envolvida no reconhecimento e ligação a PAMPs.

Por exemplo, as plantas mostram reação imune aos componentes da parede celular fúngica (quitina, glucana, ou glicoproteínas), lipopolissacarídeo bacteriano (LPS), e flagelina. Nas plantas, dois tipos de PRRs são encontrados: (1) proteínas que possuem um domínio quinase intracelular, chamado receptor-like kinases (LRR-RLKs), e (2) proteínas sem um domínio citoplasmático, denominado receptor-tipo proteínas (LRR-RLPs). PAMPs ligados a esses receptores induzem a PTI, sendo ela a primeira ativação de resposta a percepção da presença do patógeno (VAKHRUSHEVA; NEDOSPASOV, 2010).

Na co-evolução da interação patógeno-hospedeiro, os patógenos adquiriram a habilidade de produzir e introduzir proteínas efetoras nas células vegetais visando inibir essa defesa primária, por meio da interceptação dos sinais gerados pelo reconhecimento das PAMPs pelas PRRs. Por outro lado, a evolução do sistema de defesa vegetal e a produção de proteínas receptoras envolvidas no reconhecimento específico de proteínas efetoras do patógeno resultou no desenvolvimento de uma resposta imune secundária (ETI), cujas as proteínas envolvidas (receptores) monitoram diretamente ou indiretamente a presença do patógeno, através do reconhecimento desses efetores. Tais receptores são denominados proteínas R (proteínas de resistência), sendo eles produtos dos genes R. A maioria das proteínas R estruturalmente pertence à família de proteínas contendo domínios de ligação a nucleotídeos (NB - envolvido na ligação a ATP e

sua hidrólise, o que promove modificações conformacionais no receptor e subsequente ativação da cascata de sinais) e domínios ricos em leucina (porção C-terminal) (NB-LRRs), e podem ser divididas de acordo com a composição da região amino-terminal dos genes. Muitos têm um domínio com homologia aos domínios intracelulares de sinalização “Toll” de drosófila e receptores de interleucina “(IL) - 1” de mamíferos, sendo chamados de TIR (“Toll-Interleukin Receptor”), e denominados TIR-NB-LRR. Outros genes contêm domínios CC (“Coiled coil”) na região amino-terminal, sendo chamados de CC-NB-LRR. Ambos, CC e TIR, seriam responsáveis pelo reconhecimento de modo indireto do patógeno, interagindo com proteínas vegetais modificadas pela ação dos patógenos (Figura1) (VAKHRUSHEVA; NEDOSPASOV, 2010).

Figura 1 – Estrutura dos domínios de receptores imunes de planta. Fonte: VAKHRUSHEVA; NEDOSPASOV, 2011.



Contudo, muitos estudos vêm demonstrando que o reconhecimento de efetores provavelmente tenha evoluído como um mecanismo indireto (sem o contato direto entre efetor específico e o seu receptor específico como proposto na Teoria de FLOR), visto que o contato direto raramente ocorre.

Uma nova teoria foi criada, como uma complementação da teoria de FLOR, conhecida como “modelo de guarda”. Neste modelo as proteínas R “vigiarão” proteínas acessórias que são os alvos de ação dos patógenos. Nesse modelo, as proteínas acessórias interagirão com os efetores e então serão capazes de ativar os genes de resistência. A detecção indireta explicaria também a razão de existir um repertório limitado de receptores de resistência na planta para responderem a uma ampla gama de patógenos e seus efetores correspondentes. Grande parte dos efetores bacterianos já caracterizados, por exemplo, exibem

atividade enzimática capazes de modificar proteínas vegetais. Muitas evidências estão indicando que múltiplos efetores e suas funções enzimáticas atuam sobre as mesmas proteínas alvo nos hospedeiro. Deste modo, a planta não desenvolveu receptores capazes de detectar todos os possíveis mecanismos efetores, e sim as modificações sofridas por alvos comuns de patógenos. Assim, através do monitoramento de perturbações comuns, as proteínas R detectam indiretamente a atividade enzimática de efetores múltiplos (DODDS; RATHJEN, 2010).

Os 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. Alguns trabalhos demonstram que mutações que levam a perda do sistema TTSS em bactérias, e, conseqüentemente a capacidade de seus efetores atuarem, promoveu a perda de sua patogenicidade, demonstrando a importância dos efetores para o desenvolvimento da infecção. Isto indica ainda que a patogenicidade de muitas bactérias esteja diretamente relacionada à sua habilidade de suprimir as respostas de PTI (DODDS; RATHJEN, 2010). Muitas vezes ocorre que, através da eliminação ou da modificação de seus efetores ou dos produtos de suas ações, alguns patógenos podem mais uma vez escapar à detecção pela planta, deste modo causando a doença, o que indica que aparentemente o processo dinâmico de evolução ainda continua.

Na maioria dos casos, como resultado do reconhecimento dos efetores e da ativação dos genes R ocorre 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. Contudo, diferentemente da PTI, na ETI geralmente ocorre o desenvolvimento de HR, resultando em morte celular programada (VAKHRUSHEVA; NEDOSPASOV, 2010).

O principal componente da resposta de defesa ativada pela interação entre o produto do gene *R* e o produto do gene *Avr* correspondente é a Resposta de Hipersensibilidade (“Hipersensitive Response” ou HR). A HR é definida com a morte das células do hospedeiro dentro de poucas horas após o contato com o patógeno. Em interações com patógenos biotróficos, os quais desenvolvem íntima associação com as células do hospedeiro através dos haustórios, a morte celular priva o patógeno do acesso a nutrientes, controlando seu desenvolvimento. A HR é uma resposta de defesa local, ativada somente nos pontos de contato entre a planta

e o patógeno. As células mortas neste processo contêm altas concentrações de moléculas antimicrobianas e, por isso, não são subsequentemente atacadas por patógenos. A morte celular na HR parece resultar de dois processos principais: a produção de espécies reativas de oxigênio, as quais são tóxicas para a célula e levam à morte celular, e a ativação de um processo de morte celular programada (HAMMOND-KOSACK; JONES, 1996).

Os ânions superóxido (O_2^-), gerados na membrana ou nos protoplastos e secretados ao ambiente extracelular, seriam rapidamente dismutados de um modo não enzimático ou através da ação da superóxido dismutase (SOD), produzindo peróxido de hidrogênio. Este, por não ser eletricamente carregado, pode atravessar facilmente a membrana plasmática, atingindo o ambiente intracelular. Os ânions superóxido e peróxido de hidrogênio não são muito reativos, mas podem ser convertidos aos radicais altamente destrutivos hidroxiperoxil (HO_2) e hidroxila (OH^-), respectivamente, sendo este último formado na presença de íons Fe^{2+} . Todas estas ROS podem resultar em dano considerável ao hospedeiro e ao patógeno, ao reagirem com proteínas, reduzindo a atividade de enzimas; com lipídeos, desestabilizando as membranas; e com o DNA, podendo causar mutações. As espécies reativas de oxigênio ("Reactive oxygen species" - ROS), além de desencadear a morte celular vinculada à HR e controlarem o desenvolvimento do patógeno, ainda: participam de reações envolvidas no reforço da parede celular, uma vez que peróxido de hidrogênio é essencial para a formação de precursores de lignina, e radicais de oxigênio podem oxidar as glicoproteínas da parede celular, as quais são então ligadas à mesma através de ligações cruzadas; ativam uma cascata de sinalização mediada pelo ácido salicílico ("Salicylic Acid" ou SA), uma vez que o peróxido de hidrogênio aumenta a atividade da enzima BA2-H envolvida na biossíntese do SA, e levam à alteração do potencial redox na célula, o qual é conhecido por ativar fatores de transcrição, assim desempenhando uma atividade de sinalização (HAMMOND-KOSACK; JONES, 1996). O óxido nítrico ("Nitric Oxide" ou NO) é outra importante molécula que se acumula durante interações incompatíveis. O NO pode auxiliar a formação da HR por potencializar a indução da morte celular mediada por ROS.

Os transcritos de inúmeros genes relacionados a patogênese (genes PR) se acumulam dentro de minutos a horas do ataque pelo patógeno. Geralmente os mesmos genes são induzidos em interações compatíveis e

incompatíveis, entretanto mais fraca e vagarosamente nas interações compatíveis. Algumas proteínas PR são quitinases e glucanases, enzimas que degradam polissacarídeos estruturais das paredes das células fúngicas e provavelmente reduzem o crescimento fúngico. O etileno e o SA parecem agir sinergicamente levando ao aumento das proteínas PR. Outras proteínas PR, como as lipoxigenases (LOX), geram moléculas sinalizadoras secundárias tais como JA e peróxidos lipídicos, e produzem metabólitos secundários tóxicos voláteis e não voláteis. As defensinas são também proteínas com atividade antimicrobiana. A via de sinalização que controla o acúmulo de defensinas é mediada por etileno e JA (HAMMOND-KOSACK; JONES, 1996).

Poucos minutos após o reconhecimento da proteína (Avr), rotas complexas de sinalização são ativadas. Um elemento inicial comum dessas cascatas é a mudança transitória da permeabilidade iônica da membrana plasmática vegetal. A ativação dos genes *R* estimula a entrada dos íons Ca^{2+} e H^+ na célula e a saída de K^+ e Cl^- . O rápido acúmulo de Ca^{2+} citosólico regula o complexo NADPH oxidase diretamente ou indiretamente, através de aumentada produção de NADPH pela proteína de ligação a Ca^{2+} , calmodulina. Assim, o acúmulo de Ca^{2+} estaria envolvido na indução de cascata de sinalização envolvendo MAP kinases até o núcleo, culminando na ativação de fatores de transcrição e genes responsivos a patógenos, produção de espécies reativas de oxigênio (ROS) e de óxido nítrico; que resultam na deposição de calose para reforçar a parede celular no sítio de infecção, de modo a prevenir a entrada e o crescimento microbiano, no fechamento de estômatos e a síntese de compostos antimicrobianos (PRs), tais como enzimas hidrolíticas (CHISHOLM et al., 2006; DA SILVA, 2008).

De modo geral, apesar da eficiência das repostas de defesa induzidas após ataques de insetos herbívoros e micro-organismos patogênicos, elas implicam em gastos energéticos. Deste modo, sistemas complexos e elaborados de regulação que coordenam a ativação de repostas de defesa (exemplo: fitormônios), de modo específico para cada tipo de ataque, foram desenvolvidos conciliando a redução de custos energéticos e a capacidade de responder e resistir ao ataque.

Geralmente patógenos com modo de vida biotrófico são mais sensíveis às defesas mediadas por SA, ao passo que os patógenos necrotóficos e insetos herbívoros são mais sensíveis à defesa mediada por JA e ET (KOORNNEF; PIETERSE, (2007). O etileno (ET) e o ácido jasmônico (JA) são moléculas

sinalizadoras envolvidas na resposta de defesa a estresses bióticos e abióticos. Eles induzem à expressão de compostos antimicrobianos não induzidos por SA. Acredita-se que estas moléculas atuem em uma cascata antagonista à cascata envolvendo SA, embora interações sinérgicas tenham sido descritas. Deste modo como resultado da interação negativa entre SA e JA, uma ativação da resposta mediada por SA devem conferir à planta maior susceptibilidade a ataques que são detidos via defesas dependentes de ativação por JA (KOORNNEF; PIETERSE, 2007).

A produção de ácido salicílico, ácido benzóico (“Benzoic Acid” ou BA), etileno e de ácido jasmônico (“Jasmonic Acid” ou JA), está relacionada à resposta a fitopatógenos. Em locais em que ocorre a incompatibilidade mediada pelo reconhecimento R-Avr, níveis elevados de SA garantem que respostas de defesa, como a indução de genes PR, inibindo o desenvolvimento do patógeno. Uma proteína envolvida na produção de SA em plantas é a fenilalanina amonialiase (“phenilalanine amonialyase” ou PAL), a qual tem sua produção aumentada após ataque com fungos (TAIZ; ZEIGER, 2004). Entretanto, a forma mais observada de produção de SA envolve a liberação de BA, que induz a expressão de um citocromo P450 monoxigenase (BA2-H), o qual converte BA em SA. A atividade de BA2-H é fortemente aumentada após HR.

Uma vez que as respostas de defesa da planta são ativadas no local da infecção na tentativa de impedir e/ou retardar a penetração do patógeno, uma resposta de defesa sistêmica é frequentemente desencadeada em partes distais para proteger estes tecidos danificados contra posterior invasão pelo patógeno. Essa resistência à doença de longa duração e de amplo espectro é referida como resistência sistêmica adquirida (SAR) e é caracterizada pela ativação coordenada de um conjunto de agentes de defesa, incluindo genes PR. O início da SAR pode ser desencadeada por PTI ou ETI mediante o reconhecimento do patógeno e está associada com aumento dos níveis de SA, localmente nos sítios de infecção e muitas vezes também sistemicamente, em tecidos distantes em plantas que apresentaram esta resposta (PIETERSE et al., 2009).

Para que ocorra a SAR, a infecção inicial precisa resultar na formação de lesões necróticas, decorrentes da HR (acúmulo de peróxido de hidrogênio) ou como sintoma da doença (HAMMOND-KOSACK; JONES, 2000). Plantas cujas mutações prejudicam a produção de SA são incapazes de desenvolver

SAR e não mostraram ativação de genes PR após a infecção, o que indica que o SA é um agente necessário na via de sinalização de SAR (PIETERSE et al., 2009).

Evidências para o papel fundamental de fitormônios nas respostas de defesa das plantas vêm de estudos com o modelo *Arabidopsis thaliana*, nos quais vários mutantes para a biossíntese desses hormônios demonstram modificações substanciais nos níveis de resistência e tipo de resposta frente a determinados tipos de patógenos. A partir desses estudos, tornou-se evidente que os patógenos biotróficos são geralmente sensíveis para a resposta de defesa que são regulados por Ácidos Salicílico (SA), tais como respostas de hipersensibilidade, enquanto patógenos necrotróficos são comumente limitados por defesas mediados por Ácidos Jasmônico e Etileno, sendo que essas duas vias são consideradas antagonistas, onde a produção de o produto de uma das vias inibe a outra via (GLAZEBROOK et al., 2005 ; PIETERSE et al.,2009). Considera-se que as diferenças observadas nas respostas ativadas por cada uma das vias seja atribuídas à indução diferencial de genes de resistência, onde JA e ET ativariam genes não ativados por SA.

De acordo com THOMMA et al. (1998), os hormônios vegetais SA e JA, cujos níveis aumentam durante a infecção por patógenos, ativam conjuntos distintos de genes que codificam proteínas antimicrobianas em *Arabidopsis thaliana*. Os genes induzidos por patógenos, PR-1 e PR-2 (β -1,2-D-glucanase) e PR-5 (Traumatin-like protein) requerem SA como um sinalizador para a sua ativação, enquanto que os genes PDF 1.2, de defensinas de plantas, e PR-3 e PR-4 (quitinases) são induzidos por patógenos através de uma via independente de SA e dependente de JA. *NahG* e *npr1-1* são mutantes de *Arabidopsis* insensíveis a ácido salicílico (SA) que têm suas respostas suprimidas a este hormônio, visto que *NahG* codifica uma SA hidroxilase que degrada SA e *npr1-1* não é capaz de codificar a proteína *npr1*, que quando estimulada por SA se desloca para o núcleo onde interage com fatores de transcrição TGA, os quais então ativam uma série de genes de defesa. O outro mutante *coil-1* não produz um componente de tradução de sinal (receptor) que reconhece JA como um sinal e desencadeia respostas, como ativação de genes de resistência (THOMMA et al., 1998). THOMMA et al. (1998), confrontou em seu estudo a planta *Arabidopsis thaliana*, mutantes *NahG*, *npr1-1* e *coil-1* a dois fungos necrotróficos (*Alternaria brassicola* e *Botrytis cinérea*) e um biotrófico (*Penospora parasítica*). Eles observaram que plantas mutantes *coil-1* (sem a

capacidade de responder a JA) foram muito colonizadas pelo fungo necrotrófico *A. brassicola* (diâmetro médio das lesões cerca de sete vezes maiores) quando comparado a plantas selvagens capazes de responder ao JA produzido. O mesmo resultado foi observado no confronto com outro fungo necrotrófico, *B. cinerea*, no qual mutantes incapazes de responder a SA (*NahG* e *npr1-1*) se mostraram mais suscetíveis ao fungo biotrófico *P. parasítica* (presença de hifas intercelulares e oósporos) e, como o esperado, um mutante incapaz de responder a JA (no qual não há inibição da reposta mediada por SA) demonstrou resistência a este patógeno. A aplicação prévia de INA (substância análoga funciona de SA) em plantas selvagens (Col-0) infectadas com *P. parasítica* promoveu eficiente proteção a esse parasita biotrófico, mas nenhuma proteção com a adição de MeJA, indicando novamente a importância de repostas geradas por SA frente a fungos biotróficos (THOMMA et al., 1998).

Outros estudos, como os promovidos por STASWICK et al. (1998) e VIJAYAN et al. (1998) demonstram também a importância de JA em reposta em defesa frente a fitopatógenos necrotróficos (*Phytium irregulare* e *Phytrium mastophorum*). Plantas de *A. thaliana* insensíveis (mutante *jar1*) ou deficientes na síntese de precursores (mutante triplo *fad 3-2 fad 7-2 fad8*) e, conseqüentemente na síntese deste hormônio, foram altamente suscetíveis.

2.5 RESPOSTA DE RESISTÊNCIA A FAS

O primeiro estudo da interação de *P. pachyrhizi* com *Glycine max* foi a análise do transcriptoma da soja via microarranjo, onde foram analisadas plantas jovens (estágio V2 – 21 dias), com amostras coletadas em 72 hpi com o patógeno (PANTHEE et al., 2007). Dos 35.611 genes presente no microarranjo, foi observado um total de 112 genes diferencialmente expressos na interação entre soja e FAS, sendo que entre eles, 46 tiveram aumento de expressão (up - regulado) e 66 redução (down - regulado). A maioria dos genes com aumento de expressão tem função de defesa e resposta a estresse, sendo um gene relacionado à biossíntese de SA expressado sete vezes mais durante a infecção no estágio (V2) avaliado. Os demais genes expressos, tais como proteínas heat shock (HSP), proteínas kinase do tipo receptor associado à senescência (LSRK), glutationa S-transferase e uma

chalcona isomerase (CI), são todos relacionados a defesa geral e tolerância a estresse.

Posteriormente (PANTHEE et al., 2009), o mesmo grupo comparou o perfil de expressão em duas fases outras fases do desenvolvimento da soja (estágios V4 – sete semanas, estágio vegetativo e R1- nove semanas, início do estágio reprodutivo), 72 hpi. Estes estudos objetivaram demonstrar que a expressão diferencial de genes em resposta à infecção depende do estágio de desenvolvimento. A análise dos genes diferentemente expressos entre as condições de infecção e controle, considerando-se fase de crescimento e tratamento com ferrugem como fatores distintos, revelou que 2003 genes com diferença significativa de expressão são atribuídos a diferenças na fase de crescimento, e 7370 genes com diferença significativa devido à infecção. Contudo a análise estatística da interação entre ambos os fatores revelaram a existência de significância para 3998 genes diferentemente expressos quando se combinam todas as possibilidades, o que demonstra que o padrão de expressão gênica de soja em resposta a *P. pachyrhizi* é específica e diferencial para cada estágio (V4 ou R1). Este fato é reforçado pela observação de que a maioria dos genes é expressa diferencialmente, quase 10 vezes mais na fase V4 (5056) em relação a R1 (579), indicando que a concentração de novos estudos para a análise e descoberta de genes ligados a defesa durante o desenvolvimento da infecção devem ser concentrados nas fases mais tardias e preferencialmente na fase V4.

Ao compararem com os resultados obtidos no estudo de 2007, os autores ressaltam uma maior expressão nessas fases mais tardias do desenvolvimento da soja, quando comparado aos estágios iniciais (V2) (24 genes diferentemente expressos). Uma das razões apresentada está no fato de que os estágios V4 e R1 são aqueles nas quais ocorre o desenvolvimento natural de *P. pachyrhizi*. Entre os genes de defesa encontrados estão genes associados com a resposta de morte celular, biossíntese de JA, lignina e ácido salicílico e 28 (V4) e 11(R1) genes que codificam para os fatores de transcrição do tipo WRKY, que atuam em sua maioria regulando outros genes, entre os quais podem estar relacionados à defesa.

MORTEL et al. (2007) também realizou uma análise do perfil de expressão de genes durante a infecção de soja por FAS através de microarranjo contendo o genoma da soja, onde foi comparada a reação susceptível à reação de

hipersensibilidade (mediada pelo gene de resistência *Rpp2*), sendo as amostragens do tecido realizadas após 6, 12, 18, 24, 36,48, 72, 96, 120 e 168 hpi. Foram identificados 1516 e 894 genes diferentemente expressos em plantas suscetíveis e resistentes, respectivamente. Destes, 470 apresentaram mudanças significativas de expressão em ambos os genótipos, e a análise destes genes por agrupamento hierárquico permitiu observar que houve um pico diferencial para ambos os genótipos induzidos em 12hpi e, uma segunda forte fase diferencial de expressão em estágios mais avançados da infecção, com início mais cedo no genótipo resistente. Esta observação bifásica de expressão diferenciada entre os genótipos foi comprovada estatisticamente em um modelo linear, com separação temporal dos dados. Houve um pico na infecção na fase inicial, com expressão de 879 e 240 genes, e 180 e 238 nas fases tardias de infecção nos genótipos suscetível e resistente, respectivamente. A combinação de ambas as análises dos genes diferentemente expressos, em relação aos genótipos e ao tempo, revelaram que a segunda fase de expressão diferencial ocorre no tempo de 72hai, ponto no qual há um maior número de genes induzidos no genótipo resistente, fato que ocorre apenas 24hs mais tarde no genótipo suscetível (96hai).

A anotação funcional dos genes identificados revelou a ocorrência das classes funcionais de resposta de defesa, metabolismo secundário, transcrição e regulação, e transporte foram induzidos. No genótipo resistente, foram exclusivamente expressos genes pertencentes à classe de regulação da transcrição. Genes associados ao metabolismo secundário, provavelmente envolvendo a via dos flavonóides para síntese de fitoalexinas e de metabolitos de reforço de paredes, também foram mais induzidos nas duas fases, sendo igualmente expressos na fase inicial de infecção e pelo menos 24hs mais cedo nas plantas resistentes (72 hai). Outra classe funcional induzida é a dos genes de fatores de transcrição, sendo observados 127 conjuntos diferentemente regulados na interação da soja com a FAS, onde mais da metade dos mesmos (70) são pertencentes à classe de defesa e estresse. Entre eles, mais da metade (46, ou um terço dos fatores expressos) pertencem à família WRKY de fatores de transcrição, seguido da família MYB com 13 genes expressos (19%), e juntos correspondem a 50% de todos FTs, indicando um potencial papel deles na resposta de resistência à infecção. De modo geral, a resistência mediada por *Rpp2* parece provir da resposta mais rápida (1 a 2 dias) em relação à planta suscetível, começando por volta de 72 hai.

Em trabalho mais recente, SCHNEIDER et. al. (2011), procurou correlacionar essa alteração bifásica na expressão de genes em soja induzida por *P. pachyrhizi* com as etapas de infecção, correlacionando o perfil expressão gênica com análises microscópicas nos mesmos pontos de análise. Os autores utilizaram apenas um genótipo, c.v Ankur contendo o gene de resistência *Rpp3*, infectado com os isolados de *P. pachyrhizi* Hawaii 94-1 e Taiwan 80-2, originando interações compatível e incompatível, respectivamente. As análises microscópicas demonstraram que para ambos os isolados houve um desenvolvimento de infecção de forma semelhante até 72 hai, ponto no qual as diferenças entre as interações compatível e incompatível tornaram-se evidentes. As colônias da cepa virulenta TW80-2 demonstraram alto desenvolvimento de hifas e haustórios e expansão maior e mais rápida quando comparado à cepa avirulenta HW94-1, no qual se pode observar um maior número de células necróticas e descoloridas como parte da resposta de defesa. A análise dos genes diferentemente expressos ao longo do tempo revelou novamente um padrão bifásico em resposta a *P. pachyrhizi*, no qual no ensaio com HW94-1, houve 1833 genes induzidos em 12 hai, apenas um gene diferentemente expresso em 24 hai, e novo pico a partir de 72 hai, enquanto que em TW80-2, houve 726 genes induzidos em 12 hai, e no período de 24 a 72 hai, não houve expressão gênica diferencial, o que ocorreu novamente apenas a partir de 144 hai, de forma muito menos expressivo e demorado.

A anotação funcional revelou 54 categorias de processos biológicos (GO), dos quais 14 foram significativos apenas na interação incompatível e 20 induzidos em ambas as interações, contudo com maior número de genes diferencialmente expressos ocorrendo na interação incompatível. Entre os genes induzidos na interação incompatível estavam os de resposta SA, estímulo de ácido abscísico (ABA) e biossíntese JA. Na interação incompatível ainda foram observados indução de quatro classes associados à regulação e transcrição, dos quais dois, metabolismo de peróxidos de hidrogênio e regulação da transcrição, são exclusivos deste tipo de interação. A análise das classes de fatores de transcrição diferencialmente expressos permitiu identificar 2194 conjuntos de genes de fatores de transcrição, dentre os quais 479 são exclusivas das interações incompatíveis, entre eles AP2-ereb, AUX-IAA-ARF, bHLH, C2C2, homeobox, MYB, NAC, TPR e WRKY. Durante os períodos de 24 a 72 hai, nos quais se observam a repressão de muitos genes, incluindo FT na interação compatível, não se observou o mesmo na

interação incompatível, visto que seus fatores de transcrição se mantiveram expressos em altos níveis (WRKY; bHLH, MYB-HD like; C2C2). Entre eles, os FT exclusivos da interação incompatível WRKY, MYB e NAC possuem papel na resposta de defesa. Este perfil diferencial de expressão, bem como a indução apenas na interação incompatível e a não repressão nas fases diferencialmente importantes entre as interações indica a grande importância que este grupo de fatores de transcrição representa para a resposta diferenciada das plantas resistentes frente ao patógeno virulento.

CHOI et al.(2008) obtiveram o perfil de expressão da interação soja - *P. pachyrhizi* através da construção de bibliotecas SSH (suppressive subtraction hybridization) seguida de análise por microarranjo, comparando a reação suscetível com a imune (governada pelo gene *Rpp1*), nos estágios iniciais da infecção (1, 6, 12, 24 e 48 hpi). Neste trabalho foi ressaltada a importância tanto de genes responsáveis pelo balanço oxidativo (lipoxigenases e peroxidases), visto que os mesmos foram expressos em toda a infecção; além de membros de diversas famílias de fatores de transcrição. A ativação de enzimas oxidantes como peroxidases e lipoxigenases sugerem a ocorrência de uma explosão oxidativa no início da infecção, o que resultaria em modificações na parede celular (reforço pela indução de formação de lignina), indução de compostos de defesa, ativação de cascatas de sinalização mediadas pelo SA. Além do papel protetor atribuído a este tipo de resposta, o dano oxidativo da célula do hospedeiro também pode resultar em necrose, promovendo uma defesa primária a infecção e resultando em parte da resistência mediada por *Rpp1*. Entre os genes relacionadas à transcrição diferencialmente expressados, foram novamente observados as famílias de fatores de transcrição tais como WRKY, *up*-regulados em 12 hai e MYB, *up* –regulados em 12 e 24 hai, sendo *down*-regulados em 24 e 24-48 hai, respectivamente. Esta expressão de diferentes membros de fatores de transcrição de modo tanto positivo (nos picos iniciais de resposta) como negativo (na fase onde não ocorreu expressão significativa de genes) nas diferentes fases do estudo, sugere que eles podem estar desempenhando papéis cruciais na regulação das vias de defesa durante os eventos de infecção na resposta de resistência nos quais são expressos.

SORIA-GUERRA et al. (2009) procurou identificar genes associados à resposta de resistência à FAS comparando a expressão diferencial de genes em plantas de *Glycine tomentella*, infectadas e não infectadas no intervalo de 12, 24, 48

e 72 hai, em genótipos resistentes (R) e suscetíveis (S), utilizando um microarranjo com 38,400 genes de soja. A caracterização da resistência à FAS em *G. Tomantella*, uma parente selvagem de *Glycine max*, poderia identificar novos genes de resistência que poderiam ser transferidos para a soja, a fim de se desenvolver uma variedade resistente. Um total de 1342 genes foram diferentemente expressos nos quatro tempos testados nos genótipos resistente e suscetível, dentre os quais 70% foram up-regulados e 30% down-regulados. Muitos genes diferencialmente expressos estavam associados à via biossintética de flavonóides. A enzima Isoflavona redutase aparece induzida em 24 e 48 hai apenas nas plantas resistentes, sendo reprimida em plantas suscetíveis. Além desta, várias enzimas da rota foram induzidas no genótipo resistente desde as primeiras horas de infecção, tais como PAL, Chalcona sintase, isoflavona redutase, 4-coumarato-CoA ligase, UDP-glicose:flavonóide glicosetranferase. Outras classes de enzimas encontradas sendo expressas, up-reguladas apenas no genótipo resistente, em todos os pontos de tempo testados foram as lipoxigenases (up-regulado no genótipo R em 24-48 hai) e peroxidases, indicando possível explosão oxidativa como resposta mediada nas plantas resistentes nas etapas iniciais da infecção. Outros genes com mudanças maiores de expressão nas plantas resistentes foram os genes relacionados à patogênese (PR), tais como a beta-1,3-glucanase e quitinase, ambos atuantes na degradação de parede fúngica. Quitinase foi up-regulado apenas no genótipo R e down-regulado em S, em 12 hai, enquanto que beta-1,3-glucanase foi expresso duas vezes mais no genótipo resistente (0,24 em R e 0,11 em S) em 12 hai.

Cerca de 14% (4 de 28) dos genes envolvidos com processos de transcrição encontrados neste trabalho são da família de fatores de transcrição WRKY. Eles foram encontrados em todos os pontos de análise (12-72 hai) em ambos os genótipos, contudo o tempo para atingir a expressão máxima foi diferencial para ambos os genótipos, ocorrendo em 12 hai para a planta R e em 24-48 hai para as plantas S. Todos estes genes destas classes funcionais descritos estão possivelmente envolvidos diretamente com a resistência a FAS em *G. tomentella*, visto que apresentam expressão apenas nas plantas R, ou ocorrem em níveis mais altos nas mesmas.

Em trabalho recente, PANDEY et al. (2011), a partir de transcriptoma previamente publicados (MORTEL et al., 2007; CHOI et al., 2008; PANTHEE et al., 2009), conduziu a caracterização funcional de genes potencialmente envolvidos na

resposta de resistência mediada por *Rpp2* a *P. pachyrhizi*. Através da estratégia de silenciamento gênico induzido por vírus (VIGS), foram silenciados 64 FT WRKY e 140 potenciais genes, sendo identificados 11 genes que quando silenciados causaram a reversão do fenótipo de resistência para suscetibilidade.

A triagem e silenciamento dos 140 genes de defesa conhecidos resultaram na identificação de *GmEDS1* (*Enhanced disease susceptibility1*), *GmPAD4* (*Phytoalexin Deficient4*), e *GmNPR1* (*Non-expressor of PR genes 1*) como essenciais para a sinalização de defesa em *Rpp2*. O silenciamento desses três genes resultou em lesões do tipo tan, fenótipo de plantas suscetíveis, redução do mRNA transcritos dos genes alvos, quando comparados aos controles (sem inóculo e vetor vazio) e aumento de α -*tubulina* (crescimento do fungo).

O silenciamento de 64 FT WRKY resultou na identificação de três, *GmWRKY36*, *GmWRKY40*, *GmWRKY45*, relacionada à resistência mediada por *Rpp2*. O silenciamento de *GmDBFT*, que apesar de não possuir o domínio WRKY de ligação ao DNA, possui um domínio C₂H₂ encontrado em WRKY, também resultou na perda de resistência em *Rpp2*. Foi testado ainda a influencia de FT *GmWRKY36* e *GmDBFT*, os fatores com forte fenótipo de perda de resistência, para identificar genes que são potencialmente regulados por eles através da análise de expressão de genes em plantas cujo esses fatores foram silenciados. Para isso foi realizado uma busca entre 635 genes de soja induzidos durante a resposta de defesa na interação *Rpp2*-FAS nos estudos de MORTEL et al. (2007), que possuíam regiões promotoras com W-boxes, que são os elementos regulatórios de reconhecimento dos fatores de transcrição do tipo WRKY.

A análise dos níveis de transcrição destes genes em plantas com *GmDBFT* e *GmWRKY36* silenciados, revelou que a expressão em resposta a FAS foi em *GmWRKY12*, *GmWRKY19*, *GmPROT1* afetado por ambos FT; *GmFMO*, *GmO-MT*, *GmPR1* afetado apenas por *WRKY36*, em *GmPAL1* por *GmDBFTi*.

A partir destes resultados os autores especulam que *GmWRKY36* e *GmDBFT* desempenham significativo papel na resistência mediada por *Rpp2*, regulando processos de lignificação, uma vez que controlam a expressão de duas enzimas chaves no processo de produção da lignina, a fenilalanina amônia liase, que atua na conversão de fenilalanina em ácido *trans*-cinamico, e o-metiltransferase, que catalisa a transferência de um grupo metil de ácido cafeico para ferulato. Observou-se ainda a redução em 30% no conteúdo total de lignina em plantas com

GmO-MT silenciadas, o que indica papel importante para o acúmulo desta substância durante a resistência em *Rpp2*. *GmWRKY36* por regular *GmPR1* poderia atuar regulando, via PR1, a expressão de genes relacionados à patogênese (PR), tais como quitinase e beta-1,3-glucanases.

Outro grupo de fatores de transcrição testados no trabalho foram os MYBs, envolvidos na regulação do metabolismo de fenilpropanóides, dos quais entre os 20 genes testados, o silenciamento de *GmMYB84* resultou na perda de resistência mediada por *Rpp2*. A análise de expressão de seis genes preditos para serem envolvidos na via de fenilpropanóides, após silenciamento de *GmMYB84*, revelou a down-regulação de fenilalanina amônia-liase (*GmPAL1*), 4-coumarato-CoA-ligase (*Gm4-CL*), Cinnamato-4-hidroxilase (*GmC4H*) e Chalcona sintase (*GmCHS*), indicando que esta via também possui importância na defesa frente à FAS.

Em recente trabalho, LINO (2011) realizou a caracterização dos genes transcritos durante a infecção por *P. pachyrhizi* em dois genótipos de soja, PI561356 e BRS231, por meio de sequenciamento de três bibliotecas subtrativas (SSH) de cada genótipo, divididos em genes expressos 12, 24, 48 hai (Bulk 1); 72 e 96 hai (Bulk 2) e 192 hai (Bulk 3). Os resultados estão disponíveis no banco de dados do LGE (Soybean Genome Project) como parte do consórcio GENOSOJA (<http://bioinfo03.ibi.unicamp.br/soja/>). A categorização funcional permitiu agrupar os genes ativados na PI561356 em 14 principais processos, sendo os grupos com maior número de sequências o de resposta de estresse (421 sequências), resposta a estímulos abióticos (334) e transporte. A análise dos genes induzidos em BRS231 gerou 20 categorias, sendo resposta a estresse (149), óxido-redução (93) e respostas a estímulos químicos (70) as três categorias com maior número de sequências. Dentro do grupo funcional de regulação e expressão gênica foram identificados 14 diferentes fatores de transcrição (Auxin response factor, Helix-loop-helix, Bel-like Homeodomain, BZIP Ethylene responsive TF, GATA, KNOX homeobox, Leucine zipper, MYB, NAC, NDX homeobox, TFII, Zinc Finger, WRKY), entre os quais três estavam presentes em estudos de análise de transcritos na interação da soja com *P. pachyrhizi*, anteriormente citados: WRKY (MORTEL et al. 2007; PANDEY et al. 2011), NAC (MORTEL et al., 2007; TREMBLAY et al., 2010) e MYB (MORTEL et al., 2007; CHOI et al., 2008; PANTHEE et al., 2009). Ressalta-se ainda que todos eles foram induzidos em ambos os genótipos analisados.

2.6 FATORES DE TRANSCRIÇÃO MYB ENVOLVIDOS NA RESPOSTA DE RESISTÊNCIA A DOENÇAS EM PLANTAS

A regulação da expressão gênica é responsável pelo controle dos processos biológicos nos organismos, que incluem desde o desenvolvimento celular a complexos mecanismos de resposta a estímulos ambientais, dentre os quais, as respostas a estresses bióticos e abióticos observados nas plantas (FELLER et al., 2011).

Apesar do grande número de pontos sobre os quais é possível realizar o controle da expressão (regulação a nível transcricional, traducional e pós-traducional), para a maioria dos genes este controle é realizado no início da transcrição (transcrição do DNA para mRNA), visto que garante menor gasto celular de energia e assegura a produção de proteínas em situações e níveis adequados para a sua atividade biológica.

Como reflexo, uma proporção significativa do genoma de muitos organismos são dedicados a desempenhar papéis no controle da expressão gênica. Cerca de 3,5 a 5,9% do genoma de plantas e alguns eucariotos são codificadores de fatores reguladores de transcrição (RIECHMANN et al., 2000). O genoma de *Arabidopsis thaliana* inclui 27416 genes codificadores de proteínas, dos quais 6% (mais de 1700) codificam fatores de transcrição (FELLER et al., 2010). Em *Glycine max*, das 46430 sequências codificadoras de proteínas, 5671 sequências são de fatores de transcrição (12%) (WANG et al., 2010).

De modo geral, a região de controle gênico, onde ocorrem os principais processos que modulam a transcrição, envolvem três elementos: os promotores, (local onde a RNA polimerase e os fatores de transcrição basais se associam), elementos proximais a promotores que contem sítios para ligação de fatores de transcrição induzíveis (auxiliam a ligação da RNA polimerase ao promotor) e regiões *enhancer* e *silencers* (aumentam ou reprimem a transcrição). Este último elemento (*silencer*) (com menos de 20 pares de nucleotídeos) possui a capacidade de regular o início da transcrição aumentando a frequência com a qual ocorre ou reprimindo-a (ALBERTS et al., 2006).

O controle deste processo envolve ainda mais um componente: proteínas com capacidade de reconhecer e se ligar a essas regiões, os fatores de transcrição. Cabe a este elemento o papel de ativar ou desativar um conjunto

específico de genes após o reconhecimento destes sítios. Estas proteínas reconhecem na porção externa do sulco maior da dupla fita de DNA, padrões específicos dos diferentes arranjos de quatro ou mais pares de base, tais como grupamentos (doadores e aceptores de hidrogênio), sinais hidrofóbicos e a distorção local da hélice (conformação espacial) resultante da composição da região. Deste modo a especificidade encontrada na organização destes padrões de acordo com sua composição de pares de base, permite as proteínas reguladoras distinguir uma sequência de DNA de outra (ALBERTS et al., 2006). Estas regiões são denominadas cis-elementos.

Os fatores de transcrição encontrados em eucariotos são divididos em dois grupos. O primeiro corresponde a fatores basais (Fatores Gerais de Transcrição - *TFI*), que auxiliam nas etapas de iniciação da transcrição: posicionamento da RNA polimerase II no promotor; abertura da dupla fita de DNA e liberação da polimerase do promotor para iniciar a extensão do transcrito. No entanto eles não possuem a capacidade de alterar a frequência na qual ocorre o processo de iniciação. São proteínas que se associam a qualquer promotor utilizado pela RNA polimerase, e estão presentes em todas as células.

No segundo grupo estão os fatores que distinguem, de forma específica, as sequências reguladoras, sendo assim capazes de modificar a frequência (aumento ou repressão) com a qual os genes sobre o controle destas sequências reguladoras são expressos. (ALBERTS et al., 2006; SILVEIRA, 2007). A modulação da transcrição, contudo, não é produto da atividade de apenas um fator de transcrição e, muitas vezes, envolve mais de um desses fatores que desempenham papéis sinérgicos sobre o controle da expressão.

Frequentemente, estas proteínas possuem uma estrutura modular, que compreende um domínio de ligação ao DNA junto de um domínio regulador (ativador ou repressor), sendo então capazes de reconhecer o DNA de uma forma específica e que regulam a frequência de iniciação da transcrição após a ligação a estes sítios específicos da região promotora dos genes alvos. Este domínio de ligação ao DNA apresenta sequências de aminoácidos típicos, permitindo, com base na similaridade desta sequência, classificar os fatores de transcrição em diferentes famílias (STRACKE; WEBER; WEISSHAAR, 2011).

Fatores de transcrição podem ser encontrados em todos os eucariotos, contudo são muito mais numerosos em plantas (STRACKE; WEBER;

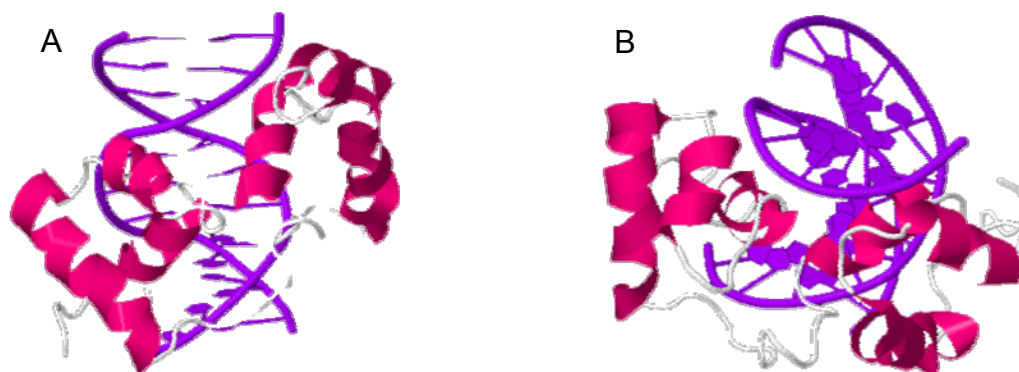
WEISSHAAR, 2001), como é o caso da família de TF MYB (RIECHMANN et al., 2000), os quais são 20 vezes mais numerosos em *Glycine max* quando comparado ao homem (FELLER et al., 2011)

2.6.1 Fatores de Transcrição MYB

A grande família de proteínas MYB é funcionalmente diversa e está presente em todos os eucariotos nos quais sua presença tem sido investigada (KRANZ et al., 1998; DUBOS et al., 2010; FELLER et al., 2011). O primeiro gene codificador de uma proteína contendo o domínio MYB foi o *v-myb* foi identificado no genoma do vírus da myeloblastose aviária (AMV), o qual também é responsável pelo nome do domínio (KLEMPNAUER et al., 1982). O primeiro MYB caracterizado em plantas foi *C1 (COLORED1)* de *Zea mays* que regula a biossíntese de antocianina na aleurona de sementes (PAZ-ARES et al., 1987).

A característica comum de todas as proteínas MYB é a presença de domínio de ligação ao DNA altamente conservado formado por um a quatro repetições imperfeitas de aminoácidos (R) que podem atuar em sinergia para se ligarem ao DNA ou individualmente em interações proteína-proteína (com o fator de transcrição basal ou maquinaria de transcrição). Cada uma das repetições MYB é composto por 50 a 53 aminoácidos de comprimento, com três clusters de triptofanos regularmente espaçados (três hélices), formando juntos um núcleo hidrofóbico. Cada repetição adota uma conformação hélice-volta-hélice (helix-turn-helix, HTH), que é um motivo estrutural capaz de se ligar ao DNA, composto por alfa-hélices unidas por uma porção curta de aminoácidos. As três repetições são referidas como R1, R2 e R3. A segunda e a terceira alfa-hélice (hélice de reconhecimento do DNA) de cada repetição se ligam ao sulco maior (principal) do DNA, e em domínios MYB com duas ou mais repetições MYB, a ligação com o DNA regulado envolve o reconhecimento sinérgico do alvo pela terceira alfa-hélice de cada repetição (FELLER et al., 2011) (Figura 2).

Figura 2 – Estrutura tridimensional de R2R3-MYB. A e B – Reconhecimento do sulco maior da molécula de DNA pela 2ª e 3ª alfa hélice de R2R3.



Fonte: Protein Data Bank (<http://www.rcsb.org/pdb/explore/jmol.do?structureId=1MSE>).

Proteínas MYBs podem ser classificadas em subfamílias dependendo do número de repetições (R) nos domínios MYB (Figura 4). Denominam-se proteínas MYB com uma repetição como 1R-MYB, com duas repetições como R2R3-MYB, com três de 3R-MYB e quatro 4R-MYB. As quatro classes são encontradas nas plantas, e, desta forma, representam o táxon com a maior diversidade de fatores de transcrição MYB. A classe 4R-MYB é a menor e menos conhecida delas, sendo formada por quatro repetições de 1R e 2R. A segunda menor classe possui MYB com repetições 1R2R3R (3R-MYB), ligados ao controle do ciclo celular, e normalmente são encontrados cinco genes para esta classe em genomas de plantas. A terceira classe é a 1/2R-MYB (MYB-related), composto por proteínas com uma repetição, inteira ou parcial, que exercem controle sobre o ciclo circadiano, morfogênese celular e desenvolvimento de órgãos. Dentre essas classes, a maior é aquela formada por proteínas com duas repetições, R2R3-MYB (DUBOS et al., 2010), formado por mais de 100 membros em algumas espécies de plantas, como *A. thaliana*, *G. max*, *Zea mays*, *Oryza sativa*, *Vitis vinifera* (Figura 5). A classe R2R3-MYB representa 56,77% e 70,05% dos MYBs em arroz e *Arabidopsis*, respectivamente (KATIYAR et al., 2012), e é estimada por FELLER et al., (2011) em 50% da família em *G. max*.

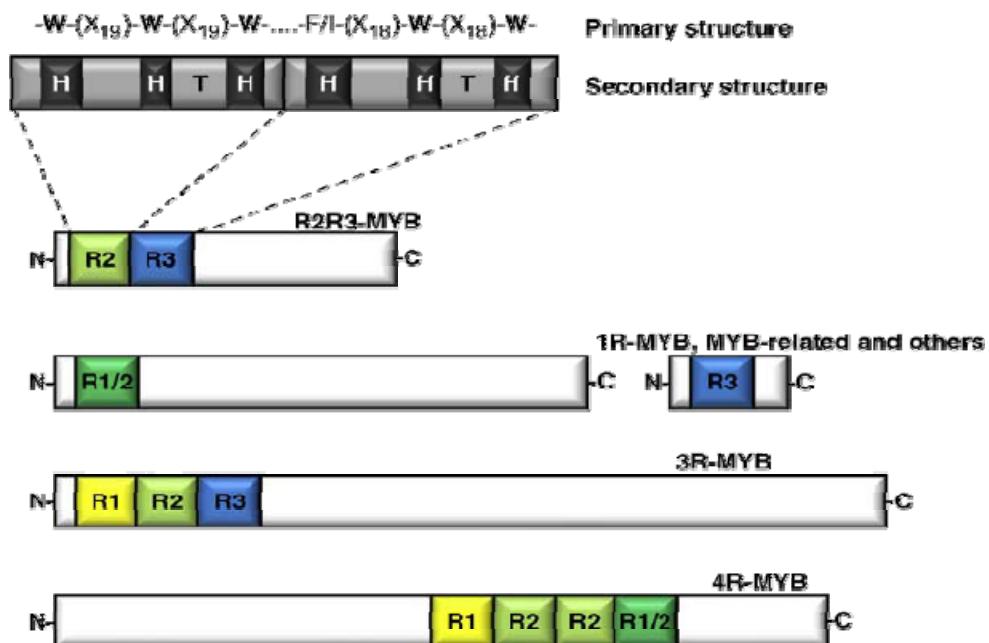
Em *A. thaliana* MYBs representam 10% dos 1600 FT identificados (PROUSE, CAMPBELL, 2012), enquanto apenas três cópias de genes desta classe são descritos no genoma de *Saccharomyces cerevisiae*, *Drosophila melanogaster* (RIECHMANN et al., 2000) e no homem (LIN-WANG et al., 2010).

Outras diferenças observadas entre MYBs R2R3 de plantas e animais são a presença de um resíduo L (leucina) entre a segunda e terceira hélice

da repetição R2 e o primeiro resíduo de triptofano (W) trocado por outro aminoácido hidrofóbico em plantas (DUBOS et al., 2010).

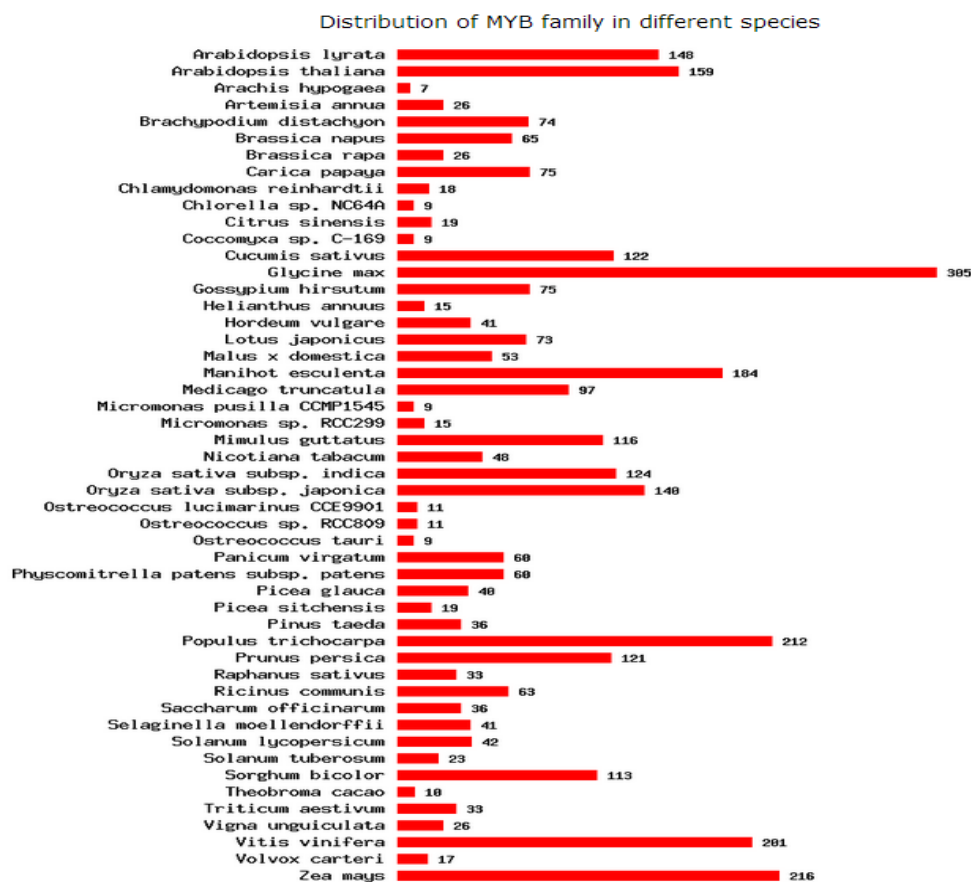
Em *A. thaliana* inicialmente foram identificados 80 genes de R2R3-MYBs a partir de uma biblioteca genômica (ROMERO et al., 1998), e posteriormente o número de membros desta família subiu para 97 (KRANZ et al., 1998). STRACKE, WEBER; WEISSHAAR, (2001) incorporaram dados de diferentes estudos de AtMYBs a buscas de similaridade realizadas a partir de ferramentas como BLAST e HMMer, o que resultou na identificação de 125 genes distintos de AtMYBs (confirmados utilizando primers específicos, através de RT-PCR). Recentemente DUBOS et al.(2010) revisou os avanços nos conhecimentos da família R2R3-MYB de *Arabidopsis thaliana*, com ênfase nos papéis funcionais em 135 identificados AtMYBs nos últimos anos.

Figura 3 – Classes de fatores de transcrição MYB em plantas. H, hélice; T, volta; W, triptofano; X, aminoácido; R, repetições



Fonte: DUBOS et al. (2010).

Figura 4 – Distribuição da classe R2R3 MYB em diferentes espécies de plantas.



Fonte: Transcription Factor Database (<http://planttf.db.cbi.edu.cn/family.php?fam=MYB>).

Tabela 1 – Classificação da família R2R3-MYB de *A. thaliana*

Subgrupo	R2R3 MYB <i>A. thaliana</i>	Funções
S20	AtMYB2,AtMYB62,AtMYB78,AtMYB108,AtMYB112,AtMYB16	Defesa
S2	AtMYB13,AtMYB14,AtMYB15	Defesa
S22	AtMYB44,AtMYB70,AtMYB73,AtMYB77	Defesa
S2	AtMYB30,AtMYB31,AtMYB60,AtMYB94,AtMYB96	Defesa
S1	AtMYB41,AtMYB74,AtMYB102, AtMYB49	Defesa
S11	AtMYB26, AtMYB67	Defesa
S18	AtMYB0,AtMYB23,AtMYB66	Desenvolvimento/defesa
S15	AtMYB36,AtMYB37,AtMYB38,AtMYB68,AtMYB84,AtM87	Diferenciação celular
S14	AtMYB21,AtMYB24, AtMYB57	Desenvolvimento
S19	AtMYB22,AtMYB64,AtMYB100,AtMYB115,AtMYB118,AtMYB119, AtMYB98	Desenvolvimento
S25	AtMYB18,AtMYB19,AtMYB45	Desenvolvimento
S16	AtMYB52,AtMYB54,AtMYB56,AtMYB69,AtMYB105,AtMYB110,AtM YB117	Desenvolvimento
S21	AtMYB50,AtMYB55,AtMYB61,AtMYB86	Metabolismo/desenvolvimento
S13	AtMYB3,AtMYB7,AtMYB32, AtMYB8, AtMYB4, AtMYB6	Metabolismo/desenvolvimento
S4	AtMYB3,AtMYB7,AtMYB32, AtMYB8, AtMYB4, AtMYB6	Metabolismo
S6	AtMYB90,AtMYB113,AtMYB114, AtMYB75	Metabolismo
S7	AtMYB11,AtMYB12,AtMYB111	Metabolismo
S12	AtMYB28,AtMYB29,AtMYB34,AtMYB51,AtMYB76,AtMYB122	Metabolismo
S3	AtMYB058, AtMYB063, AtMYB72, AtMYB10	Metabolismo
S5	AtMYB123	Metabolismo

Fonte: DUBOS et al. (2010).

Estudos em *Arabidopsis thaliana* apontam que proteínas desta classe de MYB estão envolvidas no controle de processos específicos da planta, incluindo o metabolismo primário e secundário, diferenciação celular, processos de desenvolvimento e respostas a estresse abiótico e biótico (FELLER et al., 2011), exibindo uma diversidade funcional maior que o observado em animais, onde em sua maioria estão associados apenas ao controle da proliferação e diferenciação de células e a prevenção da apoptose (ROMERO et al., 1998).

Proteínas da classe R2R3-MYB possuem uma estrutura modular, com o domínio MYB (de ligação ao DNA) na região altamente conservada N-terminal e domínio de ativação ou repressão transcricional conservados na região altamente variável C-terminal (DUBOS et al., 2010). A conservação do domínio MYB e motivos de aminoácidos na região C-terminal permite classificar estes FT em 22 subgrupos, assim como realizado por KRANZ et al. (1998); STRACKE, WEBER; WEISSHAAR, (2001); DUBOS et al. (2010) em *A. thaliana* e em *G. max* (DU et al., 2012), entre os quais, alguns com funções comuns para seus membros conhecidos, incluindo resposta a estresse biótico.

KRANZ et al. (1998) identificou 97 genes R2R3-MYBs de *A. thaliana* (AtMYB) a partir de dados do projeto de sequenciamento de *Arabidopsis* associada à análise de cDNAs obtidos a partir de banco de dados de ESTs (sequências expressas). Com base na similaridade das proteínas e das sequências dos motivos conservados na porção C-terminal, as proteínas MYB foram classificadas em 22 subgrupos (Figura 6).

O alinhamento de 125 sequências identificadas por STRACKE, WEBER; (2001) resultou nos mesmos 22 subgrupos observados por KRANZ et al. (1998) e em dois novos motivos, devido ao aumento no número de membros da família AtMYB. Os autores notaram que em alguns subgrupos os membros agrupados possuíam funções conservadas. Em trabalho de DUBOS et al. (2010), o alinhamento de 135 sequências de AtMYB conhecidas resultou nos mesmos subgrupos observados por STRACKE, WEBER; WEISSHAAR, (2001), o que permitiu relacionar os subgrupos (e seus membros) a funções específicas identificadas nos últimos anos (Figura 6).

infecção pelo patógeno induz a síntese de SA, o qual promove a expressão de AtMYB44, que parece regular diretamente o gene *PR1*.

Os fatores AtMYB60 e AtMYB96 atuam controlando a cascata de sinalização mediada por ABA, regulando a abertura e fechamento dos estômatos em resposta a invasão da planta por patógenos (DUBOS et al., 2010). AtMYB31 foi identificado em *A. thaliana* como sendo responsivo concomitantemente a genes de defesa e outros 11 AtMYBs ao elicitor quitina (hidrolisada e quito-octamero), componente estrutural da parede celular de fungos (LIBAULT et al., 2007). AtMYB51 e AtMYB27 são induzidos em plantas infectadas com *Alternaria brassicola* (McGRATH et al., 2005) e também ao tratamento com quitina (LIBAULT et al., 2007). AtMYB108 é um fator de transcrição que possivelmente controla a via de sinalização de síntese de JA, importante hormônio para respostas contra patógenos necrotróficos. Mutação no gene deste fator de transcrição promoveu aumento da susceptibilidade à *Botrytis* e *Alternaria brassicola*, dois fungos necrotróficos, fato não observado em plantas selvagens, cuja expressão deste fator e, conseqüentemente, a síntese de JA não está comprometida (MENGISTE et al., 2003).

Um gene similar a AtMYB2 (*Jamyb*), isolado a partir de uma biblioteca de genes induzidos em resposta ao ataque fúngico de *Pyricularia grisea* em *Oriza sativa* (cultivar Drew), apresentou expressão durante a formação de lesões decorrentes do processo de morte celular programada e em resposta a aplicação de JA, indicando potencial papel nas respostas de resistência induzidas contra patógenos e herbivoria (LEE; QI; YANG, 2001).

Deste modo, devido a tantas evidências do envolvimento de diversos fatores de transcrição das famílias WRKY, NAC e MYB nos mecanismos de defesa nas mais diversas espécies de plantas, assim como durante a interação soja – *P. pachyrhizi*, faz-se necessário esforços para a caracterização dos fatores de transcrição MYB em soja, visando identificar seu papel específico na resposta ao patógeno e à outras fontes de estresse biótico. Tal estudo poderá contribuir a descoberta de novos genes e rotas importantes e os mecanismos envolvidos na resposta mediada por eles.

3 OBJETIVO GERAL

O presente trabalho objetivou caracterização *in silico* e *in vivo* de genes pertencentes à família dos fatores de transcrição R2R3-MYBs em *Glycine max*, bem como identificar aqueles envolvidos na resposta à *P. pachyrhizi*.

3.1 OBJETIVOS ESPECÍFICOS

- ✓ Caracterizar *in silico* a família de fatores de transcrição R2R3-MYB em soja (*Glycine max*) quanto à filogenia, distribuição em famílias;
- ✓ Caracterizar *in silico* o perfil transcricional dos fatores de transcrição MYB R2R3 via busca em Bancos de Transcriptoma, com foco naqueles envolvidos na resposta a doenças;
- ✓ Caracterizar *in vivo* o envolvimento de fatores de transcrição R2R3-MYB na resposta a *P. pachyrhizi* via análise do perfil transcricional (RT-qPCR) em genótipos de soja contendo diferentes genes de resistência a este patógeno.

4 MATERIAL E MÉTODOS

4.1 BUSCA E ANOTAÇÃO DE GENES MYBs R2R3 DE *Glycine max* A PARTIR DE BANCO DE DADOS

Sequências proteicas preditas e anotadas como MYBs no genoma de *Glycine max* foram obtidas e combinadas a partir de três diferentes bancos de dados: Plant Transcription Factor Database - Planttfdb (<http://planttfdb.cbi.edu.cn/>), SoyDB (<http://casp.rnet.missouri.edu/soydb/>) e Phytozome (<http://www.phytozome.net/>). Para complementar a busca nos dados do Phytozome, foi realizada uma busca por similaridade pela ferramenta Blastp, a partir do perfil do Modelo de Markov (HMM – Hidden Markov Model) de MYB obtido no Pfam (<http://pfam.sanger.ac.uk/family/PF00249#tabview=tab3>), e busca por palavra-chave (MYB), a fim de se obter sequências adicionais. Os resultados das buscas foram analisados e combinados de forma a excluir dados repetidos e obter um conjunto de sequências únicas.

Todas as sequências foram submetidas a análises de domínio no Pfam (<http://pfam.sanger.ac.uk/search>), SMART (<http://smart.embl-heidelberg.de/>) e MEME (<http://meme.sdsc.edu/meme/cgi-bin/meme.cgi>), utilizando um e-value de corte de 1e-10, para exclusão de sequências sem o domínio característico R2R3-MYB e/ou domínio sem significância ou incompletos. Para complementar a análise, realizou-se ainda a inspeção manual de todas as sequências (Figura 5).

4.2 CARACTERIZAÇÃO TRANSCRICIONAL DE MYBs DE *Glycine max* EM BANCOS DE TRANSCRITOMAS

Após a confirmação das sequências como MYB-R2R3, o perfil de expressão dos genes dos candidatos foi avaliado em bancos de expressão, com o objetivo de se averiguar a sua funcionalidade e se excluir modelos gênicos incorretos. Foram consultados os seguintes bancos: Soybase – (<http://soybase.org>), Genosoja (LGE–SuperSAGE e Subtrative library-<http://bioinfo03.ibi.unicamp.br/soja/>) e Genevestigator – (<https://www.genevestigator.com/gv/>). Estes bancos reúnem ESTs de soja identificados nos seguintes experimentos:

1. Genosoja:

- a) SSH: Caracterização dos genes expressos em PI561356 durante infecção por *P. pachyrhizi*, por meio de sequenciamento de três bibliotecas subtrativas, divididos em três bulks - 12, 24, 48 hai (Bulk 1); 72 e 96 hai (Bulk 2) e 192 hai (Bulk 3);
- b) SuperSAGE-LGE: biblioteca SuperSAGE criada a partir ensaios de folhas de soja PI561356 (resistente) infectadas com *P. pachyrhizi*, nos tempos de 12, 24 e 48 hai, e controles compostos de folhas sem infecção.
- c) RNA-Seq: biblioteca criada a partir de dois genótipos, TMG 115 RR e BRSGO Chapadões, suscetível e resistente, respectivamente, à infecção pelo nematóide *Pratylenchus brachyurus*, submetidos à inoculação do patógeno, nos períodos de tratamento, 24, 48, 96 e 192 horas após a inoculação Ambos disponíveis em www.lge.ibi.unicamp.br/soja/.

2. Genevestigator:

- a) Ensaio formado por plantas de soja (V3) PI462312 (Rpp3) em três tratamentos: com cepa virulenta de *P. pachyrhizi* (TW80-2), com cepas avirulentas (HW94-1) e controles (mock), As análises foram realizadas em 6 pontos: 12hpi (a), 24hpi (b), 72hpi (c), 144hpi (d), 216hpi (e), 288hpi (f). Dados do experimento disponíveis em: http://www.plexdb.org/modules/PD_browse/experiment_browser.php?experiment=GM36;
- b) ***P. pachyrhizi* estudos 2 e 3:** Análise do perfil de expressão de soja PI200492 (Rpp1) com dois tratamentos (HW94-1 ou TW72-1) em diferentes tempos : *P. pachyrhizi* estudo 2 (early, 6-12hpi)/non-infected leaf samples (PI200492)(2a), *P. pachyrhizi* estudo 2 (intermediate-24hpi)/non-infected leaf samples (PI200492)(2b), *P. pachyrhizi* estudo 2 (late48hs)/non-infected leaf samples (PI200492)(2c), *P. pachyrhizi* estudo 3 (early, 6-12hpi)/non-infected leaf samples (PI200492)(3a), *P. pachyrhizi* estudo 3 (intermediate – 24hpi)/non-infected leaf samples

- (PI200492)(3b), *P. pachyrhizi* estudo 3 (late-48hpi)/non-infected leaf samples (PI200492)(3c). Disponível em <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE8432>;
- c) ***P. pachyrhizi* estudos 4 e 5:** Ensaio com folhas de soja PI230970 (Rpp2) e Embrapa-48, infectadas com FAS, em: *P. pachyrhizi* estudo 4 (early, 6-24hpi)/mock inoculated PI230970 leaf samples (4a), *P. pachyrhizi* estudo 4(intermediate, 36-72hpi)/mock inoculated PI230970 leaf samples(4b), *P. pachyrhizi* estudo 4 (late, 96-168hpi)/mock inoculated PI230970 (4c), *P. pachyrhizi* estudo 5 (early, 6-24hpi)/mock inoculated embrapa48 leaf samples (early)(5a);*P. pachyrhizi* estudo 5 (intermediate, 36-72hpi)/mock inoculated embrapa48 leaf samples (5b), *P. pachyrhizi* estudo 5 (late, 96-168hpi)/mock inoculated embrapa48 leaf samples (5c), *P. pachyrhizi*/non-infected leaf samples (5601t)_ (6). Disponível: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE33>;
- d) ***Aphis glycines* x soja (V3):** Análise da resposta transcricional de plantas de soja à infestação de pulgões de soja resistentes e suscetíveis. Experimento com três fatores: variedade de soja em V3 (suscetível SD01-76R, resistente LD05-16060), o tratamento de pulgões (controle, pulgões), e duração de infestação (1 dia e 7 dias). Dados disponíveis em: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE35427>. (a): 1dpi, (b): 7dpi (estudo 2);
- e) ***Heterodera glycines*:** Análise de perfis de transcrição de raízes de soja (Williams 82), via microarranjo, em três pontos de tempo de infecção (2,5,10 dias após inóculo - dpi) com nematódeos císticos de soja. Dados disponíveis em: <http://www.ebi.ac.uk/arrayexpress/browse.html?keywords=%09+E-MEXP-808>. (a): estudo 1, (b): estudo 2.
- f) ***P. pachyrhizi* x soja 5601T:** Análise de expressão gênica por microarranjo em plantas de soja infectadas com o fungo *Phakopsora pachyrhizi* (72hpi) na fase de crescimento V2.

Dados disponíveis em:
<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE7108>.

- g) ***Phytophthora sojae***: Identificação de genes diferencialmente expressos durante infecção pelo patógeno *P. sojae* em 8 cultivares de soja que diferem na resistência quantitativa (Athow, Conrad, General, Ox20-8, PI291327, Sloan, V71-370, Williams), através de microarranjo. As análises foram realizadas em tempos: 3dpi (a) e 5dpi (b). Dados disponíveis em: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE7124>; Padrões de expressão em diferentes tempos, durante a infecção por *P. sojae* (isolado PT2004C2.S1), de cultivares de soja que diferem na resistência quantitativa (resistentes a cepa – CONRAD e V71-370, suscetíveis – SLOAN e VP-RIL9): estudos (3) (24hpi), (4) (48hpi), (5) (72hpi), (6) (120hpi), disponível em: http://www.plexdb.org/modules/PD_browse/experiment_browser.php?experiment=GM3 e estudos (7)(24hpi), (8) (48hpi), (9) (72hpi), disponível em: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE11611>. Estudo 10 – Amostras de hipocótilo de soja Harosoy (susceptível) inoculados com a cepa P6497 (raça 2) de *P. sojae*, foram analisados através de microarranjo em 12 (a), 24 (b), 48 (c) horas após o inóculo (hpi). Todos os resultados foram obtidos a partir de ensaios com microarranjo
3. Soybase: os dados se referem ao atlas de expressão em soja de todos os modelos gênicos preditos a partir do sequenciamento do genoma analisados ao longo do ciclo da espécie, nos diferentes órgãos: raiz, folha, flor, vagem e semente de *G. max*, através de RNA-Seq (SEVERIN et al., 2010).

Foram avaliados um total de dois experimentos de sequenciamento de alta performance disponível no banco do Genosoja (SSH e SuperSAGE) e 9 experimentos de microarranjos (Genevestigator) que permitiram traçar o perfil transcricional da família R2R3 de GmMYBs frente a diferentes estresses bióticos em

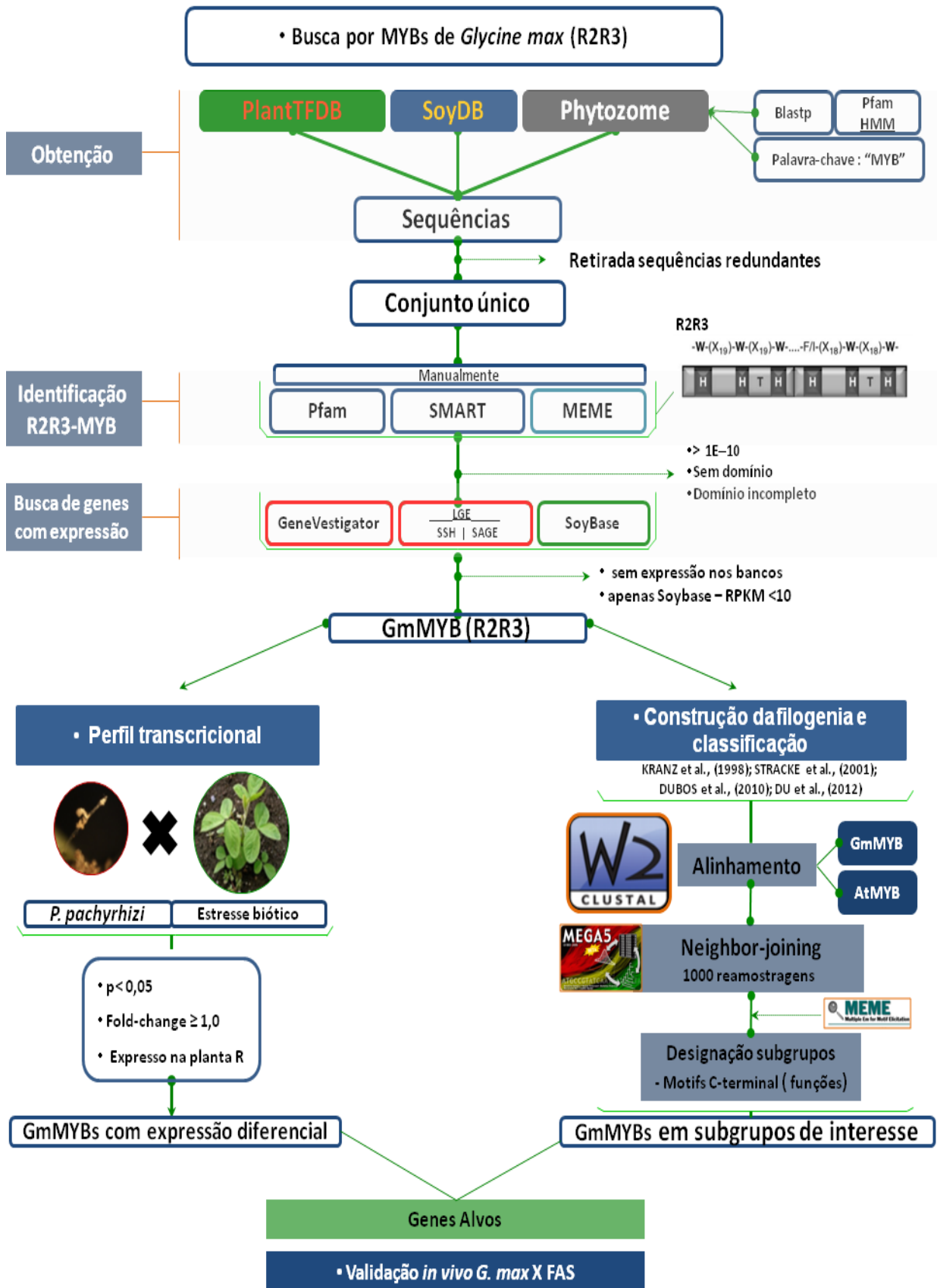
soja. Também foram analisados a expressão dos modelos gênicos em resposta a todos os demais estresses e ensaios disponíveis em todos os bancos consultados para a exclusão dos modelos gênicos que não estavam presente em pelo menos uma das fontes consultadas (não funcional).

4.3 CONSTRUÇÃO DA FILOGENIA E CLASSIFICAÇÃO EM SUBGRUPOS DOS MYBS

Para a análise filogenética, as sequências das proteínas Mybs de *Glycine max* que apresentaram domínios R2R3 completos foram alinhadas com sequências guias de Mybs (85 MYBs de *Arabidopsis thaliana*), com o propósito de classificá-los em subgrupos como o proposto para *A. thaliana* (STRACKE; WEBER; WEISSHAAR, 2001; KRANZ, 2000), através do programa ClustalX (versão 2.1). A árvore filogenética foi construída pelo método de neighbor-joining utilizando o programa MEGA (versão 5.05), realizando bootstrapping com 1000 re-amostragens para avaliar a representatividade dos agrupamentos. Após o agrupamento e a designação dos subgrupos, motivos conservados foram detectados na porção C-terminal das sequências utilizando o software MEME (<http://meme.sdsc.edu/meme/cgi-bin/meme.cgi>), com os parâmetros utilizados por JIANG; GU; PETERSON, (2004): distribuição de motivos: zero ou um por sequência; número de diferentes motivos: 16; comprimento mínimo do motivo: 6; comprimento máximo do motivo: 300 (modificado para identificar domínios R2R3 mais longos). Apenas motivos com $e\text{-value} \leq 1e-10$ foram utilizadas para todas as análises. A figura 7 apresenta o *pipeline* utilizado no estudo.

Para a identificação dos *cis*-motivos comuns da região promotora dos MYBs, 1Kb a montante do sítio de iniciação da tradução foi analisado utilizando a ferramenta MEME (Multiple Expectation maximization for Motif Elicitation analysis tool) (BAILEY et al., 2006) (MEME versão 4.9.0, <http://meme.nbcr.net/meme/cgi-bin/meme.cgi>). Foram utilizados os parâmetros adotados por KATIYAR et al. (2012), buscando número máximo de motivos: 5; comprimento mínimo do motivo: 8; comprimento máximo: 12; qualquer número de repetições, motivos encontrados apenas no sentido *forward* e com $e\text{-value} < 0,01$. Os motivos descobertos foram analisados no banco de *cis*-elementos de plantas PLACE (HIGO et al., 1999) (<http://www.dna.affrc.go.jp/PLACE>), a fim de se obter maiores informações a cerca dos mesmos.

Figura 6 – Esquema das diferentes etapas conduzidas para caracterização *in silico* e validação *in vivo* de genes MYB R2R3 em soja envolvidos na resposta às doenças



4.4 VALIDAÇÕES *in vivo* DE FATORES DE TRANSCRIÇÃO MYB R2R3 NA RESPOSTA A INFECÇÃO POR *P. pachyrhizi*

Para validar a metodologia proposta no presente trabalho referente às análises *In silico*, foram selecionados 11 genes Mybs potencialmente envolvidos em respostas às doenças em soja, que tiveram o perfil e expressão avaliados em resposta a infecção por *P. pachyrhizi* via RT-qPCR.

4.4.1 Material Biológico

Foram utilizados quatro genótipos de soja (estádio de desenvolvimento V2 – 21 dias): a cultivar Williams 82 (WH82) como padrão suscetível e as plantas introduzidas (PI) como resistentes a infecção por *P. pachyrhizi*, cada uma contendo um gene R específico, previamente caracterizados: PI 230970 (*Rpp2*) e Shiranui (*Rpp5*). Todas as sementes foram obtidas no Banco Ativo de Germoplasma da Embrapa – Soja (BAG).

As plantas foram mantidas em casa de vegetação sob condições de temperatura ($22 \pm 1^\circ\text{C}$) e umidade ($> 60\%$) controladas, e fotoperíodo de 12 h. O experimento seguiu delineamento inteiramente casualizado (DIC) com três repetições (cada repetição contendo cinco replicatas), e estrutura de tratamentos em fatorial com três fatores: tempo de coleta (12, 24, 48, 72 e 96 horas após inoculação - hai), genótipos (PI230970, Shiranui e Williams 82) e tipo de infecção (inoculado – “inoc”, e inoculado com água – “mock”). No total, 120 vasos foram utilizados, contendo cinco plantas cada, totalizando 600 plantas.

4.4.2 Inoculação de *P. pachyrhizi*

O inóculo consistiu de uma população fúngica coletada em campos experimentais da Embrapa Soja e mantida em casa de vegetação em cultivar suscetível. Os urediniósporos foram coletados batendo-se as folhas infectadas sobre uma bandeja plástica e diluídos em água destilada com 0,05% (v/v) de “Tween 20” (Uniqema) para uma concentração final de aproximadamente 5×10^4 esporos/mL. As plantas foram inoculadas quando atingiram o estágio de desenvolvimento V3-V4 de acordo com a escala proposta por FEHR; CAVINES, (1981). A inoculação das

plantas, utilizando-se pulverizador manual, foi realizada ao anoitecer para evitar a inviabilização dos uredósporos pela falta de água livre nas folhas e baixa umidade do ambiente, causados pelas temperaturas mais altas na casa de vegetação durante os horários de maior insolação.

4.4.3 Avaliação da Expressão dos Genes MYBs

4.4.3.1 Coleta do material

Para a análise dos perfis de expressão dos genes selecionados, foram coletados os 2º trifólios de cada planta de todos os vasos, e cada repetição, composto pelos trifólios das cinco plantas, foram reunidas (*bulk*). Os trifólios foram embalados e armazenados em envelopes de papel alumínio, e congelados através de imersão em nitrogênio líquido e armazenados em ultrafreezer a -80 °C em caixas.

4.4.3.2 Extração de RNA total

Cerca de 100 mg de cada folha das 90 amostras do ensaio de validação das cultivares PI230970, Shiranui e Williams82 foram macerados em graal, previamente resfriado em nitrogênio líquido, com auxílio de pistilo e armazenados em microtubo de 1,5 mL. Posteriormente, 1 mL de reagente TRIZOL (Invitrogen™) (55 °C) foi adicionado e, após mistura em vórtex (temperatura ambiente) por 10 minutos, centrifugou-se a amostra por 15 minutos a 14000 RPM (20000xg) (4 °C). O sobrenadante foi transferido para novo microtubo, e após a adição de 400 µL de clorofórmio, a amostra foi misturada em vórtex por cinco minutos em temperatura ambiente. Decorrido este tempo, o material foi centrifugado a 14000 RPM (20000xg) por 10 minutos. Cerca de 400 µL do sobrenadante foram transferidos para novo microtubo, e repetiu-se o processo.

Aproximadamente 200 µL do sobrenadante foram transferidos para microtubo novo, ao qual se adicionou 250 µL de tampão (0,8M citrato de sódio; 1,2M cloreto de sódio) e 250 µL de isopropanol. O material foi misturado em vórtex (temperatura ambiente) por 10 minutos e centrifugado por 10 minutos a 14000 RPM (20000xg) a 4 °C. Ao final da centrifugação o sobrenadante foi descartado e o *pellet* dissolvido em 100 µL de água deionizada. Adicionou-se 10 µL de acetato de sódio e

250 μL de etanol 100%, seguido de mistura em vórtex por cinco minutos e centrifugação por 15 minutos (20000xg, 4 °C). Cuidadosamente o sobrenadante foi removido e ao *pellet*, cerca de 400 μL de etanol 75% foram adicionados. Após nova centrifugação por 15 minutos (20000xg, 4 °C), o sobrenadante foi descartado e o *pellet* seco em vácuo por 20 minutos. Após a extração, a amostra foi ressuspensa em 30 μL de água deionizada e armazenada em ultrafreezer a -80°C.

A concentração do RNA foi determinada após leitura da amostra em espectrofotômetro (nanodrop) a 260 nm e a integridade da amostra após eletroforese em gel de agarose 1%, utilizando o tampão de corrida TBE (400 μL de EDTA 0,5M, 10,8g de Tris e 5,5g de ácido bórico) e marcador (ladder) de 100pb.

4.4.3.3 Síntese do cDNA

Primeiramente 1 μg de RNA de cada amostra foi tratada com 0,5 μL DNase I, adicionado de 0,5 μL de tampão 10X e água para completar o volume final de 5,0 μL . A reação foi incubada por sete minutos em temperatura ambiente. Em seguida foi adicionado 0,8 μL de EDTA e uma nova incubação foi realizada por cinco minutos a 65°C.

Em microtubos uma reação contendo 1 μg de RNA previamente tratado com DNase, 1 μL de oligo dT (concentração final de 50 μM) e água tratada com DEPC para completar o volume de 10 μL foram incubadas a 65°C por cinco minutos. Em seguida adicionou-se o mix para a síntese do cDNA composto por 2 μL de tampão RT 10X, 4 μL de MgCl_2 (25mM), 2 μL de DTT (0,1M), 1 μL de inibidor de RNase (RNase OUT 40U/ μL) e 1 μL de enzima transcriptase reversa (200U/ μL). A reação foi incubada em termociclador (VERITI[®] 96, Applied-Biosystems) a 50 °C por 50 minutos e a 85°C por cinco minutos para a interrupção da síntese. Em seguida as amostras foram mantidas em gelo para a adição de 1 μL de RNase H e nova incubação em termociclador a 37°C por 20 minutos. Ao final do processo, as amostras de cDNA foram armazenadas em freezer a -20°C.

Subsequente à transcrição reversa, realizou-se uma reação de PCR utilizando o primer para o gene da β -actina a fim de se verificar possíveis contaminações das amostras de cDNA com DNA genômico. A reação foi constituída de 2,5 μL de tampão 10X (10mM Tris e 50mM KCl), 1,0 μL de MgCl_2 , 1,0 μL (50mM) de dNTP's (2,5mM), 1,0 μL de primers *forward* e *reverse* (10 μM), 0,2 μL de Taq-

Polimerase (5U/ μ L) e 1,0 μ L de cDNA, com volume final de 25 μ L ajustado com água Milli-Q. Os parâmetros de ciclagem para as reações de amplificação foram: 94°C por 1 minuto; seguidos de 35 ciclos de 94°C por um minuto, 56°C um minuto e 72°C por um minuto, sendo a última extensão a 72°C por sete minutos. A reação de PCR foi analisada por eletroforese em gel de agarose 1,2% corado com brometo de etídio (0,1 mg/ml) utilizando o tampão de corrida SB 1X.

4.4.3.4 PCR quantitativo em tempo real (RT- qPCR)

As reações de PCR foram realizadas em termociclador 7900 Real Time System (Applied Biosystem), utilizando-se o Kit Platinum[®] SYBER[®] Green qPCR SuperMix UDG (Invitrogen – Life Technologies), de acordo com as instruções do fabricante, e como referência endógena o gene que codifica a β -actina. Para cada par de primers dos genes alvos e para a referência endógena (normalizador), foram realizadas reações de amplificação em triplicata, utilizando cDNA (bulk) com diluição seriada (1, 1:5, 1:25, 1:125, 1:625, 1:3125) e variando – se a concentração dos primers (0,25; 0,50; 0,75), a fim de se determinar e ajustar a eficiência de amplificação pelos pares de primers (curva de eficiência). A inclinação da curva obtida a partir da interpolação dos valores de Ct em função do logaritmo de cada diluição foi utilizada no cálculo da eficiência de amplificação dos primers, a partir da fórmula $E = [10^{-1/\text{slope}}] - 1$, no qual se espera que sejam similares e próximos de 100% (valor 1) para ambos os genes. A concentração de primer que promover a melhor eficiência foi selecionada para ser utilizada nos ensaios de quantificação relativa.

As reações foram feitas em triplicatas e compostas de 4,25 μ L de água MilliQ, 6,25 μ L de SYBER[®] Green qPCR SuperMix contendo ROX, 0,5 μ L de cada primer e 1 μ L de cDNA (~1,5 μ g) em bulk. Os parâmetros de ciclagem para as reações de amplificação foram: 50°C por 2min; 95°C por 10min; seguidos de 40 ciclos de 95°C por 15 seg., 60°C por 1min e 72°C por 30 seg. Após o término da quantificação relativa, realizou-se uma curva de dissociação com o intuito de averiguar se, ao longo da reação, não ocorreram amplificações inespecíficas, contaminações e formações de dímeros. Para tanto, a reação foi incubada a 95°C, por 15 seg., 60°C por 15 seg. e 95°C, por mais 15 seg.

Os valores de quantificação relativa e análise estatística foram determinados utilizando o Software Relative Expression Software Tool (REST)

versão 2.0.7 (PFAFFL *et al.*, 2002), sendo o gene β -actina utilizado como normalizador e as amostras falso inoculadas (*mock*) como calibradoras, em cada tempo e tratamento correspondente.

5 RESULTADOS E DISCUSSÃO

IDENTIFICAÇÃO DE GmMYB EM SOJA E FILOGENIA DA FAMÍLIA MYB

5.1 IDENTIFICAÇÃO DE MYBs R2R3 EM SOJA

Um total de 1709 sequências anotadas como MYB de soja foram obtidas a partir de três bancos públicos consultados: 305 sequências (PlantTFDB), 448 (SoyDB), 637 (Phytozome – palavras-chaves) e 319 (Phytozome – busca através da ferramenta Blastp do perfil HMM de MYB). Após a retirada das sequências redundantes entre os 3 bancos (comuns) obteve-se uma lista com 472 sequências potencialmente MYB-R2R3 únicas.

A análise detalhada do domínio R2R3 das sequências obtidas, realizada nos programas Pfam, SMART, MEME, em conjunto com a inspeção visual, e pelo cruzamento dos resultados obtidos, resultou em uma listagem com 264 sequências nas quais se identificou o domínio MYB R2R3 (Anexo 1). Um número significativo, uma vez que, entre os 791 MYBs preditos em *G. max* (WANG et al., 2010), esta classe representa 33% desta família de fatores de transcrição. A família MYB é ainda a maior entre os 5671 fatores de transcrição preditos na soja, representando 14% do total. A exclusão de 198 sequências se deu pela ausência do domínio característico ou devido a um *e*-value acima do ponto de corte adotado ($1e-10$). Também foram excluídas 10 sequências com três repetições (3-R) e quatro repetições (4-R).

O número de GmMYBs R2R3 identificados se apresentou superior a aquele encontrado por DU et al. (2012) (244 GmMYBs R2R3), possivelmente devido ao maior rigor na nota de corte adotado pelo autor para a identificação do domínio ($1e-20$), e inferior ao estipulado por FELLER et al., (2011) (288 GmMYBs R2R3), com base em buscas de sequências proteicas homologas ao perfil HMM de MYBs R2R3 (HMMER) em bancos de dados de sequências (Protein DB, Swiss Prot, Uniprot). O número de sequências encontradas também foi inferior às anotados como MYB-R2R3 nos bancos de dados consultados: SoyDB - (448 sequências); PlantTFDB - (305 sequências) e Phytozome - (305 sequências). Ambas as observações indicam que muitas sequências com domínios incompletos ou similares a MYBs estão presentes em diferentes bancos de sequências proteicas, o que

reforça a importância da realização de análises manuais e mais detalhadas que confirmem a identidade das mesmas e melhorem as informações dos bancos.

Apesar da redução do número de integrantes destes fatores de transcrição em soja, esse número ainda é maior que o encontrado em outras plantas como *Zea mays* com 188 e *Populus trichocarpa* com 198 (FELLER et al., 2011), e é mais que o dobro que observado em *Sorghum bicolor* com 104, *Minulus guttatus* com 119 sequências, *Vitis vinifera* com 131 (FELLER et al., 2011), *A. thaliana* com 126 (DUBOS et al., 2010) e em *Oryza sativa* com 109 MYBs (JIANG; PETERSON, 2004; YANHUI et al., 2006), tornando *GmMYB* uma das maiores famílias R2R3-MYB entre as plantas (Figura 5).

Entre as 198 sequências excluídas, 143 sequências não possuíam o domínio R2R3, 24 sequências possuíam apenas a segunda repetição (R2), quatro sequências possuíam apenas a terceira repetição (R3), seis sequências com R2 incompleto, 11 sequências com a terceira repetição R3 incompletas, cinco sequências cujo último aminoácido triptofano (W) foi substituído por outro aminoácido hidrofóbico (F ou C) e cinco sequências por apresentarem espaçamento entre os resíduos de triptofano anormais (maiores que aqueles encontrados caracteristicamente no domínio MYB).

Examinando a expressão das 198 sequências sem domínio completo R2R3 nos bancos de expressão (Anexo 4), pode-se observar que 105 (53,03%) não apresentaram qualquer expressão em nenhum das condições aos quais foram submetidas e analisadas; 51 (25,75%) apresentaram apenas uma expressão mínima em diferentes órgãos (Soybase) e 42 (21%) sequências apresentaram expressão diferencial durante a infecção de *G. max* por *P. pachyrhizi*, indicando que alguns deles (MYB-related) possivelmente apresentam algum papel nesse tipo de estresse, sendo potenciais alvos de estudos futuros mais detalhados desta classe pouco estudada até o momento.

Das cinco sequências com variação no espaçamento entre os resíduos de triptofano, apenas o Glyma18g39760.2, que possui o padrão mais próximo encontrado comumente no domínio R2R3 (21 resíduos de aminoácidos entre os "W", espaçamento normal varia entre 18 a 19) apresentou expressão. Não se observou expressão em nenhum dos bancos de expressão consultados nas cinco demais sequências cuja variação do padrão do domínio foi maior. Este fato constitui um indicativo de que eles possam ser apenas erros de sequenciamento ou

sequências contendo mutações mantidas no genoma do organismo (pseudogenes), contudo não funcionais. Diferença encontrada entre os genótipos estudados, assim como dos parâmetros experimentais adotados não estão descartados como possíveis explicações para a falta de expressão dos mesmos.

Cinco sequências apresentaram uma substituição do último triptofano (W) da repetição R3 por outro aminoácido hidrofóbico. Em quatro (Glyma02g02310.1, Glyma08g40950.1, Glyma18g16040.1, Glyma01g05190.1) ocorreu a substituição de W por F (fenilalanina), um aminoácido com estrutura semelhante (ambos com grupo hidrocarboneto aromático), e todas apresentaram expressão. Apenas o Glyma15g14190.1, cuja a substituição de W se deu por um resíduo de cisteína (C), um aminoácido diferente estruturalmente (hidrocarboneto alifático) não apresentou nenhuma expressão, indicando uma possível importância dessa região na conformação estrutural original da proteína.

A lista contendo todas as 208 sequências excluídas, bem como detalhes estruturais e perfil de expressão (com base nos bancos de dados analisados) pode ser consultada nos anexos (Anexo 4).

5.2 CARACTERIZAÇÃO FILOGENÉTICA E DISTRIBUIÇÃO EM SUBGRUPOS

A árvore filogenética construída a partir do método *neighbor-joining* permitiu o agrupamento de 263 das 264 sequências MYB-R2R3 identificadas no presente trabalho, juntamente com os *AtMYBs* guias utilizados em estudos prévios, em 43 subgrupos (Figura 8).

A distribuição dos MYBs nos 43 subgrupos não foi uniforme, e metade das sequências MYB ficaram agrupadas em apenas 11 subgrupos, com número de sequências variando entre 24 (subgrupo 14) a 1 por subgrupo (Subgrupo 5). Em ordem decrescente de tamanho, o subgrupo 14 agrupou 24 sequências, seguido do subgrupo 20 com 19 sequências, C11 com 18, subgrupo 2 com 17, subgrupos 9 e 21 com 16, subgrupos 13, 18, 22 e 1 com 15, subgrupos 4, 11, 25 e C34 com 13, subgrupo 6 com 12, subgrupo 7 com 11, subgrupos 10 e 20 com nove sequências, 2?, C16 e C40 com oito, subgrupos 16 e 23 com 7, subgrupos 12, 24 e C4 com 6 sequências cada, subgrupos C25, C35, C7, C32, C13 e C8 com quatro, subgrupos C11, C5, C3, ?1, C8, C36 e 15 com três sequências cada e subgrupo 5

com uma sequência agrupada. A distribuição dos subgrupos pode ser observada na Figura 8.

Dos 43 subgrupos formados, 25 subgrupos foram similares aos reportados por STRACKE et al. (2001) e DUBOS et al. (2010), bem como 14 novos subgrupos (C3, C4, C7, C5, C8, C11, C25, C35, C40, C36, C34, C35, C13, C15) reportados por DU et al. (2012), dos quais dois (C5 e C8) foram formados apenas por MYB de soja (subgrupos exclusivos). No presente trabalho foram identificados mais três novos subgrupos (?1, ?2 e ?3), não observados anteriormente, também formados apenas por GmMYBs, que podem contribuir, juntamente de C5 e C18, em futuras análises para a identificação e classificação de outras sequências pertencentes a estes subgrupos, de acordo com seus motivos, em *G.max*, *A. thaliana* e em outras plantas. O número maior de membros da classe R2R3 utilizados no presente trabalho explica a formação destes novos subgrupos, visto que os mesmos foram formados por sequências descartadas por DU et al. (2012) em suas análises. Apenas duas sequências, um AtMYB (AtMYB 82) e um GmMYB (Glyma13g09010) não foram agrupados em nenhum grupo e/ou apresentaram baixo *bootstrap* (< 500) para formarem um novo. Contudo, esse número de AtMYB não agrupados é menor que o observado nos trabalhos de STRACKE et al., (2000) e DUBOS et al. (2010), onde 35 AtMYB não se agruparam, incluindo AtMYB 82, Enquanto no presente trabalho estes 32 AtMYB remanescentes formaram novos subgrupos (identificados por C e ?). DU et al. (2012) também analisando a família MYB em soja, observou que Glyma13g09010, juntamente do Glyma07g14480 e Glyma08g27660, não pertenciam a nenhuma subfamília identificada. No presente trabalho, no entanto, Glyma07g14480 e Glyma08g27660 formaram um novo subgrupo (?3). Possivelmente, a utilização de guias provenientes de um maior número de espécies distintas de plantas por DU et al. (2012), em relação ao presente trabalho (apenas *A. thaliana* e *G. max*) explique a formação de um maior número de subgrupos e exclusão de sequências em seu trabalho. Uma vez que o agrupamento tem como base a similaridade dos membros, uma maior diversidade entre as sequências guias garante uma maior separação das sequências (de acordo com seus guias).

Alguns dos subgrupos identificados por DU et al. (2012) em soja não foram formados. C6 (Subgrupo5) não agrupou as sequências de GmMYB191, GmMYB193 e GmMYB205, que formaram um novo subgrupo (?1), restando apenas o guia AtMYB123, assim como o observado por DUBOS et al. (2010). Os grupos C12 e C34 não foram formados, pois as suas sequências foram agrupadas juntamente das sequências de C11, formando um único grupo. O mesmo foi observado para C18 (sequências agrupadas em C17), C28 e C30 (sequências agrupadas em C29). Outras pequenas variações dos subgrupos podem ser consultadas no Anexo 5.

Ocorreu ainda a formação de subgrupos compostos apenas por AtMYB: subgrupo 5 (AtMYB123), Subgrupo 12 (AtMYB028, AtMYB029, AtMYB034, AtMYB051, AtMYB076) e Subgrupo 15 (AtMYB000, AtMYB023, AtMYB066).

Interessantemente, a formação de agrupamentos contendo genes MYB exclusivos de *A. thaliana*, oferece indícios de que *G. max* aparentemente não apresenta genes MYBs que exerçam controle sobre algumas funções, tais como subgrupo S5 que agrupa apenas genes MYB envolvidos na síntese de protoantocianina, o S12 a referente rota de biossíntese de glicosinolatos, rota sabidamente predominante da Família Brassicaceae, em espécies de Cruciferae (MATTHAUS; LUFTMANN 2000), e, finalmente, o S15, com os fatores de transcrição MYB envolvidos na diferenciação celular de tricomas e distribuição dos pelos radiculares, indicando que o controle sobre essas funções podem estar sendo realizados por outros fatores de transcrição.

A distribuição dos genes GmMYB em subgrupos com base na similaridade entre as sequências de *G. max* e *A. thaliana*, permitiu ainda que se atribuíssem possíveis funções para os *clusters*, uma vez que alguns deles agrupam AtMYBs guias com funções já caracterizadas. Entre estes subgrupos, destacam-se aqueles relacionadas com a biossíntese de lignina, que agrupa pelo menos 23 sequências de soja, bem como diferentes subgrupos da classe de defesa e de controle de vias metabólicas de interesse nesse trabalho. A distribuição dos subgrupos de acordo com as funções designadas pelas sequências guias pode ser observada na tabela 2.

Relacionando a distribuição dos MYBs de soja nos subgrupos e a suas funções putativas, conclui-se que as principais funções sobre as quais GmMYBs podem desempenhar controle em soja são: Defesa – 24,58% das

sequências GmMYB (Subgrupos 20, 2, 18, 22, 1, 11, 20); Desenvolvimento – 23,38% (subgrupos 14, 9, 21, 13, 18, 25, 19, 16, 15) e Metabolismo com 17,66% das sequências (subgrupos 21, 13, 4, 6, 7,12) (tabela 2).

Em *A. thaliana*, onde os papéis dos MYBs são bem definidos, observa-se uma distribuição diferente das funções conhecidas sob o controle destes fatores de transcrição, com um número maior de MYBs atuando em rotas metabólicas (30,95%), ao passo que em soja este tipo de controle é o menos observado. Em seguida se destaca genes MYB envolvidos no controle sobre o desenvolvimento (29,36%) e defesa (18,25%) (DUBOS et al., 2010).

O destaque da participação da família MYB no controle putativo sobre respostas de defesa em *G. max* ainda é reforçado pelos resultados observados na análise do perfil transcricional dos GmMYBs em bancos de expressão, onde se observou que, entre as 213 sequências que possuem o domínio R2R3 MYB e com expressão mínima, 69 (32,4%) e 81 (40,8%), foram induzidos durante ensaios de infecção de soja por apenas *P. pachyrhizi* e *P. pachyrhizi* e outros agentes de estresse biótico (*Aphis glycines*, *Heterodera glycines*, *Phytophthora sojae*, *Pratylenchus brachyurus*), respectivamente (Tabela 4).

Outro fato que valida a filogenia obtida da família GmMYB é a análise das estruturas conservadas na porção C-terminal, região a jusante (*downstream*) ao domínio MYB, onde são encontrados motivos responsáveis pela atividade da proteína na regulação transcricional (FELLER et al., 2011). Desta maneira, apesar de não existirem informações que liguem cada um destes motivos a funções exatas, devido à escassa quantidade de trabalhos investigando o papel destas estruturas através de análise de seu perfil de expressão ou funcional, o compartilhamento destas estruturas conservadas por integrantes de um mesmo *cluster*, agrupados com base em sua similaridade estrutural, fornece um meio de inferir funções putativas aos GmMYBs, de acordo com a função de seu guia (*A. thaliana*). Os resultados da análise da região C-terminal estão apresentadas na tabela 3.

Tabela 2 – Subgrupos de GmMYB identificados no genoma de soja e funções putativas, atribuídas a partir do agrupamento com FT de *Arabidopsis thaliana*:

Subgrupo	N° seq.	Função (papel desempenhado)	
S14	24	desenvolvimento	Raízes, órgãos laterais
S20	19	defesa	Resposta a estresse abiótico mediado por Seca, alta salinidade, mediados por ABA, mediados por GA, mediados por JA, desidratação e a patógenos e herbivoria
C11	18	?	?
S2	17	defesa	estresse abiótico
S9	16	diferenciação celular	Controle sobre etapas de diferenciação
S21	16	metabolismo e desenvolvimento	Espessamento da parede celular e regulação de meristemas axilares
S13	15	metabolismo e desenvolvimento	Deposição e extrusão de mucilagem, Síntese de lignina, fechamento estomático
S18	15	desenvolvimento e defesa	Desenvolvimento do estâmen e antera, resposta a estresse abiótico (ABA)
S22	15	Defesa	estresse abiótico (seca, luz e lesão)
S1	15	defesa	estresse abiótico (seca) / estresse biótico (AS - HR)
S4	13	metabolismo	via dos fenilpropanóides
S11	13	defesa	estresse abiótico (osmótico)
S25	13	desenvolvimento	embriogênese
C34	13	?	?
S6	12	metabolismo	via dos fenilpropanóides / biossíntese de antocianina
S7	11	metabolismo	via dos fenilpropanóides / biossíntese de flavonóides
S10	9	?	?
C15	9	?	?
S19	8	desenvolvimento	Desenvolvimento
C13	8	?	?
2?	8	?	?
C40	8	?	?
S16	7	desenvolvimento	Desenvolvimento
S23	7	?	?
S12	6	metabolismo	biossíntese de glicosinolato
S24	6	?	?
C4	6	?	?
C16	6	Metabolismo	Via dos fenilpropanóides / síntese de lignina
C25	4	?	?
C35	4	?	?
C7	4	?	?
C32	4	?	?
C13	4	?	?
C8	4	?	?
C3	3	?	?
3?	3	?	?
C5	3	?	?
?1	3	?	?
C36	3	?	?
S15	3	diferenciação celular	Diferenciação celular
S3	2	metabolismo	Via dos fenilpropanóides / síntese de lignina

A análise das sequências proteicas dos MYBs de cada subgrupo pelo software MEME permitiu a identificação, além do domínio R2R3, de todos os motivos observados nas árvores filogenéticas construídas por STRACKE et al. (2000) e KRANZ et al. (1998) em *AtMYB*, na região de estudo, a identificação de 13 motivos novos em dez subgrupos previamente descritos por STRACKE et al. (2000) e KRANZ et al. (1998) em *AtMYB* (subgrupos 3, 4, 14, 20, 21, 22, 24 e 25 com um motivo cada; subgrupo 2 e 10 com três cada), e 11 (C3, C25, C11, C5, C8, C36, C13 com um cada; C40 e C35 com dois cada) nos subgrupos novos, dos quais dois (C5 e C8) são formados apenas por *GmMYBs*, totalizando 25 motivos conservados identificados. Entre os onze motivos dos subgrupos novos (C), oito foram observados por DU et al. (2012): C11, C3, C25, C8, C40 e C13 com um motivo novo cada e C35 com dois novos (tabela: 3).

Apenas os subgrupos C5, C40 e C36 apresentaram os mesmos motivos identificados por DU et al. (2012) em soja. Assim como observado por ele, os subgrupos C4 e C15 não apresentaram nenhum motivo. A distribuição dos motivos em cada subgrupo pode ser observada na Tabela 3.

Diferenças nos parâmetros adotados para a busca dos motivos na ferramenta MEME, tais como o número máximo de motivos para a busca menor (16 em vez de 50), possivelmente expliquem a ausência dos demais motivos observados por DU et al., (2012) em soja. Todos os motivos identificados neste trabalho foram únicos para os seus subgrupos, assim como o observado por STRACKE et al. (2000) e KRANZ et al. (1998), não ocorrendo o compartilhamento destas estruturas entre os mesmos, fato que reforça esta forma de classificação. Com exceção dos subgrupos 2 - Glyma10g30860.1, Glyma06g45570, Glyma12g11330, Glyma12g11340, Glyma12g11390 e Glyma06g45540, sem o motivo [E/N]ddMdFwynvfi observado por KRANZ et al. (1998); Glyma03g41100.1, Glyma19g43740.1, Glyma10g30860.1, Glyma06g45570 e Glyma12g11330 sem o motivo IDxSFW - -MxFWFD observado por STRACKE et al. (2000), e o motivo IDeSFWX[E/D]xIstd observado por KRANZ et al.,(1998); e subgrupo 21 - *AtMYB89*, *AtMYB111* sem o motivo FxDFL - (STRACKE et al., 2000), todos os demais subgrupos apresentaram as estruturas proteicas conservadas em todos as sequências que integram o grupo, demonstrando que, de fato, os *GmMYBs* estão agrupados de acordo com suas semelhanças estruturais.

Tabela 3 – Sequências consenso dos motivos da região C-terminal de sequências completas de GmMYBs identificados através de análise na ferramenta MEME.

Grupos	Nº ID	Referências	Sequência consenso	E-value
S1	1	STRACKE et al. (2000) – YASSxxNI KRANZ et al. (1998) – YaSS ^T / _A Eni ^A / _s ^R / _k LI	YASS[TA][ED]NI[AS][RK]LL[EKQ][GN]WM[KR][NS][SPT]P[KS]	5.7e-167
	2	STRACKE et al. (2000) – SL[F/I]EKWLF[D/E]	Q[MVGP][PS][LF][ST][LF][LF]E[KN]WL[LF]D[DE][AQ][AGS][GC]Q[CQE][KV][EL]D	3.0E-108
S2	3	KRANZ et al. (1998) – [E/N]ddMdFwynvfi Exceto:	[DS][AHNY][AHNS][NGY][SD][SCK][NQ][FLY][DNS][DE][GED]M[DE]FW[YF][DN][IV][FLY] Glyma10g30860.1, Glyma06g45570, Glyma12g11330, Glyma12g11340, Glyma12g11390, Glyma06g45540	1.7e-052
	4	STRACKE et al. (2000) – IDxSFW - -MxFWFD KRANZ et al. (1998) – IDeSFWX[E/D]xlstd Exceto:	[EF][ET][F]I[LM][PAD]V[I][DN]ESFVSE[LVA][LT][LS][DS]D[EIN]S Glyma03g41100.1, Glyma19g43740.1, Glyma10g30860, Glyma06g45570, Glyma12g11330	4.8e-50
S3	5	STRACKE et al. (2000) – WFKHLESELGLEExDNQQQ KRANZ et al. (1998)	[SV][HDE][DE][AD]E[AV][KD]KW[FS][KHQ][DH][FL]E[NS]E[FL]G[LV][EV]	3.3e-018
	6	QEVDKPE/D]LLE[/M]PFDWFKHLESELGLEE[N/D]DN Novo	L[IL]E[IM]P[FW][DE][PLS]D[PY]D[F]I[W]K[S][FL][LI]D[DG]T[DS][ND]S[FL][QEG][QS][FPS][Q ET][SAN][NR][EA][VLN][QS]	7.0e-016
S4	7	STRACKE et al. (2000) – LNL[E/D]L KRANZ et al. (1998) – pdLNL[D/E]Lxi[G/S] Du et al. (2012) - EIV][ELS]E[RP][CH][PH][DQ]LNL[ED]L[SRT]I[SG][PL]P[S RW]Q[PQ]Q	ISERCPDLNL[ED]L[TS][SG][PL]PSQPQ	1.2e-083
	8	Novo	N[TS]HI[RK]RKL[LI][SN]RGIDP[AN]THRP[IL]	7.3e-199
S5	9	STRACKE et al. (2000) - DExWRLxxT	EDWLR[CD][FQ]T	3.8e+000
S6	10	STRACKE et al. (2000) – KPRPR[S/T]F	AE[AP][KT][ST]TV[KQR][KP][HI][EQ]VIKPRPRS[FLI][ST] [IK][EQ]T[AS]E[APD][KTR][STP][TIP][VAT][KQR][KPD][HIV][EQ]VI[KR]P[RIQ]PR[SNT][FIL][STV][KV][TKNS][SN][PSV][CW]	6.1e-039 1.3e-047
S7	11	STRACKE et al. (2000) – GRTxRSxMK KRANZ et al. (1998) – KRR[L/P]GRT[G/S]RSAMKPK	[AK][AIN][V][NTPY][PLV][PAQ][PA][KR][RK][KR][RL]GR[TG]SR[SW][AS][MP]	8.3e-020
S9	12	STRACKE et al. (2000) – AQWESARxxAEeRLxRES KRANZ et al. (1998) – HmaQWeSARleAEaRLxR[E/Q]SxL DU et al., (2012) - [KT]N[DE][AT]LLS[ST][DEL]G[QP]SK[DN]T]AANLSHMAQ WESARLEAEARL[VA]RESK[LI]	GQ[SA]KS[AS]A[NS][LT][SR]HMAQWESARLEAEARL[SV]RESKL	2.1e-272
	13	KRANZ et al. (1998) – MGIDPvTHkp	NTHLKKRLT[KR]MG[IL]DP	5.3e-150

Tabela 3 – Continuação

Grupos	Nº ID	Referências	Sequência consenso	E-value
S10	14	STRACKE et al. (2000) – DDxxSDSxWK	W[EGC][DK][FL][LMI]D[DE]E[TA][ST][DG]S[YS][WSY][KW]	2.8e-019
	15	Novo	[SP]LP[NA]LV[SA][AF]SP[EDK][EMR][SG][TKP][FE]	3.9e-23
	16	Novo	[RP][RTS][DT][HD]LN[HLV]L[MSA][NAS]L[SQP]QLL[AG][ALT][SAT]N[LF]	1.5e-24
	17	Novo	T[QDS]L[AN]K[LAI]QLL[QH]N[IML][LF]I[Q][VL]	2.4e-20
S11	18	STRACKE et al. (2000) – PRLDLLD DU et al. (2012) - L[RK]MGIDPVTH[ST]PRLDLLD[IL]SSIL[NRS][SA]S[LF]I[Y G][NG]	NTHI[RK]K[RK]L[RK]MGIDPVTH[ST]PRLDLLD[IL]SSIL	4.1e-253
S12	19	STRACKE et al. (2000) – [L/F]LN[K/R]VA	[KC]K[VK][KR][FK]KR[SG]SS[ST][AS]R[LFLN[KR]VA][AN][RK][AF][GS]S[IR][GI]NI	3.0e-019
S13	20	KRANZ et al. (1998) – GIDPxTHKPxSEV	NSCLKKKL[RM][QK][RKQ]GIDP[VN]THKPLSE[VA]EN[GD][EK][DE]	1.2e-255
S14	21	Novo	NT[KR]LKKKL[L]M[GKQ][RP]K[ES]	5.0e-171
S15		-	-	-
S16	22	STRACKE et al. (2000) – PxLxFSEW DU et al. (2012) - PKL[LF]F[SA]EWL	LPKL[LF]F[SA]EWLS[LSV][DS][HQ][LPV][HN]	3.6e-24
S18	23	STRACKE et al. (2000) – GLP[L/V]YP	NTR[IML]KR[RL]QR[AQ]GLPLY[P]D[E][VI]x[QL]	3.0e-172
S19	26	KRANZ et al. (1998) – yWs[V/M][E/D]DIW[P/S] DU et al., (2012) - NYWS[MI][E/D]DIWS[MS]	YWS[MIV][ED][D]I[W]SP[MS][MQ][LQ]L	5.0e-038
S20	27	KRANZ et al. (1998) – AkqLkcdvNSkqFkdtmrylWmPRL	RTR[VI]QK[QH]A[KR]QL[KN][CI][DE][VS]NSKQF[KR]	2.0e-258
	28	Novo	[DE][TA][VM][RK]Y[LFV]WMPRL[LV]E[RK]QA[SA][SA][SA][SA]	6.1e-170
S21	29	STRACKE et al. (2000) – FxDFL	[DEK][EK][NK]G[VP]P[F][F]DFL[GS]VGA	6.6e-061
		Exceto	AtMYB89, AtMYB111	
	30	Novo	HVIMAR[KR]I[QRE][QR]S[KS][SL][YR]	3.9e-092
		Exceto	AtMYB69	
S22	31	KRANZ et al. (1998) – GEFM[V]VQEMikaEVRSYM	F[SG][AG][EG]F[MLW][ADT][VM][MV][QR][ED][MV][RAK][KRI]EVRSYM[V][SEA]S[LM]	3.3e-132
	32	Novo	GRTDNA[IV]KNHWNSTLKR[KR][CR]S[GS]	1.3e-218
S23	33	DU et al. (2012) - [CDY][CDGSY][ED][PQR][FIMQS][VI]P[SHQ][KQT]CG[HG] GC[CS]	[QE][GS]P[LV][IL]Q[SA][SP]KPD[MSV][EGL][AFS][CG][KR][LF][LF][EDQ][GS][IL][CD][CG][ED][PQ][MS][VI]P[SH][QK]CGHGCC	1.3e-075
	34	Novo	[NS]FEL[AI][AS]IATD[IL][NS]N[IL]AW[IL][KR]SGLE[NS][SC]SV[KLR]V[MA]	1.8e-040
	35	Novo	[GET][SR][FHK]S[EH][SG]S[LV]LGPEFV[DE]	5.9e-020
S24	36	Novo	[QL][AT][ER][ALM][AQ][EK]L[AQ][KL][LF]Q[YC]L[QL]Q	2.9e-013
	37	Novo	NATKRRQ[NY]S[KR]RK[CN][KR][SKR]	7.7e-084
S25	38	DU et al. (2012) - [PKV][RG][PN][NS][IL]L[EQ][DN]YI[KR]S[IL][TNE][ILQ][NI]	YP[RG][PGN][SN][IL]L[QE][DNE]YI[KR]S[IL][TN][L][ND]	7.1e-055
C11	39	novo	NTHIKKKL[KL]K[ML]G[IL]DP[VL]TH[KE]PL[NS][KQN][QK][ET][AE][SQ]	3.8e-252
Grupos	Nº ID	Referências	Sequência consenso	E-value
C3	40	novo	[YF][YC][SY]DD[VK]FSSFLNSLIND[V][AG][FD]	1.4e-013
?2	41	novo	-	-

Tabela 3: Continuação

C25	42	novo	[PS][NS][LST][TF][SE][PS][SC][HPS]S[ISTV][GT]S[AS][KR][LF]LNK[LM]A[AT][GS][IL][PS][HS][KR][VVKQ][HY][GS][L][ED][ARS][IAV]K[ATY][VI][FIL]	2.8e-014
C35	43	novo	RT][TH][HLM][MR][RK]KKAQ[ED]KK[RH][PAG][EMS][AS][AP][ST][SP][SV][SF]S[NS][CV][CDQ]	2.1e-020
C4	-	-	[YF]S[ML]DDIW[KR][DE][DM]	6.0e-013
C5	45	DU et al. (2012) - KT[QL]VYLPKPIRVKA[ML][YS][LS][QC] [RI][TP][DR]	KT[QL]VYLPKPIRVKA[ML][YS][LS][QC][RI][TP][DR]	5.4e-016
?3	-	-	-	-
C34	-	-	-	-
C16	-	-	-	-
?1	-	-	-	-
C7	-	-	-	-
C8	46	novo	[GE][NF][RK]WSLIAGRLPGRTDNEVKNYWNSHIR[RK]KLI[KS][MKN]GIDPN[NS]H[KR][LP][HKN][HQ][ST][FI]P[RS][SPV][HLY][QAV]	1.1e-072
C32	-	-	-	-
C40	47	DU et al. (2012) – (motivo 36 e 43) GHRA[LW][AG][TD]HKKEAAWRL[S R]R[VL]ELQLESEK[TA][NC]R[RQ]REK[I M]EE[FI]EAK[IM]KAL[QR]EE[EQ][KL][AN]A[LM][GE][RK]IE[AG]EYREQL[ADV][AG]LRRDAE[NA]K[ED]QKLA	[LV][SIR]R[VL]ELQLESEK[AT][GNC]R[RQ]REK[MI]EE[IF]EAK[IM]KAL[RQ]EE[QE][KTL][AN]A[LM][DGE][RK]IE[AG]EYREQL[ADV][GA]LRRDAE[NSA]K[ED]QKLA[ED]QW[ADT][AS][KR]H[LI]R[LF]T[KR][FL]LEQ[LQSV][GDM][CGT]	1.6e-145
	48	novo	EVFKEKQRE[KTE]K[EGD][NISV][NS][CRK][ITR][IASV][DE]PI[NDS][DNE]SKY[ED][HR]ILESFA EKLVKER	1.2e-077
C36	49	DU et al. (2012) - [WF][EG][LV][GMR][SK]SPYE[TN]RI[LS]DWI[AS]E[LI][QS][NT][GD][YQ][GS][E D][AKN][EN]L[ES][QE]D[CH][NS]S[TN] [S][TC][CS][NE][NY][YN]	SPYE[TN]RI[LS]DWI[AS]E[LI]	7.7e-015
C13	50	novo	[IM]EVQSC[ST][IM]DE[EQ]G[DE][IM]ALECL[QR]R[QR][GD]LNEW[VM][EV]	4.6e-45
C15	-	-	-	-
C32	-	-	-	-

5.3 PERFIL TRANSCRICIONAL DE GMMYB R2R3 EM REPOSTA A DOENÇAS EM SOJA

Com o propósito de inferir os possíveis papéis dos genes GmMYBs nos mecanismos de defesa de *G. max* em resposta a doenças, em especial a infecção por *P. pachyrhizi*, as 264 sequências de MYB-R2R3 foram analisadas em três bancos de expressão, o que permitiu a obtenção 69 modelos gênicos GmMYB (32,39%) com expressão diferencial em pelo menos um tratamento em resposta a infecção por *P. pachyrhizi*, tanto em interações compatíveis como incompatíveis ($p < 0,05$; Fold-change $\geq 1,0$; ratio $\geq 0,5$). Os níveis de expressão variaram de 1,49 (fold-change) (Glyma13g04030.1) a 81,94 (fold-change) (Glyma16g02570.1) (Anexo 6).

Outro fato observado que reforça o potencial destes FTs em contribuir para a defesa de soja durante infecção com patógenos vem da observação da expressão diferencial dos GmMYBs em ensaios com outros três agentes de estresse, com o pulgão-da-soja, *Aphis glycines*, o nematóide *Heterodera glycines*, o oomiceto *Phytophthora sojae*, causador da podridão da raiz da soja (dados do banco Genevestigator) e com o nematóide *Pratylenchus brachyurus* (Genosoja) (Tabela 5). Dos 213 GmMYBs induzidos em *G. max*, 87 (40,85%) apresentaram expressão diferencial em pelo menos um dos outros três estresses bióticos avaliados. Entre eles, 45 foram induzidos em resposta a apenas um dos agentes de estresse, dos quais 23 em ensaios com *Phytophthora sojae*, 9 em *Aphis glycines* e 6 em resposta a *Heterodera glycines*. Outros 29 GmMYBs foram encontrados sendo induzidos comumente por dois agentes, dos quais 12 por *P. sojae* e *A. glycines*, nove por *P. sojae* e *H. glycines* e sete por *A. glycines* e *H. glycines*. Por fim, 16 GmMYBs apresentaram expressão em ensaios com todos os três agentes (Tabela 4). Correlacionando esses resultados com aqueles encontrados anteriormente para os ensaios de *G. max* com FAS, observou-se que 37 GmMYBs induzidos por *P. pachyrhizi* (Glymas identificados com “*” tabela 4) foram responsivos também a outros agentes de estresse biótico. Entre estes 37, 10 GmMYBs apresentaram expressão para todos os agentes, representando a maior parte dos genes com esse padrão de indução. Outros nove GmMYBs foram induzidos por *P. sojae*, seis por *P. sojae* e *H. glycines*, cinco por *P. sojae* e *A. glycines*, quatro por *A. glycines*, dois por *A. glycines* e *H. glycines* e um por *H. glycines*. Nota-se o destaque dos 10 GmMYBs com expressão em todos os três

agentes, onde apresentaram os maiores níveis de expressão e se mostraram presentes em um maior número de ensaios. O MYB84 silenciado por PANDEY et al. (2011), envolvido na síntese de fenilpropanóides, novamente apresentou destaque com indução pelos três agentes, estando presente em todos os ensaios de *P. sojae*, em dois dos três para *H. glycine* e, um dos dois para *A. glycine*. O nível de expressão também foi elevado, variando entre 2,01 a 77,0 (fold-change).

GmMYBs também foram observados sendo expressos em resposta a *Heterodera glycines*, em trabalho de KANDOTH et al. (2011), que analisava os mecanismos de resistência em soja, através de microarranjo e RT-qPCR, onde o modelo gênico Glyma18g49360, ortólogo de AtMYB30, apresentou expressão em raízes de NIL-R (resistentes a *H. glycines*) superior em três vezes quando comparado a NIL-S/PA3 (suscetível), após cinco dias de infecção.

Considerando todos os dados de transcriptoma da soja em resposta a estresses disponíveis nos três bancos de dados acessados para ensaios com soja, 51 sequências não apresentaram nenhum nível de expressão sob as diferentes condições às quais foram submetidas ou apresentaram níveis de expressão muito baixo, restrito a uma parte da planta. Tais sequências possivelmente representam genes com expressão muito baixa ou ativados apenas em situações ainda não testadas ou não disponíveis nos bancos de dados, ou ainda possíveis erros de sequenciamento e montagem.

Glyma20g09140.1						
Glyma20g11040.1	+	✓				
Glyma20g11110.1		✓				
Glyma20g12290.1				+		
Glyma20g20980.1	+					
Glyma20g21670.1						
Glyma20g21680.1						
Glyma20g22230.1*	+					
Glyma20g24600.1				+		
Glyma20g25110.1a	+	✓				+
Glyma20g29730.1	+	✓				
Glyma20g32500.1	+		+			
Glyma20g32510.1a	+	✓				+
Glyma20g34140.1	+	✓				
Glyma20g35180.1*a	+					+
Glyma20g36600.1	+			+		
Glyma20g36600.2		✓				
Gnl ug gma#s34273466		✓				
Gnl ug gma#s39312923						
Gnl ug gma#s53086590;put-169a-glycine_max-63505	+					
Glyma14g21490.1	+					
Put-169a-glycine_max-102940						
Put-169a-glycine_max-179128;gnl ug gma#s53090519	+					
Put-169a-glycine_max-81255;gnl ug gma#s48315927	+					

SSH (LINO, 2011): 12, 24 e 48 HAI (A); 72 e 96 HAI (B) e 192 HAI (C); SuperSAGE-LGE: biblioteca SuperSAGE criada a partir ensaios de folhas de soja PI561356 (resistente) infectadas com *P. pachyrhizi*, e controles compostos de folhas sem infecção. Ambos disponíveis em www.lge.ibi.unicamp.br/soja/.

Genevestigator: Ensaio formado por plantas de soja (V3) PI462312 (Rpp3) em três tratamentos: com cepa virulenta de *P. pachyrhizi* (TW80-2), com cepas avirulentas (HW94-1) e controles (mock), As análises foram realizadas em 6 pontos: 12hpi (a), 24hpi (b), 72hpi (c), 144hpi (d), 216hpi (e), 288hpi (f). Dados do experimento disponíveis em: http://www.plexdb.org/modules/PD_browse/experiment_browser.php?experiment=GM36

Todos os demais dados foram obtidos através de consulta ao Genevestigator (<https://www.genevestigator.com/gv/plant.jsp>), utilizando os seguintes parâmetros para a análise dos resultados: $p < 0,05$; $\text{ratio} \geq 0,5$; $\text{fold-change} \geq 1,0$. Gene IDs acompanhados de “*” representam os 69 modelos com indução na interação soja-FAS; “a” – *P. sojae*; “b” – *A. glycines*; “c” – *H. glycines*; “d” – *P. sojae* + *A. glycines*; “e” – *P. sojae* + *H. glycines*; “f” – *H. glycines* + *A. glycines*; “g” – *P. sojae* + *A. glycines* + *H. glycines*;

Aphis glycines x soja (V3): Análise da resposta transcricional de plantas de soja à infestação de pulgões de soja resistentes e suscetíveis. Experimento com três fatores: variedade de soja em V3 (susceptível SD01-76R, resistente LD05-16060), o tratamento de pulgões (controle, pulgões), e duração de infestação (1 dia e 7 dias). Dados disponíveis em: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE35427>. (a): 1dpi, (b): 7dpi (estudo 2);

Heterodera glycines: Análise de perfis de transcrição de raízes de soja (Williams 82), via microarranjo, em três pontos de tempo de infecção (2,5,10 dias após inóculo - dpi) com nematódeos císticos de soja. Dados disponíveis em: <http://www.ebi.ac.uk/arrayexpress/browse.html?keywords=%09+E-MEXP-808>. (a): estudo 1, (b): estudo 2.

P. pachyrhizi x soja 5601T: Análise de expressão gênica por microarranjo em plantas de soja infectadas com o fungo *Phakopsora pachyrhizi* (72hpi) na fase de crescimento V2. Dados disponíveis em: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE7108>.

Phytophthora sojae: Identificação de genes diferencialmente expressos durante infecção pelo patógeno *P. sojae* em 8 cultivares de soja que diferem na resistência quantitativa (Athow, Conrad, General, Ox20-8, PI291327, Sloan, V71-370, Williams), através de microarranjo. As análises foram realizadas em tempos: 3dpi (a) e 5dpi (b). Dados disponíveis em: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE7124>; Padrões de expressão em diferentes tempos, durante a infecção por *P. sojae* (isolado PT2004C2.S1), de cultivares de soja que diferem na resistência quantitativa (resistentes a cepa – CONRAD e V71-370, suscetíveis – SLOAN e VP-RIL9): estudos (3) (24hpi), (4) (48hpi), (5) (72hpi), (6) (120hpi), disponível em: http://www.plexdb.org/modules/PD_browse/experiment_browser.php?experiment=GM3 e estudos (7)(24hpi), (8) (48hpi), (9) (72hpi), disponível em: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE11611>. Estudo 10 – Amostras de hipocótilo de soja Harosoy (susceptível) inoculados com a cepa P6497 (raça 2) de *P. sojae*, foram analisados através de microarranjo em 12 (a), 24 (b), 48 (c) horas após o inóculo (hpi). Todos os resultados foram obtidos a partir de ensaios com microarranjo.

Pratylenchus brachyurus: biblioteca (RNA-Seq) criada a partir de dois genótipos, TMG 115 RR e BRSGO Chapadões, suscetível e resistente, respectivamente, à infecção pelo nematóide *Pratylenchus brachyurus*, submetidos à inoculação do patógeno, nos períodos de tratamento, 24, 48, 96 e 192 horas após a inoculação

Correlacionando as sequências expressas em resposta a algum dos patógenos em soja a sua classificação na árvore filogenética, observou-se que 09 subgrupos (S25, ?3, C8, S6, S15, C36, C32, S24 e C3) não apresentaram qualquer sequência expressa. Para alguns destes subgrupos este resultado era esperado, devido o controle predito sobre funções não relacionadas a respostas de defesa (S25, embriogênese; S15, diferenciação celular de tricomas e raízes; C32, desenvolvimento de estâmen e antera; C3, metabolismo de mucilagem).

Entre os subgrupos com sequências expressas em resposta a patógenos, observou-se que o controle sobre a síntese de fenilpropanóides (S20, subgrupo do MYB84) (S7, flavonóides), síntese de lignina (S13), defesa (S1, morte celular), ABA (S11 e S22), espessamento da parede celular (S21), assim como um subgrupo de função desconhecida (C5), foram as respostas mais comuns observadas em soja.

A análise da região promotora (1Kb a montante do sítio de iniciação da tradução) dos 264 genes GmMYBs preditos teve como objetivo a busca de *cis*-motivos conservados. Foram avaliados os genes previamente identificados como induzidos por patógenos (*P. pachyrhizi*, *Aphis glycines*, *Heterodera glycines*, *Pratilenchus brachyurus* e *Phytophthora sojae*) (90 GmMYB-R2R3) a partir da análise de Bancos de Dados, bem como aqueles não induzidos, visando a identificação de *cis*-elementos conservados naqueles responsivos a infecção. No total foram identificados 30 *cis*-motivos, nenhum previamente reportado ou com descrição no banco de dados do PLACE, tornando-os potenciais novos *cis* elementos envolvidos na indução de genes em resposta a estresse bióticos em soja. Nenhum dos *cis*-motivos na região promotora de MYBs de milho identificados (KATIYAR et al., 2013) foram observados em soja (Tabela 5).

Tabela 5 – *cis*-motivos conservados presentes a montante a região promotora de genes de MYB de soja e Arabidopsis.

Sequência consenso dos motivos (IUPAC)	Log Likelihood Ratio	E-value	Conservação do Motivo (%)	Descrição PLACE	Nº do motivo	Logos do Motivo
Motivos de GmMYB induzidos por FAS e patógenos						
GRGRGWGWGRRRA	528	1,30E-31	47,36	motivo novo	1	
YYTCTCYYYYY	537	1,50E-30	52,63	motivo novo	2	
CHCMCHCM MMM	437	2,40E-07	57,89	motivo novo	3	
ATATATATATAT	294	1,30E-06	29,82	motivo novo	4	
YTTTTYTTYTTT	457	6,90E-02	52,63	CCA1	5	
YYTCTYTCTCTC	508	8,90E-46	97,95	motivo novo	6	
GAGRGRRAGRGR	310	1,10E-19	53,06	motivo novo	7	
AARDAARAAAAA	428	9,90E-10	100,00	motivo novo	8	
YTTTTYTYTTCT	406	1,10E-04	93,88	motivo novo	9	
Motivos de GmMYB induzidos por patógenos						
CYCTYTCTCTCT	603	2,00E-65	67,44	motivo novo	11	
AGARRRARARAR	555	3,50E-44	51,16	motivo novo	12	
GWGGKGGTGMK	267	4,20E-18	30,23	motivo novo	13	
TATATATATATA	218	4,90E-05	23,26	motivo novo	14	
SSGDYACGTRSC	135	9,10E-05	13,95	motivo novo	15	
Motivos dos demais GmMYBs (não induzidos por estresse biótico)						
TCTCTCTCTCTC	344	4,30E-32	39,53	motivo novo	16	
RRRARARAGR	499	3,10E-19	48,84	motivo novo	17	
YCCYCYCCMMHC	208	1,50E-05	32,56	motivo novo	18	
YTTTTYTTYTY	447	4,20E-03	69,77	motivo novo	19	
RGDGAGRG	234	9,00E-03	39,53	motivo novo	20	
CTCTCTYTCTC	359	6,60E-37	36,95	motivo novo	21	
GARRRRRRGARR	532	2,20E-36	60,87	motivo novo	22	
ATATATATATAT	412	1,20E-22	39,13	motivo novo	23	
GKTRGWKRGRG	248	2,30E-06	39,13	motivo novo	24	
SCTAGCTAGCW	208	1,10E-02	32,61	motivo novo	25	
YYYYTYTCTYTY	532	3,90E-28	55,17	motivo novo	26	
KGWGKTKKTD	461	1,70E-13	48,28	motivo novo	27	
CCACYCWHBBC	447	1,60E-10	50,00	motivo novo	28	
GGGCAGKWSAG	333	3,00E-03	41,38	motivo novo	29	
TCTCTCTCTCTC	108	6,70E-02	12,07	motivo novo	30	

*AT denota *A. thaliana* e GM denota *G. max*. FAS - *P. pachyrhizi*. Patógenos - *Aphis glycines*, *Heterodera glycines* e *Phytophthora sojae*

Nenhum *cis*-motivo foi encontrado em 100% das sequências e apenas o elemento TATATATATATA foi observado nos GmMYBs induzidos por FAS e patógenos, GmMYBs induzidos por patógenos e GmMYBs não induzidos por estresse biótico, contudo sem função conhecida.

Comparando-se os *cis*-motivos de sequências responsivos a um mesmo estímulo, não se observou qualquer conjunto comum ou padrão de disposição dos motivos na região de estudo. A disposição dos *cis*-elementos foi

muito variada na maioria dos grupos (GmMYBs responsivos a patógenos e não induzidos por estresse biótico). Contudo, observou-se que alguns GmMYBs induzidos por FAS e ao menos por um dos demais patógenos (Tabela 4) apresentam conjunto de *cis*-motivos iguais aos guias (AtMYBs) e comuns entre integrantes do subgrupo aos quais pertencem, apesar da disposição e o número dos elementos não ser a mesma (S23 - GRGRGWGWGRRA, S13 - YTTTTYTTYTTT e C35 - CHCMCHCMMMM, GRGRGWGWGRRA, YTTTTYTTYTTT, YYTCTCYYYYY). Observou-se ainda que os conjuntos de *cis*-motivos não foram únicos, sendo compartilhada entre diferentes subgrupos, que não apresentam função comum. Os subgrupos S2, S7 e S20 apresentaram o mesmo conjunto de *cis*-motivos (AARDAARAAAAA; YTTYTYTYTTCT e YYTCTYTCTCTC), enquanto os subgrupos S1, S14 e S18 apresentaram além destes três elementos, o motivo GAGRGRRAGRGR, e o S22 o *cis*-motivo CAGCTGTATCCC .

Desta forma, não foi possível identificar nenhum conjunto específico ou forma de organização (posição e número de elementos) dos *cis*-motivos para a região promotora dos GmMYBs responsivos aos agentes de estresse bióticos foco deste estudo e nem para os demais GmMYBs (não responsivos a estresse biótico) característico de cada subgrupo.

Contudo, a constatação de que alguns subgrupos dos GmMYBs apresentam o mesmo conjunto de *cis*-motivos que seus guias AtMYB (S2, S7, S20; S1, S14, S18; S22; S13; S23; C35), fornece um indício de que os mesmos podem sofrer o mesmo controle de expressão que seus guias AtMYBs em *Arabidopsis*, além de exercerem a mesma função (putativa).

5.4 SELEÇÃO DE GMMYBS PARA VALIDAÇÃO EM ENSAIOS DE RT-QPCR

De modo geral, a comparação das 69 sequências GmMYB com expressão diferencial em ensaios após infecção com *P. pachyrhizi* identificadas revelou que as mesmas foram as que apresentam maior similaridade com a sequência guia de seu grupo AtMYB cujas as funções estão relacionadas com defesa e metabolismo de flavonóides e fenilpropanóides. Tal fato indica uma possível similaridade funcional dos MYBs de soja com seu guia e que, possivelmente, estão envolvidos na resposta a infecção por patógenos em soja, tornando-os potenciais alvos para estudos funcionais futuros. Essa observação

ainda é reforçada pelo fato de que algumas dessas sequências, compõe o grupo de GmMYBs observados por DU et al. (2012) (36% dos MYBs de soja), presentes em blocos de sintonia de alta confiança entre soja e *Arabidopsis*, e que representam MYBs ortólogos destes guias. Isto sugere que alguns MYBs, após a divergência de soja e *Arabidopsis*, podem ter mantido suas propriedades e funções.

Os GmMYBs mais similares dentro de um mesmo subgrupo, em alguns casos, ainda apresentaram padrões similares de expressão frente à FAS. Alguns desses são GmMYBs duplicados no genoma (identificados por DU et al., 2012), tais como Glyma04g33720 (ortólogo putativo do guia AtMYB31) e Glyma06g20800 (subgrupo S1), Glyma05g35050 e Glyma08g04670 (subgrupo S20), Glyma07g05960 e Glyma16g02570 (subgrupo C5), Glyma11g11450 e Glyma12g03600 (subgrupo S4). Tais genes apresentaram um padrão (tempo e nível) de expressão similar em um mesmo ensaio com *P. pachyrhizi* (tabela 6). Esses fatos validam a organização das sequências em subgrupos, de acordo com a similaridade proteica, bem como a escolha de alvos potenciais baseada na função compartilhada entre elas e com seus guias, para estudos *in vivo* neste trabalho.

Com base nas funções dos genes regulados pelos AtMYBs guias de cada subgrupo, cujos genes alvos e o papel já são estabelecidos o suficiente para inferir funções putativas aos GmMYBs seus análogos em soja foram selecionados para validação. Foram selecionados 12 fatores MYB guias (Tabela 6) envolvidos nas rotas relacionadas a resposta a doenças e infecção por patógenos, bem como subgrupos que atuam na potencial regulação sobre as mesmas rotas metabólicas: Subgrupo1 (defesa), com três guias (AtMYB 30, AtMYB96 e AtMYB31), Subgrupo3 (metabolismo), com dois guias (AtMYB58 e AtMYB63), Subgrupo7 (metabolismo) com três guias (AtMYB11, AtMYB12 e AtMYB111), Subgrupo13 (metabolismo) com um guia (AtMYB61), e três grupos novos (?) cujos guias foram AtMY005 (metabolismo), AtMY046 (metabolismo) e AtMYB085 (metabolismo). Fixando-se um *Bootstrap* de corte de ≥ 500 para a formação dos agrupamentos, 33 sequências reunidas em um mesmo *cluster*, juntamente dos 12 AtMYB guias, foram selecionados como possíveis alvos das análises funcionais, e 12 delas estão entre aquelas identificadas nos bancos de expressão com perfil diferencial em resposta a *P. pachyrhizi* (fold-change $\geq 1,0$).

Entre os 12 GmMYBs, 10 foram escolhidos (Tabela 6), seguindo-se os seguintes critérios: maior similaridade com o guia, ser ortólogo putativo do guia,

apresentar expressão em ensaios de soja com FAS. Apenas os subgrupos C11, C3 e S3 e suas sequências foram selecionados por apresentarem funções preditas de interesse (mesmo não seguindo os critérios previamente citados). Todos apresentam ainda o mesmo padrão de íntrons e éxons (DU et al., 2012), e o mesmo motivo característico na porção C-terminal que os demais integrantes do seu subgrupo, indicando que são estruturalmente similares. Também foram incluídos como alvos o MYB84 silenciado por PANDEY et al. (2011) e um identificado por LINO, (2011), ambos potencialmente envolvidos a resposta a FAS.

O subgrupo 1 foi composto por cinco AtMYBs (AtMYB30, AtMYB96, AtMYB94, AtMYB31 e AtMYB60) e dez GmMYBs (Glyma13g05370, Glyma19g02600, Glyma18g49360, Glyma09g37340, Glyma06g20800, Glyma04g33720, Glyma17g10820, Glyma05g01080, Glyma19g29750 e Glyma03g00890). Os AtMYBs contidos nesse grupo são descritos por estar envolvido em respostas de defesa, controlando respostas a estresse biótico ativando a morte celular programada (resposta de hipersensibilidade - AtMYB30 e AtMYB96) e por serem expressos após indução por quitina (AtMYB31), um dos principais constituintes de parede celular fúngica e também um elicitador de defesa (Pathogen-Associated Molecular Patterns - PAMP) (DUBOS et al., 2010). DANIEL et al. (1999), infectando plantas de *A. thaliana* resistentes (Col-0) e suscetíveis (Sf-2) com uma cepa avirulenta de *Xanthomonas campestris* (cepa 147) capaz de induzir a HR, observaram a expressão de AtMYB30 apenas nas plantas resistentes (interação incompatível), não sendo observadas na interação compatível e controle, uma hora após a infecção e, precedendo o início, sendo expresso especificamente, durante a HR. O mesmo resultado foi observado quando as mesmas plantas de *A. thaliana* (Col-0 e Sf-2) receberam inóculos de cepas do patógeno *Pseudomonas syringae* PV. *Tomato* (DC3000) com os genes de avirulência *avrRpm1* ou *avrB*, onde AtMYB30 foi expresso 1 hora após o inóculo e 9 a 7 horas, respectivamente, antes da HR. Uma evidência mais concreta deste controle foi observada pela mesma equipe (VAILLEAU et al., 2002) quando infectaram plantas transgênicas de *A. thaliana* (Ws-4) expressando AtMYB30 em *sense* (overexpressão) e *antisense* (supressão), novamente infectada com o patógeno *Pseudomonas syringae* PV. *Tomato* (DC3000) virulento e avirulento (DC3000/*avrRpm1*). Nas plantas com superexpressão de AtMYB30 houve uma acelerada HR em resposta a cepas avirulentas, enquanto em

plantas com AtMYB30 em *antisense* a morte celular foi totalmente suprimida em reposta as mesmas cepas.

RAFFAELE et al. (2006) demonstrou que a HR é modulada pela expressão de AtMYB30 em *A. thaliana* infectada por *Pseudomonas syringae* pv. *Tomato (pst)* avirulento(*avrRpm1*), em resposta ao ácido salicílico (SA). Em plantas transgênicas incapazes de acumular SA (com a enzima salicilato hidrogenase – *NahG*) e mutantes (*sid1* e *sid2*) e infectadas com *Pst*, a expressão de AtMYB30 foi reduzida, quando comparada as plantas selvagens, demonstrando que este FT é dependente da síntese de SA. Plantas superexpressando AtMYB30 exibiram altos níveis de HR, expressão de genes de defesa e fenótipo de resistência, e apenas as plantas duplo mutantes, superexpressando esse FT e sem capacidade de acumular SA, não apresentaram esses resultados, o que indica que HR e o fenótipo de resistência conferidos pela expressão de AtMYB30 (modulador) são dependentes da síntese deste hormônio.

O outro guia AtMYB96 , assim como AtMYB30, está envolvido controle de respostas de defesa à estresses bióticos através do controle sobre a síntese de SA. SEO; PARK, (2010) trabalhando com *A. thaliana* superexpressando AtMYB96, observou que o aumento de expressão de AtMYB96 resultava em regulação positiva das enzimas *PAL1* e *SID2* ligadas a síntese de SA. A infecção desta plantas com *P. syringae* aumentou a resistência ao patógeno e a expressão de *SID2* e genes *PR* (*PR1*, *PR2* e *PR5*) regulados por SA. O tratamento das plantas com um eficiente elicitador de defesa de plantas (flagelina Flg22) demonstrou que a ordem de expressão dos três genes (AtMYB96 → *SID2* → *PR1*) permite concluir que o FT possivelmente promove a síntese de SA, o qual induz os genes *PR*, tornando esse FT o controlador principal desta via. O uso de plantas superexpressando AtMYB96 e sem a enzima *SID2* (*sid2*) promoveu uma redução da expressão dos genes *PR*, bem como perda da resistência à *P. syringae*.

Tabela 6 – GmMYBs com expressão diferencial distribuídos nos subgrupos juntamente de guias AtMYB:

Subgrupo	MYB guia / Função (DUBOS et al., 2010)	Similaridade com guia	Glymas	Similaridade entre as sequências				Expressão ^A (Fold-Change)	
				(A)	(B)	(C)	(D)		
S1 Bootstrap (925)	AtMYB030	Defesa ¹	74,0%	Glyma09g37340.1 (A)	-	88.0%	73.0%	63.2%	-
			74,0%	Glyma18g49360.1 (B) *	88.6%	-	72.1%	64.7%	1.59
			65,0%	Glyma13g05370.1 (C) *	73.2%	71.9%	-	83.7%	-
			51,1%	Glyma19g02600.1 (D) *	64.2%	65.3%	84.7%	-	-
S1 Bootstrap (709)	AtMYB031 e AtMYB96	Defesa ^{1 e 2}	60,6%	Glyma04g33720.1 (A) *	-	93.1%	69.3%	66.8%	2.15
			62,1%	Glyma06g20800.1 (B) *	87.1%	-	74.9%	69.1%	2.67
			57,9%	Glyma05g01080.1 (C)	69.1%	69.9%	-	82.5%	-
			60,9%	Glyma17g10820.1 (D)	70.3%	68.1%	87.1%	-	-
S3 Bootstrap (983)	AtMYB058	Metabolismo ³	68,6%	Glyma08g17860.1 (A)	-	74.0%	-	-	-
			60,6%	Glyma15g41250.1 (B)	75.3%	-	-	-	-
	AtMYB063	Metabolismo ³	62,6%	Glyma08g17860.1 (A)	-	74.0%	-	-	-
			59,2%	Glyma15g41250.1 (B)	75.3%	-	-	-	-
S7 Bootstrap (925)	AtMYB11	Metabolismo ⁴	57,1%	Glyma02g01740.1 (A) *	-	67.7%	66.1%	-	3.19
			35,6%	Glyma03g37640.1 (B) *	60.7%	-	75.3%	-	-
	35,6%	Glyma19g40250.1 (C) *	61.8%	78.5%	-	-	1.62		
	AtMYB12	Metabolismo ⁴	53,4%	Glyma02g01740.1 (A) *	-	67.7%	66.1%	-	-
			35,0%	Glyma03g37640.1 (B)	60.7%	-	75.3%	-	-
	34,2%	Glyma19g40250.1 (C)	61.8%	78.5%	-	-	1.62		
AtMYB111	Metabolismo ⁴	40,1%	Glyma17g03480.1 (A) *	-	62.1%	69.2%	-	-	
		54,1%	Glyma09g04370.1 (B) *	71.7%	-	85.1%	-	2.24	
32,2%	Glyma15g15400.1 (C)	75.8%	80.7%	-	-	-			
S13 Bootstrap (944)	AtMYB61	Metabolismo ³	69,4%	Glyma20g22230.1 (A) *	-	87.4%	72.7%	73.5%	2.06
			65,8%	Glyma10g28250.1 (B) *	87.6%	-	67.5%	73.5%	1.56
			73,2%	Glyma03g38660.1 (C) *	71.0%	65.7%	-	86.9%	-
			70,2%	Glyma19g41250.1 (D) *	74.5%	74.4%	90.2%	-	-
C3 Bootstrap (1000)	AtMYB005	Metabolismo ⁵	53,0%	Glyma13g09980.1 (A)	-	78.9%	-	-	-
			43,8%	Glyma14g24500.1 (B)	72.2%	-	-	-	-
C16 Bootstrap (1000)	AtMYB046	Metabolismo ⁶	75,0%	Glyma15g03920.1 (A)	-	76.0%	73.6%	-	1.65
			57,1%	Glyma11g14200.1 (B)	67.4%	-	69.6%	-	-
			54,3%	Glyma12g06180.1 (C)	60.8%	64.9%	-	-	-
C11 Bootstrap (1000)	AtMYB085	Metabolismo ⁶	68,8%	Glyma01g42050.1 (A) *	-	80.4%	87.5%	74.5%	-
			70,3%	Glyma05g03780.1 (B) *	76.2%	-	90.2%	90.5%	-
			67,7%	Glyma11g03300.1 (C) *	80.8%	91.5%	-	85.4%	-
			63,2%	Glyma17g14290.1 (D) *	71.3%	91.5%	88.6%	-	-
PANDEY et al.(2011)	Glyma08g04670.1	Metabolismo ³	90,1%	Glyma05g35050.1	-	-	-	-	4,62
				Glyma08g04670.1	-	-	-	-	5,02

¹ - Resposta a estresse abiótico / resposta HR, mediada por AS; ² - Resposta a quitina; ³ - Regulação da via de fenilpropanóides (biossíntese de lignina); ⁴ - Regulação da via de fenilpropanóides (biossíntese de flavonóide); ⁵ - Regulação da via de fenilpropanóides (biossíntese de proantocianina); ⁶ - Espessamento da parede celular vegetal.

^{*} - gene MYB putativo ortólogo do guia (s), identificado por Du et al., (2012), determinado após análise da relação de sintonia com base no conteúdo de genes, ordem e orientação entre os genomas de soja e Arabidopsis. GmMYBs sombreado em cinza são genes duplicados. ^A - padrão de expressão de genes em um mesmo ensaio e tempo de inoculo (Dados do banco Genevestigator). As sequências selecionadas para os ensaios de validação estão marcadas em vermelho.

Assim, os FT de soja Glyma09g37340, Glyma13g05370, Glyma19g02600, Glyma18g49360 agrupados juntamente de AtMYB30, apesar de não possuírem funções conhecidas, compartilham todos os motivos, o padrão de éxons e íntrons encontrados caracteristicamente neste grupo (DU et al., 2012) e, possivelmente, podem desempenhar o mesmo controle sobre as respostas de resistência em soja. O Glyma18g49360 apresentou maior similaridade com o guia (74%) e, além de ser um ortólogo putativo do mesmo (DU et al., 2012), foi o único a apresentar expressão diferencial durante a infecção da soja por FAS.

LIBAULT et al. (2007) associando dados de um microarranjo a um RT-qPCR, utilizando primers para 2077 fatores de transcrição (FT) de *A. thaliana*, identificou 118 FT de 29 famílias, expressos em resposta ao tratamento com quitina pura. Entre os FT regulados positivamente, incluíam AP2/ERE, C2H2 zinc-finger, WRKYs e 10 MYBs, entre os quais o AtMYB31, induzido (nível de expressão: 2,7) apenas na planta tratada. Juntamente de AtMYB31 se agruparam os GmMYBs Glyma04g33720, Glyma06g20800 Glyma17g10820, Glyma05g01080, onde Glyma04g33720 e Glyma06g20800 apresentaram a maior similaridade com o guia (60% ambos) e também expressão em soja durante infecção por FAS. Contudo, Glyma04g33720 é um ortólogo do guia AtMYB31.

O subgrupo 3 foi composto por dois AtMYB (AtMYB58 e AtMYB63) e dois GmMYB (Glyma08g17860 e Glyma15g41250). ZHOU et al. (2009) realizou a superexpressão de AtMYB 58 e AtMYB63 em plantas de *Arabidopsis* e observou deposição ectópica de lignina nas células mesofilicas e epidérmicas de folhas, assim como em células corticais do pecíolo (lignificação de tecidos). Análises químicas também indicaram que o aumento de expressão destes FT elevou o conteúdo de lignina na folha em comparação com a planta selvagem em 142% (AtMYB63) a 212% (AtMYB58). O silenciamento através de RNAi de ambos os FT promoveu a redução na espessura da parede secundária em fibras interfascicular e fibras do xilema, bem como o conteúdo total de lignina em 38% (AtMYB58) e 26% (AtMYB63).

Análise do perfil de expressão entre plantas superexpressando ambos os fatores de transcrição e plantas selvagens demonstrou que ambos induzem a expressão de nove dos dez genes (exceto a enzima F5H1 – ferulato 5-hidroxilase) da via de síntese de lignina. Ambos os guias desse grupo desempenham em *A. thaliana* regulação positiva sobre nove enzimas (PAL1, fenilalanina amônia liase 1; C4H, cinamato 4-hidroxilase; 4CL1, 4-coumarato CoA

ligase 1, enzimas comuns de todas rotas dos fenilpropanóides; HCT, hidroxicinamoil CoA:shikimato/quinato hidroxicinnamoiltransferase; C3H1,p-coumarato 3-hidroxilase 1; CCoAOMT1, caffeoil CoA 3-O-metiltransferase 1; CCR1, cinamoil CoA redutase 1; F5H1, ferulato 5-hidroxilase 1; COMT, acido caffeico O-metiltransferase; CAD6, cinamil álcool dehidrogenase 6, enzimas da rota de síntese de lignina) da via dos fenilpropanóides (ZHOU et al., 2009), e fazem parte de um complexo regulatório que combina a interação de múltiplos FTs que controlam a ativação de genes de biossíntese de lignina. Ambos FT, AtMYB58 e AtMYB63 ativam diretamente a expressão, reconhecendo elementos AC encontrados na região promotora dos genes das enzimas (ZHONG; YE, 2009).

Finalmente, todos os MYBs deste subgrupo foram agrupados em um único clado onde Glyma08g17860 apresentou similaridade de 68% com AtMYB058, e 62% AtMYB63, enquanto Glyma15g41250 60% com AtMYB058 e 59% com AtMYB63. Contudo nenhum deles apresentou expressão diferencial em ensaios de soja com FAS nos bancos de expressão consultados.

O Subgrupo 13 é outro grupo envolvido no controle da síntese de lignina e foi formado por quatro AtMYBs (AtMYB86, AtMYB61, AtMYB50 e AtMYB55) e onze GmMYBs (Glyma04g36110, Glyma06g18830, Glyma05g02550, Glyma20g22230, Glyma10g28250, Glyma19g41250, Glyma03g38660, Glyma10g27940, Glyma19g41010 e Glyma03g38410). Apenas AtMYB61 apresenta função conhecida entre os AtMYBs deste grupo. NEWMAN et al. (2004) promovendo superexpressão deste FT em *A. thaliana* (Col-0), observaram um aumento na lignificação de pecíolos, xilema e fibras interfasciculares. De acordo com ZHONG; YE (2009), este FT é homólogo a MYB de *Pinus taeda*, PtMYB8, cuja superexpressão também promove a deposição ectópica de lignina durante a formação da madeira em *Pinus* e eucalipto.

Em trabalho de DU et al. (2012) AtMYB61 foi agrupado juntamente com PtMYB8, igualmente no subgrupo13. BOMAL et al., 2008 analisando por RT-qPCR de *Pinus taeda* superexpressando o gene de PtMYB8, observou a regulação positiva sobre seis das dez enzimas da via de síntese de lignina (PAL1, fenilalanina amônia liase 1; C4H, cinamato 4-hidroxilase; 4CL1, 4-coumarato CoA ligase 1; C3H1,p-coumarato 3-hidroxilase 1; CCR1, cinamoil CoA redutase 1; CAD6, cinamil álcool dehidrogenase 6).

O AtMYB61 foi agrupado juntamente de Glyma20g22230, Glyma10g28250, Glyma19g41250 e Glyma03g38660, dos quais, apenas Glyma20g22230 e Glyma10g28250 apresentaram expressão na interação de soja com *P. pachyrhizi* nos bancos de expressão. O Glyma20g22230 foi escolhido como alvo por ser ortólogo ao guia AtMYB61.

No subgrupo C16 foram agrupados dois AtMYB (AtMYB46 e AtMYB83) e seis GmMYBs (Glyma15g03920, Glyma12g06180, Glyma11g14200, Glyma19g05080, Glyma13g27310, Glyma12g36630). O guia AtMYB46 é um regulador positivo da síntese de lignina (fibras e vasos) e de deposição de celulose e xilanas em paredes secundárias em *A. thaliana*. O único a apresentar expressão em soja durante algum estresse biótico foi o Glyma15g03920, que também foi o MYB com maior similaridade com AtMYB46. Este fator de transcrição desempenha o papel de principal regulador da biossíntese dos três principais componentes de parede secundária, que incluem a celulose, xilana e lignina. AtMYB46 controla os ativadores diretos da transcrição dos genes de síntese de lignina, tais como os FT AtMYB58, AtMYB63 e AtMYB85, que regulam positivamente 9 das dez enzimas da via de síntese da lignina, ligando aos *cis*-elementos AC destes genes. (ZHONG; RICHARDSON; YE, 2007).

O subgrupo C11 reuniu seis AtMYBs (AtMYB85, AtMYB42, AtMYB20, AtMYB43, AtMYB99, AtMYB40) e doze GmMYBs (Glyma05g03780, Glyma17g14290, PUTT69aGlycinemax102940, Glyma01g42050, Glyma11g03300, Glyma11g33620, Glyma18g04580, Glyma14g39530, Glyma02g41180, Glyma05g06410, Glyma19g07830, Glyma16g06900). Apenas AtMYB85 possui função conhecida em *A. thaliana*, onde atua regulando positivamente enzimas da via de síntese de lignina em célula de fibras. Assim como AtMYB58 e AtMYB63, este FT faz parte de uma complexa rede de controle da síntese de lignina, e está sobre controle de AtMYB46, que por sua vez é regulado pelo FT NAC (SND1), o regulador principal desta rede, com capacidade de ativar as vias biossintéticas inteiras de celulose, xilana e lignina, conduzindo a formação de paredes secundárias (ZHONG et al., 2007). Os GmMYBs Glyma05g03780, Glyma17g14290, Glyma01g42050 e Glyma11g03300 foram agrupados juntos desse MYB, contudo o Glyma05g03780 foi escolhido como alvo por apresentar maior similaridade e ser ortólogo do guia.

No subgrupo 7 foram agrupados três AtMYB (AtMYB11, AtMYB12 e AtMYB111) e oito GmMYB (Glyma03g37640, Glyma19g40250, Glyma02g01740,

GmaS53090519, Glyma07g37140, Glyma17g03480, Glyma09g04370 e Glyma15g15400). Os três AtMYBs guias controlam quatro enzimas chave da via dos fenilpropanóides (Chalcona sintase - CHS, chalcona isomerase - CHI, flavona 3-hidroxilase – F3H e flavonol sintase - FLS) que levam a produção de flavonóides em *A. thaliana*, (STRACKE et al., 2007), substâncias com atividade antifúngica, tais como as fitoalexinas (TREUTTER , 2006). MEHRTENS et al. (2005) obtiveram plantas de *A. thaliana* (Col-0) superexpressando (*MYB12* –OX) e plantas com o mesmo gene não funcional (*myb12* –KO). A análise do perfil de expressão destas plantas revelou uma forte ativação e redução em paralelo, respectivamente, das enzimas CHS, CHI, F3H e FLS, em comparação com a planta selvagem. O mesmo padrão foi observado para dois produtos destas enzimas, kaempferols e quercetinas, quantificados nestas plantas por meio de HPLC. STRACKE et al., (2007) por sua vez, avaliou uma planta triplo mutante para os três FTs do subgrupo 7, através de microarranjo, e constatou que, em comparação com as plantas selvagens, houve redução da expressão das enzimas CHS, FLS1, CHI e F3H. O efeito da redução de atividade das enzimas também refletiu sobre a produção de flavonóides produzidos e identificados na planta selvagem através de HPLC.

AtMYB12 e *AtMYB11* formaram um grupo juntamente de Glyma02g01740, Glyma03g37640, Glyma19g40250, Glyma17g03480, sendo Glyma02g01740 o *GmMYB* de maior similaridade com ambos os guias (57% - *AtMYB12*; 53,4% - *AtMYB11*) e também o que exibiu maior expressão em *G. max* durante ensaios com FAS. Outro grupo formado dentro do subgrupo 7 foi *AtMYB111* e Glyma17g03480 (similaridade de 40%), Glyma09g04370 (54%) e Glyma15g15400 (32,2%). Novamente os MYBs de maior similaridade com seus guias foram os únicos expressos em soja durante a interação com *P. pachyrhizi* e, assim como o primeiro alvo deste subgrupo, é um ortólogo de seu guia.

O subgrupo C3 reuniu apenas um *AtMYB* (*AtMYB005*) e dois *GmMYBs* (Glyma14g24500 e Glyma13g09980). *AtMYB005* é conhecido por atuar em *A. thaliana* controlando a síntese de mucilagem e proantocianidina, composto com atividade antifúngica encontrada na semente e nas bordas das folhas jovens, assim como o desenvolvimento de tricomas em *A. thaliana*, importante mecanismos de defesa contra herbivoria. Em cevada, experimentos *in vitro* de infecção de grão de plantas mutantes (sem síntese de proantocianidina) e utilizando os extratos destes grãos frente aos patógenos *Fusarium poae*, *Fusarium culmorum*, e *Fusarium*

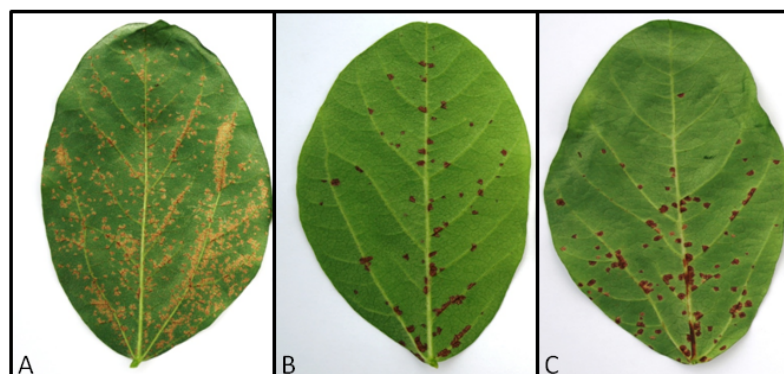
graminearum, demonstraram perda da capacidade de inibir o crescimento de todas as cepas nas plantas mutantes, em teste de halo de inibição. Plantas selvagens cuja síntese ocorreu normalmente, a inibição foi observada normalmente (SKADHAUGE; THOMSEN; WETTSTEIN, 1997). Plantas mutantes para esse FT também apresentaram redução de 40% do número de tricomas, enquanto plantas superexpressando o FT tiveram um aumento no número dessas estruturas (GONZALEZ; MENDENHALL; LLOYD, 2009). Nenhum dos GmMYBs nesse grupo apresentou expressão em soja durante a infecção por *P. pachyrhizi*.

5.5 VALIDAÇÃO *in vivo* DO PERFIL TRANSCRICIONAL DE GMMYB ENVOLVIDOS NA RESPOSTA A FAS

Com base nas análises *in silico*, foram selecionados inicialmente onze GmMYBs (a partir dos 12 AtMYBs guias previamente selecionados no item 5.4), potencialmente envolvidos na resposta a doenças, para serem validados via RT-qPCR frente à infecção por *P. pachyrhizi*. Entretanto apenas sete GmMYBs dos onze foram avaliados por apresentarem *primers* funcionais. Amostras de RNAm extraídos de folhas de WH82 (Williams 82, genótipo suscetível), PI230970 e Shiranui (genótipos resistentes) em 12, 24, 48, 72 e 96 hai (horas após inóculo) foram utilizadas para a validação dos genes após análise fenotípica da resposta a infecção a FAS. Os genótipos PI230970 e Shiranui apresentaram lesões do tipo RB indicando resistência e, como o esperado para o genótipo suscetível WH 82, observaram-se lesões do tipo TAN.

Os fenótipos observados 21 dias após a inoculação podem ser observados na Figura 9.

Figura 8 – Perfil fenotípico de genótipos de soja após infecção por *P. pachyrhizi*. A. WH82 com intensa esporulação característica no fenótipo de plantas suscetíveis (lesão do tipo TAN); B. PI230970 com lesões do tipo Rb, indicando resistência à população de isolados utilizados; C. Genótipo Shiranui com lesões do tipo Rb, indicando resistência do genótipo à infecção por *P. pachyrhizi*.



A análise da expressão dos GmMYBs selecionados resultou em cinco modelos gênicos com perfil diferencial após a infecção com *P. pachyrhizi*, em pelo menos um genótipo/tempo após a inoculação.

No genótipo suscetível WH 82, 4 dos 7 GmMYB avaliados foram diferencialmente expressos em pelo menos um tempo após a infecção: Glyma04g33720, Glyma12g113910, Glyma08g04670 e Glyma20g22230 (Figura 10)

No genótipo resistente PI230970 (Figura 11), os genes que apresentaram indução em pelo menos um tempo foram Glyma20g22230 e Glyma08g17860, ortólogos de AtMYB61 (NEWMAN et al., 2004) e AtMYB58 (ZHONG; YE, 2009), respectivamente, potencialmente envolvidos na regulação da biossíntese de lignina, e o Glyma12g11390, identificado na biblioteca em resposta a ferrugem, utilizando o mesmo genótipo, em todos os *bulks* (que abrangem o tempo de 12 a 192hai) (LINO, 2011) (Figura 11).

Em Shiranui, genótipo contendo o gene Rpp5, destacaram-se os fatores MYB Glyma08g04670, envolvido potencialmente na regulação da síntese de fenilpropanóides (PANDEY et al., 2011), e o Glyma04g33720, (ortólogo de AtMYB31) (LIBAULT et al., 2007), modulando respostas a quitina (Figura 12).

O Glyma04g33720 apresentou expressão diferencial no genótipo suscetível WH82 e no genótipo resistente Shiranui (Rpp5), sendo regulado positivamente no genótipo suscetível WH82 nos tempos tardios de 72hai (1,33) e 96hai (2,26), enquanto no genótipo resistente em 48 hai (1,67) e 96 hai (1,35). Em

PI230970, o gene foi negativamente regulado em 12 hai (-0,6) e 48 hai (-0,7) (Figura 11).

O gene Glyma08g17860, com similaridade ao AtMYB58, que exerce controle sobre a via dos fenilpropanóides e síntese de lignina, foi diferencialmente expresso apenas nos genótipos resistentes, sendo regulado positivamente em 24 hai (5,25) e 72 hai (2,48) na PI230970, e negativamente em 96 hai (-0,3) em Shiranui. Pelo fato de que AtMYB63 exercer o mesmo controle sobre enzimas da rota de síntese de lignina observados em AtMYB58, existe a possibilidade de que em Shiranui a indução desta rota estaria sendo realizada pelo ortólogo de AtMYB63, Glyma08g04670, uma vez que este foi induzido neste genótipo tanto em 12 como 72 hai (dados abaixo). É possível que haja a sobreposição no controle de vários pontos da rota de biossíntese de fenilpropanóides, que inclui a biossíntese de lignina, devido à redundância de função entre MYBs de um mesmo subgrupo (S3) e diferentes subgrupos (C16 e C11).

O Glyma12g11390, que apresentou expressão diferencial em resposta a infecção de soja por FAS no genótipo PI561356 demonstrado pelos valores elevados de FPKM (26,49) (LINO, 2011), também se mostrou significativamente expresso no presente trabalho, tanto no genótipo suscetível como no genótipo resistente PI230970. Em WH82 foi regulado positivamente em 12 hai (3,20), enquanto em PI230970 ele foi regulado positivamente mais tardiamente, em 24 hai (2,0), contudo a expressão se manteve até 72 hai (2,0). No genótipo Shiranui ele não apresentou expressão significativa.

O gene que codifica o fator de transcrição MYB84, silenciado no trabalho de PANDEY et al., (2011), Glyma08g04670, e que parece controlar a biossíntese de fenilpropanóides através de possível regulação sobre as enzimas fenilalanina amônia-liase (*GmPAL1*), 4-coumarato-CoA-ligase (*Gm4-CL*), Cinnamato-4-hidroxilase (*GmC4H*) e Chalcona sintase (*GmCHS*) da via, foi diferencialmente expresso durante a infecção de soja por FAS, tanto no genótipo WH82 como em Shiranui. Ele foi regulado positivamente em 12 hai (7,34) e 96 hai (2,20) e negativamente regulado em 72 hai (-0,56) em WH82. No genótipo Shiranui houve regulação positiva igualmente em 12 hai (8,62) e novamente em 72 hai (5,97).

Figura 9 – Níveis de expressão gênica dos GmMYBs no genótipo Williams 82 (WH82) em cinco tempos após inoculação com *P. pachyrhizi* (12, 24, 48, 72 e 96 hai). As barras de erro representam o desvio padrão para três experimentos independentes, com três repetições técnicas. “Up” e “Down” indicam tempo nos quais os genes apresentam expressão significativa ($p < 0,05$).

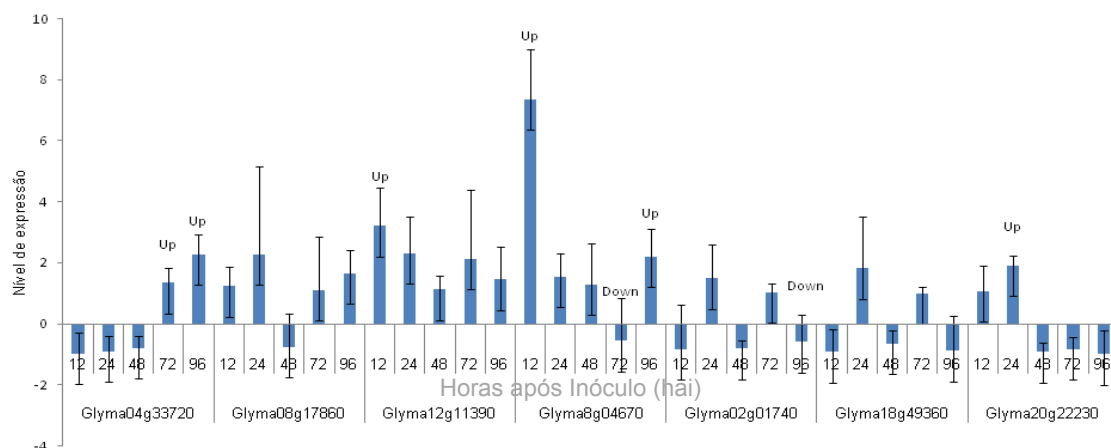


Figura 10 – Níveis de expressão gênica dos GmMYBs no genótipo resistente PI230970 (Rpp2) em cinco tempos após inoculação com *P. pachyrhizi* (12, 24, 48, 72 e 96 hai). As barras de erro representam o desvio padrão para três experimentos independentes, com três repetições técnicas. “Up” e “Down” indicam tempo nos quais os genes apresentam expressão significativa ($p < 0,05$).

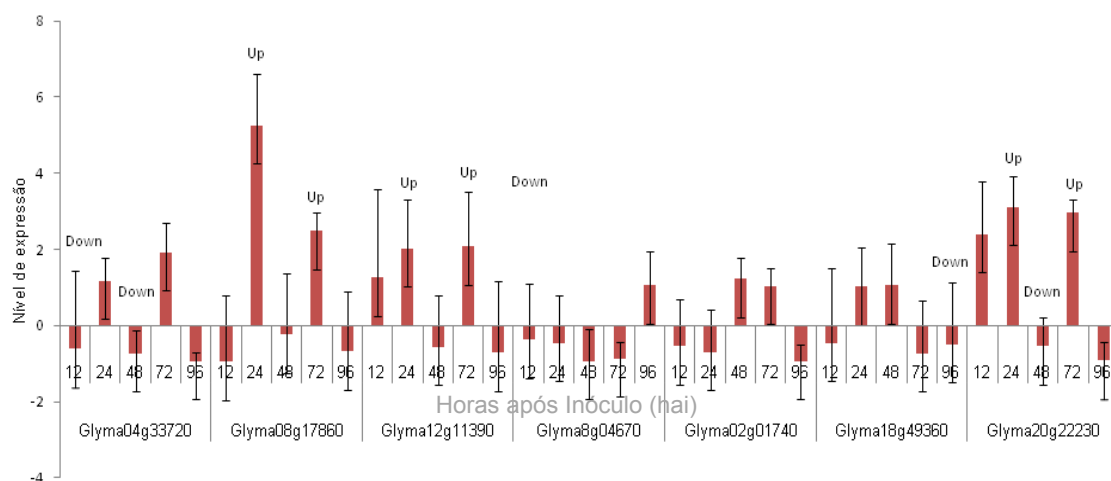
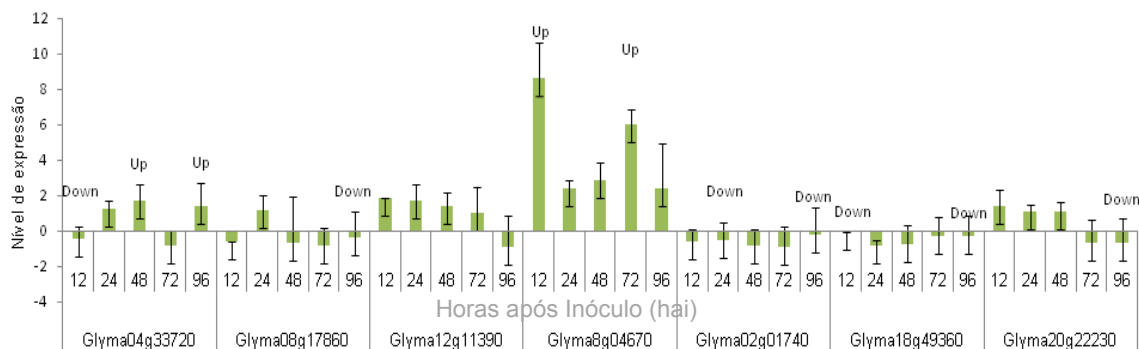


Figura 11 – Níveis de expressão gênica dos GmMYBs no genótipo resistente Shiranui (Rpp5) em cinco tempos após inoculação com *P. pachyrhizi* (12, 24, 48, 72 e 96 hai). As barras de erro representam o desvio padrão para três experimentos independentes, com três repetições técnicas. “Up” e “Down” indicam tempo nos quais os genes apresentam expressão significativa ($p < 0,05$).



De modo inesperado, diferentemente do observado por PANDEY et al. (2011), o gene do MYB84 não apresentou indução no genótipo PI230970, sendo no entanto regulado negativamente (-0,45) no tempo de 24hs (hai). Possivelmente, como o proposto anteriormente, a regulação da via de biossíntese de fenilpropanóides na PI230970 pode estar sendo realizada pelo MYB Glyma08g17860, ou mesmo outro fator de transcrição, considerando a complexidade da rota. Apesar das inúmeras evidências demonstrarem a participação da lignina como parte do mecanismo de defesa, a indução dos genes de regulação da via de síntese deste composto, entre os quais MYBs, permanece pouco compreendida e estudada, mas espera-se que diferentes fatores possam regular a mesma via.

O gene Glyma02g01740 agrupado com o *AtMYB11*, envolvido na regulação da via dos fenilpropanóides e biossíntese de flavonóides, não apresentou expressão diferencial em PI230970, e apresentou regulação negativa em 96 hai (-0,59) no genótipo WH82, e nos tempos de 24 hai (-0,55) e 96 hai (-0,21) em Shiranui. Uma possível explicação para a falta de indução da via dos flavonóides seria a maior indução da síntese de lignina, uma vez que a indução de GmMYBs potencialmente envolvidos na regulação da síntese de lignina foi observada nos três genótipos, o que promoveria o desvio dos intermediários comuns das duas vias para a rota da lignina, em detrimento da produção dos flavonóides. Esta hipótese é reforçada por ZAMBALA et al. (2006), que observou a expressão elevada de genes das rotas de fitoalexinas e lignina em detrimento da expressão de genes da rota de flavonóides quando estudou a resposta de soja a *Pseudomonas syringae*.

O Glyma18g49360 não apresentou indução em nenhum dos genótipos em resposta a FAS. Em PI230970 ele foi regulado negativamente em 96 hai (-0,48) e em Shiranui em 12 hai (-0,04) e novamente em 96 hai (-0,27). Tal fato não era esperado, uma vez que baseado na similaridade com o guia AtMYB30, que em *A. thaliana* desempenha um controle sobre respostas de defesa mediadas por SA (HR) (DANIEL et al., 1999; VAILLEAU et al., 2002, RAFFAELE et al., 2006), tal gene desempenharia importante papel na resposta a fungos biotróficos, principalmente mediada por SA.

O Glyma20g22230, agrupado no grupo 13 juntamente de AtMYB61, envolvido no controle sobre a síntese de lignina em *Arabidopsis thaliana* (NEWMAN et al., 2004; ZHONG; YE, 2009), foi regulado positivamente em WH82 em 24 hai (1,91) e em PI230970 em 24 hai (3,11) e 72 hai (2,96), enquanto nos genótipos resistentes PI230970 e Shiranui apresentou repressão em 48 hai (-0,55) e 96 hai (-0,69) após a infecção, respectivamente. O perfil de expressão diferencial dos genes positivamente regulados em pelo menos um genótipo/tempo está sumarizado na tabela 7.

Tabela 7 – Expressão geral dos *GmMYB* com indução nos três genótipos de soja após a infecção com *P. pachyrhizi*

Expressão comum entre genótipos				
<i>GmMYB</i>	Função Putativa ¹	Genótipo	Tempo (hai)	Expressão (RQ) *
Glyma04g33720	Resposta à quitina	WH82	72	1,33
			96	2,26
		Shiranui	48	1,67
			96	1,35
Glyma12g11390	?	WH82	12	3,20
			24	2,01
		PI230970	72	2,08
Glyma08g04670	PANDEY et al., (2011) Síntese de lignina	WH82	12	7,34
			96	2,20
		Shiranui	12	8,62
			72	5,97
Glyma20g22230	Síntese de lignina	WH82	24	1,91
			24	3,11
		PI230970	72	2,96
Expressão única				
Glyma08g17860	Síntese de lignina	PI230970	24	5,25
			72	2,48

*RQ significativo ao nível de 5% nas análises de expressão no Software REST 2009 ¹De acordo com o seu guia AtMYB; “?” Função desconhecida

Ao se analisar o perfil de expressão em termos quantitativos e também quanto à duração e pico de resposta, observa-se que mesmo havendo indução nos genótipos R e S, os genótipos R normalmente apresentam valores mais elevados ou resposta mais prolongada, conforme observado para os genes Glyma04g33720, Glyma12g11390 e Glyma20g22230.

O Glyma04g33720 foi expresso em ambos os genótipos WH 82 (suscetível) e Shiranui (resistente), sendo levemente superior em WH 82 (96hs), mas 24 horas mais rápida em Shiranui (48 hai). Esta indução rápida e intensa na planta resistente era esperada, visto que enzimas como as quitinases compõem a primeira linha de defesa da planta (respostas basais), que visa inibir o crescimento da hifa de infecção e, desta forma, a propagação e colonização pelo fungo. As respostas basais (PTI) são respostas rápidas (primeira resposta ativada) desencadeadas pelo reconhecimento de PAMPs, de forma imediata ao contato (VAKHRUSHEVA et al., 2010). Uma expressão tardia e intensa de um FT possivelmente controlando a expressão de genes de quitinases em planta suscetível também era esperada. Nestas plantas a demora na resposta de defesa contra o patógeno permite uma maior colonização pelo mesmo e, conseqüentemente, uma maior indução dos genes responsivos à quitina em resposta a está intensa propagação. De modo geral, plantas suscetíveis apresentam incapacidade ou demora no reconhecimento de PAMPs e efetores, promovendo respostas lentas ao invasor e culminando na doença (GARCIA-BRUGGER, 2006). Desta maneira, um ponto chave da importância deste tipo de resposta na interação de soja com FAS aparenta residir na velocidade e intensidade de indução de quitinases, possivelmente controlada por MYB.

Diferentemente de Shiranui, a regulação sobre genes responsivos à quitina foi reprimida em PI230970 já nos tempos iniciais de infecção avaliados. Desse modo, parece que os mecanismos de resposta a infecção por FAS mediada em Rpp2 diferem quanto ao envolvimento de genes de resposta a quitina. É possível que neste genótipo as respostas relacionadas com a biossíntese de lignina e fenilpropanóides sejam as mais importantes, conforme observado pelo maior número de MYBs regulando tais vias ativadas ao longo da infecção e sejam eficientes para inibir a propagação, o que não resulta na indução de respostas de genes responsivos ao elicitor quitina.

O Glyma12g11390 foi expresso em maior nível em WH82 (3,20), no entanto se manteve expresso por um período mais longo em PI230970, às 24 hai e

novamente em 72 hai, enquanto o Glyma20g22230 apresentou uma expressão maior e mais prolongada no genótipo resistente PI230970 (24 e 72 hai) quando comparado ao genótipo WH82, no qual houve uma leve indução apenas no tempo de 24 hai.

Interessantemente, o GmMYB sem função conhecida Glyma12g11390, e os GmMYBs ligados ao controle de lignina, Glyma08g04670, Glyma20g2223 e Glyma08g17860, foram todos expressos em pelo menos um dos genótipos resistentes em 72hs (hai). Em trabalhos recentes (MORTELL et al., 2007; SCHNEIDER et al., 2011) a análise da expressão de genes relacionados a defesa ao fungo *P. pachyrhizi* em soja, demonstrou haver um pico de resposta caracteristicamente observado em plantas com resistência em 72 hai.

Com três GmMYBs possivelmente controlando a rota de biossíntese de lignina induzidos em resposta a *P. pachyrhizi*, esse processo se destacou como principal mecanismo na resposta de defesa a este patógeno nos genótipos estudados, e principalmente em PI230970.

A lignina é o segundo biopolímero mais abundante no planeta, depois da celulose, e representa 30% do carbono orgânico da biosfera (BOERJAN; RALPH; BAUCHER, 2003). É um complexo polímero aromático formado a partir de três unidades de fenilpropano hidroxilado e metilado, sinapil, E-coniferil e p-coumaril, ligados através de acoplamento oxidativo, que é catalisado por enzimas peroxidases e lipoxigenases. Reações oxidativas são essenciais para a formação dos precursores da lignina e o reforço da parede celular, promovendo ligações cruzadas entre componentes, e é também comumente observado nos sítios de penetração fúngica como resposta a infecção. O processo de lignificação apresenta formas variadas de atuar na defesa da planta frente a um patógeno, tais como a criação de barreiras mecânicas. A deposição de lignina nos sítios de infecção (papilas) limitam a invasão e retém o patógeno e são acompanhadas de modificações químicas que reforçam a parede celular para resistir à ação de enzimas que degradam esta estrutura e que impedem a passagem de toxinas produzidas pelo patógeno e a saída de nutrientes do hospedeiro. Observa-se ainda produção de precursores tóxicos e radicais livres neste processo (NICHOLSON; HAMMERSCHMIDT, 1992; BHUYAN et al., 2009; NAOUMKINA et al., 2011).

Além da atuação da lignina como componente de barreiras estruturais pré-formadas da planta (reforço da parede), evidências apontam que sua

síntese e acúmulo podem ser induzidos como um mecanismo que compõem a resistência de algumas plantas frente a patógenos, uma vez que a ativação da via dos fenilpropanóides que formam os monolignóis é observada após contato com patógenos e seus elicitores, tais como quitosana, extrato de parede fúngica e soluções de quitina (BRUCE; WEST, 1989). Uma evidência da ativação da via de síntese de lignina foi apresentado por MAUCH-MANI; SLUSARENKO (1996), que infectou *A. thaliana* (col-0), transformada com o promotor da enzima chave da rota fenilalanina amônia-liase1 (*PAL1*) regulando a expressão do gene marcador GUS (β -glicosidase), com o fungo biotrófico *Peronospora parasitica* (cepa virulenta NOCO, avirulenta EMWA). O autor observou indução do promotor de *PAL1* nas células próximas às estruturas fúngicas em locais de penetração interrompidas com HR e lignificação epidérmica, apenas na interação incompatível. O pré-tratamento das plantas com um inibidor de PAL (ácido 2-aminoindano-2-fósfonico, AIP), bem como por um inibidor específico da via de síntese de lignina (2-hidroxifenil-aminosulfonil ácido acético, OH-PAS), para a enzima CAD, tornaram as plantas suscetíveis, demonstrando que a enzima PAL e CAD desempenham um papel na resistência em *A. thaliana* à *P. parasitica*, provendo precursores necessários para a lignificação nos locais de penetração.

MENDEN; KOHLHOFF e MOERSCHBACHER, (2007) injetando um elicitor do fungo biotrófico *Puccinia graminis* (Pgt), causador da ferrugem do colmo, em folhas saudáveis de trigo de alta resistência (Prelude-Sr5), observaram como resultado da elicitação um aumento no conteúdo de lignina, determinado através de dois diferentes métodos, de 50 a 160% superiores ao tratamento controle (tratamento com água).

Evidências da indução de genes R2R3-MYBs por elicitores é dada por SUGIMOTO; TAKEDA; HIROCHIKA, (2000) que observaram a expressão de um MYB de *Nicotiana tabacum*, NtMYB2 (R2R3), em resposta a elicitores (Onozuka celulase R10, oligômero de quitinase e xilanase). Os autores observaram que NtMYB2 possui a capacidade de se ligar ao motivo L-box e H box-like, encontrado na região promotora do gene *PAL2*, o que indica que o FT tem papel no controle sobre a via dos fenilpropanóides no processo de resposta à estresse ambiental.

Geralmente a tentativa de fungos invadirem a planta é acompanhada pelo desenvolvimento local das paredes celulares, formando as papilas que têm como um dos principais constituintes a lignina. BHUIYAN et al., (2009) estudaram o

papel funcional de unidades monolignóis precursores da lignina, na resposta de defesa em trigo (*Triticum monococcum*) com capacidade de impedir a penetração do oídio *Blumeria graminis* (*Bgt*), através de silenciamento via RNAi (RNA interferente) de genes da via de síntese destas substâncias. Oito genes de síntese de monolignol foram identificados em bibliotecas de cDNA derivados de trigo infectado com *Bgt*, e seis (*TmPAL*, *TmCCR*, *TmCCoAMT*, *TmF5H*, *TmCAOMT*, e *TmCAD*) foram fortemente acumulados, de modo sincronizado, e expressados em resposta ao oídio em plantas resistentes, predominantemente na epiderme. O silenciamento (simples e duplo) de todas as enzimas e a redução de monolignóis resultante, confirmada pela perda de fluorescência das papilas onde elas são depositadas como reforço (enfraquecimento da barreira), aumentou a eficiência de penetração foliar do patógeno em 40 a 78 % (construções duplas). A combinação do silenciamento da enzima comum da via dos fenilpropanóides PAL, com cada uma das duas enzimas biossintéticas finais de monolignóis (silenciamento duplo) demonstraram que, durante a penetração da planta, o único produto relevante da via foram os monolignóis e que estes produtos são componentes essenciais de papilas efetivas.

O processo de biossíntese de monolignóis, e consequentemente da lignina, é um processo metabólico dispendioso que investe carbono de modo irreversível, o que exige um controle forte sobre a ativação desta via e justifica a necessidade de fatores de transcrição atuando sobre várias enzimas. Vários genes da via dos fenilpropanóides, incluindo os que codificam as enzimas PAL, C4H, COMT, CCR e CAD, compartilham um motivo conservado (elemento AC) na região promotora, nos quais se ligam fatores de transcrição, incluindo MYBs, que regulam toda a síntese de lignina. Em *A. thaliana*, uma rede complexa de fatores de transcrição é responsável pela regulação da biossíntese dos constituintes da parede secundária. Dois FTs NAC (NST1 e NST3) são os controladores principais que ativam a síntese de todos os componentes, incluindo a via dos monolignóis e lignina. Entre os genes controlados por NAC estão alguns MYBs (*AtMYB58* e *AtMYB63*) entre outros desconhecidos, que ativam direta e especificamente as enzimas de produção destas substâncias (NAOUMKINA et al., 2011).

Além da ativação da via para lignificação constitutiva, evidências sugerem que alguns dos genes que compõem a via desempenham papéis também em resposta a infecção patogênica, fornecendo indícios de que cabe a fatores de transcrição essa fina regulação da expressão de determinados pontos da via em

resposta as circunstâncias específicas nas quais a planta se encontra. LAUVERGEAT et al. (2011) caracterizaram funcionalmente dois clones de cDNA da primeira enzima específica para a síntese de lignina CCR (AtCCR1 e AtCCR2) em *A. thaliana*, e observaram que os dois genes são expressos diferencialmente durante o desenvolvimento (AtCCR1 em pecíolos) e fortemente em resposta a infecção à cepa avirulenta de *Xantomonas campestris* pv. *Campestris* (AtCCR2). NAOUMKINA et al. (2011) anotou 35 genes que codificam CCR, após análise por similaridade utilizando a ferramenta BLASTX de sequências de soja no banco de dados da Uniprot. De modo geral, *G. max* apresenta um maior número de genes das enzimas da via de síntese da lignina (exceto C3H) quando comparado a *A. thaliana*, o que poderia estar relacionado com o fato de existir um maior número também de representantes GmMYBs nos subgrupos de controle de síntese deste composto (exemplo S13 e C11) em soja.

Em trabalhos anteriores, dez genes classificados na classe de processo de síntese de lignina (Gene ontology), bem como a primeira enzima da via dos fenilpropanóides, fenilalanina amônia liase (PAL), foram significativamente expressos nos genótipos PI230970 e Embrapa-48 durante infecção por *P. pachyrhizi* em trabalho de MORTEL et al. (2007). A análise do perfil de expressão destes genes em ambos os genótipos indicou similaridade nas etapas iniciais de infecção com aumento em 24 hai (resposta de defesa geral), divergindo posteriormente, com segundo pico de expressão, contudo, com um dia de antecedência na planta resistente (72hai). Neste genótipo a expressão diferencial de genes, como os de síntese de lignina e defesa, foi acompanhada por expressão de genes de regulação da transcrição, classe expressa unicamente na planta resistente. A família MYB foi a segunda mais expressa entre todos os FTs, representando 10,23%, e entre aqueles associados com respostas à estresses (18,57%) no tempo de 72 hai.

Devido à evidência concreta do papel do controle exercido por um MYB (MYB84 - fenilpropanóides), assim como da lignificação (GmWRKY36 e GmDBFT) na resistência de PI230970 à *P. pachyrhizi*, fornecida pelo trabalho de PANDEY et al. (2011), somados a todos os demais indícios encontrados nos outros trabalhos citados até o momento, que reforçam a importância da lignificação como mecanismo que comumente compõem a defesa de plantas, o presente estudo focou em subgrupos de MYBs de *A. thaliana* que exercem controle sobre a síntese deste composto. Soma-se a estas evidências a importância da necessidade da realização

de maiores estudos dos integrantes da família MYB em soja. Como já relatado em *A. thaliana* (ZHOU et al., 2009; ZHONG; YE, 2009; NEWMAN et al., 2004; ZHONG; RICHARDSON; YE, 2007; ZHONG et al., 2008), esta família de FTs são capazes de regular toda a via de síntese da lignina, de modo que são potenciais candidatos para exercerem este controle em soja e, conseqüentemente, se tornam alvos de estudos funcionais frente à FAS.

Como destacado anteriormente (SCHNEIDER et al., 2011), a redução da velocidade de desenvolvimento da infecção a partir de uma resposta rápida parece ser a chave para a resistência em soja, tornando, conseqüentemente, o controle exercido sobre a resposta um importante fator. Como os resultados do presente estudo sugerem, o controle putativo sobre três pontos da via de biossíntese dos fenilpropanóides (fenilalanina amônia liase; 4-coumarato-CoA-ligase; Cinnamato-4-hidroxilase), comuns para os três MYBs induzidos nas plantas resistentes (MYB84, AtMYB58 e AtMYB61) durante infecção por FAS, indicam que possivelmente, em soja, a síntese de lignina também é um fator importante para a defesa, e talvez à resistência, bem como que FTs MYBs podem também estar contribuindo diretamente através do controle do tempo de expressão deste componente.

No presente trabalho, genes GmMYBs possivelmente exercendo controle sobre a via dos fenilpropanóides (Glyma08g04670, Glyma20g2223 e Glyma08g17860) apresentaram o mesmo perfil de expressão observado por MORTEL et al. (2007) para genes dessa via e FTs, com uma indução em um segundo pico de expressão (72hai), mais rápida (Glyma08g04670) e intensa em plantas resistentes. Para o autor, a resistência mediada por Rpp2 em PI230970, de modo geral, provém desta velocidade do controle diferencial dos genes por FTs. SCHNEIDER et al., (2011) também observou este padrão, associando a expressão diferencial de genes de defesa no tempo de 72hai e a capacidade da planta reduzir o desenvolvimento fúngico neste ponto.

A outra função sobre qual GmMYB pode exercer controle e que foi induzida de forma mais rápida no genótipo resistente (Shiranui) foi sobre genes de resposta a quitina (AtMYB31), pelo Glyma04g3372.

A quitina é um PAMP, sendo um constituinte comum encontrado em patógenos fúngicos e, como os demais PAMPs, é capaz de induzir vias de defesa em plantas (elicitação). Estas estruturas são capazes de induzir desde morte celular

(HR), síntese de substâncias antimicrobianas (fitoalexinas), geração de espécies reativas de oxigênio (ROS), a ativação de proteínas PR (pathogenesis-related), tais como quitinases, e reforços da parede tais como processos de lignificação e formação de calose (SHIBUYA; MINAMI, 2001). ZHANG et al. (2002) demonstrou que mudas de *A. thaliana* (Col-o; No-o para mutantes *npr1*) mutantes para a via de sinalização de JA (*jar1*), etileno (*ein2*) e SA (*pad4*, *eds5*, *npr1* e *cpr1*) tratadas com quitina pura em pó (caranguejo), não foram capazes de bloquear a rápida expressão de genes induzidos por quitina, indicando que as respostas de defesa desencadeadas por esse elicitor são independentes e possivelmente fazem parte de uma via de resposta única.

RAMONELL et al. (2005) identificou, através de microarranjo e por northern-blot, uma série de genes responsivos a quitina em *A. thaliana* tratada com esse elicitor (quito-octâmero hidrolisado). Criando mutações (via T-DNA) para alguns desses genes, um *zinc-finger* e duas proteínas de resistência à doença (Disease resistance protein-like), ele observou aumento na susceptibilidade ao oídio *Erysiphe cichoracearum*, comprovado através da constatação de aumento do número de esporos por colônia e na superfície foliar, apenas nas plantas mutantes com supressão destes genes de resposta a quitina.

Esta capacidade da quitina elicitar a expressão gênica, algumas vezes de forma específica, têm sido atribuído ao envolvimento de fatores de transcrição. Como citado anteriormente em um trabalho de análise de FTs responsivos a quitina em *A. thaliana*, LIBAULT et al. (2007) identificou 11 MYBs, entre os quais AtMYB31, sendo induzidos exclusivamente na planta tratada com quitina. Os autores ainda correlacionaram 18 FT responsivos ao tratamento com quitina (identificados por eles) à capacidade de responder à *Alternaria brassicola*, incluindo a família MYB (AtMYB51 e AtMYB27), observados por McGRATH et al. (2005), o que representa uma evidência de que estes genes FT possivelmente atuam integrando e controlando respostas específicas de defesa que compõem a resistência, ao reconhecimento de um elicitor, neste caso a quitina.

SORIA-GUERRA et al. (2010) comparando a expressão diferencial de genes entre plantas de *Glycine tomentella* resistentes (R) e suscetíveis (S), através de microarranjo com 38.400 genes de soja, em resposta de resistência à FAS, observou 57 genes diferencialmente expressos entre os genótipos R e S em 12 (hai), dos quais 12 genes de defesa, incluindo uma quitinase. Esta enzima foi

regulada positivamente no genótipo resistente e reprimida no genótipo suscetível, ambas em estágios iniciais da infecção (12 hai). Embora o GmMYB analisado no estudo (Glyma4g33720), ortólogo de AtMYB31, apresenta expressão em ambas as plantas, WH82 (S) e Shiranui (R), esta foi um dia mais rápido na planta resistente (48hai), possivelmente modulando a expressão de genes de resposta a quitina nas fases iniciais de respostas (PTI), compondo assim a respostas de defesa geral expressas da planta nas primeiras horas de infecção, quando ocorre a penetração por *P. pachyrhizi*. Isto condiz com a proposição de SCHNEIDER et al. (2011) de que *P. pachyrhizi* elícita respostas similares a PTI (imunidade desencadeada por patógenos), não específicas, nestas fases iniciais de penetração, através de elicitores do tipo PAMP, tais como quitina, como comentado anteriormente.

Pelo fato da soja possuir um impactante valor econômico ao homem, em especial ao nosso país, estudos de potenciais genes que contribuam para a resistência desta planta, e que representam um meio de aumentar a produção deste produto, necessitam ser valorizados. Desta maneira, GmMYB, bem como o seu papel no controle sobre a síntese de lignina e a resposta a elicitores como a quitina, se tornam candidatos obrigatórios de novos estudos.

Os resultados obtidos neste trabalho podem contribuir para a elucidação de funções de outros membros da família GmMYB, bem como o papel dos mesmos durante a resposta de defesa de soja à FAS e a outros agentes de estresse biótico.

6 CONCLUSÕES

No presente estudo a busca em bancos de sequências de MYBs no genoma de soja, através de bioinformática permitiu realizar um estudo mais detalhado, o que resultou na redução do número total de integrantes da vasta família de FT MYBs em soja para 264, contudo, melhorando o conhecimento a cerca dos mesmos, visto que, muitas das sequências alocadas nos bancos não apresentavam a estrutura mínima conservada da família. Estes dados ainda tornam GmMYB a maior família R2R3-MYB conhecida entre as plantas.

O alinhamento e a construção da árvore filogenética a partir das sequências R2R3-MYBs de soja (GmMYBs) e AtMYBs da planta modelo *Arabidopsis* permitiu observar grande similaridade em sequências de FT desta família entre as duas espécies de plantas. Com base na similaridade estrutural observada entre os membros de cada subgrupo na porção conservada N-terminal, aliada a identificação de estruturas comuns conservadas na região altamente variável C-terminal (motivos) e promotora (*cis*-motivos), pode-se atribuir funções putativas aos GmMYBs, incluindo defesa. Análises de expressão validaram tal agrupamento, uma vez que apenas GmMYBs com maior similaridade com seus guias apresentam expressão diferencial em soja durante a interação com os patógenos. Foi possível ainda identificar dois subgrupos (C5 e C8) novos para a família MYB de plantas, formados exclusivamente de GmMYBs.

A análise *in silico* do perfil transcricional dos GmMYBs durante a interação da soja com diferentes patógenos demonstrou o envolvimento relevante destes FTs nas respostas de defesa em *G. max*, uma vez que, cerca de um terço (32,4%) dos GmMYBs apresentaram expressão frente a *P. pachyrhizi* e 40% a ao menos um dos demais patógenos (*Aphis glycines*, *Heterodera glycines*, *Phytophthora sojae* e *Pratylenchus brachyurus*).

Análises do padrão de expressão via RT-qPCR de GmMYBs potencialmente envolvidos na resposta de defesa a doenças em soja, com base na similaridades com seus guias em *Arabidopsis*, demonstraram que 5 dos 7 GmMYBs selecionados foram induzidos em pelo menos um tempo e/ou genótipo em resposta a infecção com *P. pachyrhizi*.

Estes dados apontam que os processos celulares potencialmente regulados por este GmMYBs, biossíntese de lignina, fenilpropanóides e quitina,

podem contribuir na resposta de resistência a FAS mediada por Rpp2 e Rpp5. Deste modo, GmMYBs pertencentes a estes subgrupos representam promissores candidatos de futuros estudo para comprovação deste potencial papel na defesa de soja.

REFERÊNCIAS BIBLIOGRÁFICAS

- AEN, 2009. Disponível em <<http://www.aen.pr.gov.br/modules/noticias/article.php?storyid=47923>>. Acesso em 12 de Maio de 2011.
- ALVES, S.A.M. **Quantificação de parâmetros da pré-penetração e monocíclicos relacionados ao patossistema *Phakopsora pachyrhizi*-soja**. 2007. 64p. Tese (Doutorado) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2007.
- ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. Controle da expressão gênica. In: ALBERTS, B. *Biologia Molecular da Célula*. 4. Ed. Porto Alegre: Artmed, 2006. p. 376-466.
- ANP - AGÊNCIA NACIONAL DE PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS. **Anuário Estatístico 2010** Disponível em: <<http://www.anp.gov.br/?pg=31286>>. Acesso em 16 de outubro de 2011.
- ASSUMPÇÃO, L. DE CASTRO. **Diversidade da comunidade bacteriana endofítica de sementes de soja e o seu potencial biológico**. 2008 Dissertação (Mestrado – Agronomia) – Universidade de São Paulo, Escola Superior de Agricultura Luiz de Queiroz, Piracicaba. 2008.
- BAILEY, T.L.; WILLIAMS, N.; MISLEH, C.; LI, W.W. MEME: discovering and analyzing DNA and protein sequence motifs. **Nucleic Acids**, Inglaterra, v. 34, p. 369-373, Jul. 2006.
- BALARDIN, R.S. **Doenças da soja**. Santa Maria: Edição do Autor, 2002. 100p.
- BHUIYAN, N. H.; SELVARAJ, G.; WEI, Y.; KING, J. Gene expression profiling and silencing reveal that monolignol biosynthesis plays a critical role in penetration defense in wheat against powdery mildew invasion. **Journal of Experimental Botany**, Inglaterra, v. 60, n. 2, p. 509-521, Nov. 2009.
- BOERJAN, W.; RALPH, J.; BAUCHER, M. Lignin Biosynthesis. **Annual Review of Plant Biology**, v.54, p. 519-546, Jun. 2003.
- CONAB: Acompanhamento da safra brasileira – grãos: Décimo segundo levantamento Safra 2011/2012, **Companhia Nacional de Abastecimento: CONAB**, Brasília, Set. 2012
- CONAB: Acompanhamento da safra brasileira – grãos: Quinto levantamento Safra 2012/2013, **Companhia Nacional de Abastecimento: CONAB**, Brasília, fev. 2013
- CONAB. Balança Comercial do Agronegócio. Disponível em <<http://www.conab.gov.br/conteudos.php?a=547&t=2>> Acesso em 10 de Fevereiro de 2013.
- CHOI, J.J.; ALKHAROUF, N.W.; SCHNEIDER, K.T.; MATTHEWS, B.F.; FREDERICK, R.D. Expression patterns in soybean resistant to *Phakopsora*

pachyrhizi reveal the importance of peroxidases and lipoxigenases. **Functional & Integrative Genomics**, v. 8, n. 4, p. 341-359, mar. 2008

CHISHOLM, S. T., COAKER, G., DAY, B. AND STASKAWICZ, B. J. Host-microbe interactions: shaping the evolution of the plant immune response. **Cell**, v.124, n. 4, p. 803-814, Fev. 2006.

DANIEL, X.; LACOMME, C.; MOREL, J. B.; ROBY, D. A novel *myb* oncogene homologue in *Arabidopsis thaliana* related to hypersensitive cell death. The **Plant Journal**, v.20, n. 1, p. 57-66. Ago.1999.

DA SILVA, D.C.G. **Mapeamento de genes de resistência da soja à ferrugem asiática e análise transcricional na interação patógeno hospedeiro**. Tese (Doutorado – Agronomia) Faculdade de Ciências Agrárias e Veterinárias – UNESP, Jaboticabal, 2008.

DONG, W.; NOWARA, D.; SCHWEIZER, P. Protein polyubiquitination plays a role in basal host resistance of barley. **Plant Cell**, Rockville, v. 18, n. 11, p. 3321-3331, Nov.2006.

DODDS, P.N.; RATHJEN, J.P. Plant immunity: Towards an integrated view of plant-pathogen interactions. **Nature Reviews: Genetics**. v.11, p.539-548, Ago. 2010.

DU, H.; YANG, S.S.; LIANG, Z.; FENG, B.R.; LIU, L.; HUANG, Y.B.; TANG, Y. X. Genome-wide analysis of the MYB transcription factor superfamily in soybean. **Plant Biology**, v. 12, p. 1-22, Jul. 2012

DUBOS, C.; STRACKE, R.; GROWOLD, E.; WEISSHAAR, B.; MARTIN, C.; LEPINIEC, L. MYB transcription factors in *Arabidopsis*. **Plant Science**, v. 15, n. 10, p. 573-581, Out. 2010.

EMBRAPA SOJA, 2008. Disponível em <<http://www.embrapa.br/imprensa/noticias/2008/outubro/3a-semana/brasil-aprende-a-combater-ferrugem-da-soja/>> Acesso em 12 de Maio de 2011.

EMBRAPA SOJA, 2009. Disponível em <http://www.cnpso.embrapa.br/noticia/ver_noticia.php?cod_noticia=557&desl=121> Acesso em 6 de Junho de 2011.

EMBRAPA, 2010. Disponível em <http://www.agencia.cnptia.embrapa.br/gestor/especies_arboreas_brasileiras/arvore/CONT000fu17wvyo02wyiv807nyi6s9ggg9il.html>. Acesso em: 16 de Out. de 2011.

EMBRAPA, 2011. História: a origem do grão. Disponível em <http://www.cnpso.embrapa.br/index.php?op_page=112&cod_pai=33>. Acesso em: 16 de Out. de 2011.

FELLER, A.; MACHEMER, K.; BRAUN, E. L.; GROTEWOLD, E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. The **Plant Journal**, v.66, p.94-116, Dez. 2011.

FEHR, W.R.; CAVINESS, C.E. Stage of soybean development. **Iowa State University. Special report 80**, Mar. 1981.

FLORS, V.; TON, J. JAKAB, G.; MAUCH-MANI. Abscisic acid and callose: Team players in defense against Pathogens **Journal of Phytopathology**, v. 153, p. 377-383, Ago. 2005

FREITAS, M.C.M. A cultura da soja no Brasil: O crescimento da produção brasileira e o surgimento de uma nova fronteira agrícola. **Enciclopédia Biosfera, Centro Científico Conhecer**, v.7, n.12, p.1-12, Maio 2011.

GARCIA-BRUGGER, A.; LAMOTTE, O.; VANDELLE, E.; BOURQUE, S.; LECOURIEUX, D.; POINSSOT, B.; WENDEHHENE, D.; PUGIN, A. Earling Signaling events induces by elicitors of plants defense. **Molecular Plant Microbe Interactions**, v.19, n.7, p. 711-724, Jul. 2006.

GERRATS, B.P.J.; BAKKER P.A.H.M.; VAN LOON, L.C. Ethylene insensitivity impairs resistance to soilborne pathogens in *Tobacco* and *Arabidopsis thaliana*. **Molecular Plant Microbe Interactions**, v.14, n.10, p. 1078-1085, Out. 2002

GERRATS, B.P.J.; BAKKER P.A.H.M.; LAWRENCE C.B.; ACHUO, E.A.; HÖFTE, M.; VAN LOON, L.C. Ethylene-insensitive Tobacco shows differentially altered susceptibility to different pathogens. **Genetics and Resistance**, v. 93, n. 7, p .813-821, Jul. 2003.

GLAZEBROOK, J. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. **Annual Review of Phytopathology**, v. 43, p. 205-227, Jul. 2005

GODOY, C.V. Eficiência de fungicidas aprovados na RPSRCB no controle da ferrugem asiática da soja, em Londrina, PR. Embrapa Soja. In Resultados dos ensaios cooperativos para controle da ferrugem asiática da soja. Safra 2008/2009.p .66-71, 2009.

GONZALEZ, A.; MENDENHALL, J.; HUO, Y.; LLOYD, A. TTG1 complex MYBs, MYB5 and TT2, control outer seed coat differentiation. **Development Biology**, v. 325, n. 2, p. 412-421. Jan. 2009.

HAMMOND-KOSACK, K. E.; JONES, J. D. G. Resistance gene-dependent plant defense response. **The Plant Cell**, Rockville, v. 8, n. 10, p. 1773-1791, Out. 1996.

HAMMOND-KOSACK, K. E.; JONES, J. D. G. Responses to plant pathogens. In: BUCHANAN, B.; GRUISSEM, W., JONES, R. (Ed.) **Biochemistry and Molecular Biology of Plants**. Rockville, Maryland, American Society of Plant Physiologists, cap. 21, p. 1102-1157, 2000.

HENNING, P. Einige neue japanische Uredineen IV. **Hedwigia Beiblatt**, v. 42, p. 107-108, 1903.

HENNING, A.A.; GODOY, C.V. Situação da Ferrugem da Soja no Brasil e no Mundo. In **Ferrugem Asiática da Soja**. Editor: Laércio Zambolim. p.1-14, 2006.

HIGO, K.; UGAWA, Y.; IWAMOTO, M.; KORENAGA, T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. **Nucleic Acids Research**, v. 27, p. 297-300, Jan. 1999.

IMEA 2012. Instituto Mato-Grossense de Economia Agropecuária. Disponível em <<http://www.imea.com.br/>>. Acesso em 08 de Fevereiro de 2013.

JIANG, C.; GU, X.; PETERSON, T. Identification of conserved gene structure and carboxy-terminal motifs in the Myb genes family of *Arabidopsis thaliana* and *Oryza sativa* L.spp.*indica*. **Genome Biology**, v. 5, p. 1-11. Jun. 2004.

JUNIOR, S. L. B. **Análise de expressão gênica induzida por *Phakopsora pachyrhizi* em soja**. 2007. Dissertação (Mestrado – Genética e Melhoramento). Universidade Federal de Viçosa, Viçosa, 2007.

KANDOTH, P. K.; ITHAL, N.; RECKNOR, J.; MAIER, T.; NETTLETON, D.; BAUM, T. J.; MICTHUM, G. M. The soybean *Rhg1* locus for resistance to the Soybean cyst nematode *Heterodera glycines* Regulates the expression of a large number of stress and defense-related genes in degenerating feeding cells. **Plant Physiology**, v.155, n. 4, p. 1960-1975. Abr. 2011.

KLEMPNAUER, K.; GONDA, T.; BISHOP, J. Nucleotide sequence of the retroviral leukemia gene *v-myb* and its cellular progenitor *c-myb*: the architecture of a transduced oncogene. **Cell**, v.31, p. 453-463. Dez. 1982.

KÖHLE, H.; YOUNG, D. H.; KAUSS, H. Physiological changes in suspension-cultured soybean cells elicited by treatment with chitosan. **Plant Science Letters**, v. 33, p. 221-230. Fev. 1984.

KRANZ, H. D.; DENEKAMP, M.; GRECO, R.; JIN, H.; LEYVA, A.; MEISSNER, R. C.; PETRONI, K.; URZAINQUI, A.; BEVAN, M.; MARTIN, .; SMEEKENS, S.; TONELLI, C.; PAZ-ARES, J.; WEISSHAAR, B. Towards functional characterization of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. **The Plant Journal**, v. 16, p. 263-276. Ago. 1998.

LAUVERGEAT, V.; LACOMME, C.; LACOMBE, E.; LASSERRE, E.; ROBY, D.; GRIMA-PETTENATI, J. Two cinnamoyl-CoA reductase (CCR) genes from *Arabidopsis thaliana* are differentially expressed during development and in response to infection with pathogenic bacteria. **Phytochemistry**, v. 57, p. 1187-1195. Ago. 2001.

LAZZAROTTO, J. J.; HIRAKURI, M. H. **Evolução e perspectivas de desempenho econômico associadas com a produção de soja nos contextos mundial brasileiro**. Londrina: Embrapa Soja, p. 46, 2010. (Embrapa Soja. Documentos, 319).

LEE, M.W.; QI, M.; YANG, Y. A novel jasmonic acid-inducible Rice *myb* gene associates with fungal infection and host cell death. **Molecular Plant Microbe Interaction**, v. 14, p. 527-535. 2001.

LEWIS, G.P.; SCHRIRE, B.D.; MACKINDER, R.B.; LOCK, M. **Legumes of World**. London: Royal Botanic Gardens, Kew, 2005. 577p.

LIBAULT, M.; WAN, J.; CZECHOWSKI, T.; UDVARDI, M.; STACEY, G. Identification of 118 *Arabidopsis* transcription factor and 30 ubiquitin-ligase genes responding to chitin, a plant –defense elicitor. **Molecular Plant Microbe Interaction**, v. 20, p. 900-91. Ago. 2007.

LINO, E. J. **Identificação e caracterização de transcritos relacionados a mecanismos de resistência à ferrugem asiática da soja**. 2011. Dissertação (mestrado – genética e biologia molecular). Universidade Estadual de Londrina, Londrina, 2011.

LI-WANG, K.; BOLITHO, K.; GRAFTON, K.; KORTSTEE, A.; KARUNAIRETNAM S.; MCGHIE, T.K.; ESPLEY, R.V.; HELLENS, R.P.; ALLAN, A.C. An R2R3-MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. **BMC Plant Biology**, v. 10, p. 1-17. Mar. 2010.

McGRATH, K. C.; DOMBRECHT, B.; MANNERS, J.M.; SCHENK, P.M.; EDGAR, C.I.; MACLEAN, D.J.; SCHEIBLE, W.R.; UDVARDI, M.K.; KAZAN, K. Repressor and activator type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genome-wide screen of *Arabidopsis* transcription factor gene expression. **Plant Physiology**, v. 139, n. 2, p. 949-959. Out. 2005

MATTHAUS, B.; LUFTMANN, H. Glucosinolate in members of the family Brassicaceae: separation and identification by LC/ESI-MS-MS. **Journal of Agricultural and Food Chemistry**, v. 48, p. 2234-2239. Maio. 2000.

MAUCH-MANI, B.; SLUSARENKO, AL. Production of salicylic acid precursors is a major function of phenylalanine ammonia-lyase in the resistance of *Arabidopsis* to *Pernospora parasitica*. **The Plant Cell**, v. 8, n. 2, p. 203-212. Fev. 1996.

MENDEN, B.; KOHLHOFF, M.; MOERSCHBACHER, B.M. Wheat cells accumulate a syringyl-rich lignin during the hypersensitive resistance response. **Phytochemistry**, v. 68, p. 513-520. Fev. 2007.

MEHRTENS, F.; KRNAZ, H.; BEDNAREK, P.; WEISSHAAR, B. The *Arabidopsis* transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. **Plant Physiology**, v. 138, n. 2, p. 1083-1096. Maio 2005.

MENGISTE, T.; CHEN, X.; JOHN, S.; DIETRICH, R. The *BOTRYTIS SUSCEPTIBLE1* Gene encodes an R2R3MYB transcription factor protein that is required for biotic and abiotic stress response in *Arabidopsis*. **The Plant Cell**, v. 15, n. 11, p. 2551-2565. Nov. 2003.

NAOUMKINA, M.A.; ZHAO, Q.; GALLEGU, G.L.; DAI, X.; ZHAO, P.X.; DIXON, R.A. Genome-wide analysis of phenylpropanoid defense pathways. **Molecular Plant Pathology**, v. 11, n. 6, p. 829-846. Jul. 2010.

NEWMAN, L.J.; PERAZZA, D.E.; JUDA, L.; CAMPEBELL, M.M. Involvement of the R2R3-MYB, AtMYB61, in the ectopic lignifications and dark-photomorphogenic components of the *det3* mutant phenotype. **The Plant Journal**, v. 37, p. 239-250. Jan. 2004

NICHOLSON, R.L. Phenolic compounds and their role in disease resistance. **Phytopathology**, v. 30, p. 369-389. Set. 1992.

PASCHOLATI, S. F.; LEITE, B. Hospedeiro: mecanismos de resistência. In: BERGAMIM FILHO, A.; KIMATI, H.; AMORIN, L. (Ed.) Manual de fitopatologia. 3. ed. São Paulo: Ceres, 1995. p. 417-454.

PANDEY, A.K.; YANG, C.; ZHANG, C.; GRAHAN, M.A.; HORSTMAN, H.D.; LEE, Y.; ZABOTINA, O.A.; HILL, J.H.; PEDLEY, K.F.; WHITHAN, S.A. Functional analysis of the soybean rust resistance pathway mediated by *Rpp2*. **Molecular Plant Microbe Interaction**, v. 24, n. 2, p. 194-206. Fev. 2011.

PANTHEE, D.R.; YUAN, J.S.; WRIGHT, D.L.; MAROIS, J.J.; MAILHOT, D.; STEWART JR, C.N. Gene expression analysis in soybean in response to the casual agent of Asian soybean rust (*Phakopsora pachyrhizi* Sydow) in a early growth stage. **Functional & Integrative Genomics**, v. 7, p. 291-301, Out. 2007.

PANTHEE, D.R.; MAROIS, J.J.; WRIGHT, D.L.; NARVÁEZ, D.; YUAN, J.S.; STEWART JR, C.N. Deferential expression of genes in soybean in response to the casual agent of Asian soybean rust (*Phakopsora pachyrhizi* Sydow) is soybean growth stage –specific. **Theoretical Applied Genetics**. v. 118, p. 359-370, Jan. 2009

PAZ-ARES, J.; GHOSAL, D.; WIENAND, U.; PETERSON, P.A.; SAEDLER, H. The regulatory *c1* locus of *Zea mays* encodes a protein with homology to *myb* proto-oncogene products and with structural similarities to transcriptional activators. **The EMBO Journal**, v. 6, p. 3553-3558. Dez. 1987.

PFAFFL, M.W.; HORGAN, G. W.; DEMPFLER, L. Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. **Nucleic Acids Research**, v. 30, p. 1-10. Maio 2002.

PIETERSE, C.M.J.; LEON-REYES, A.; ENT, S.V.D.; WEES, C.M.V. Networking by small-molecule hormones in plant immunity. **Nature Chemical Biology**, v. 5, n. 5, p. 308-316, Abr. 2009

RAFFAELE, S.; RIVAS, S.; ROBY, D. An essential role for salicylic acid in AtMYB30-mediated control of the hypersensitive cell death program in Arabidopsis. **FEBS Letters**, v. 580, p.3498-3504. Jun. 2006.

RAMONELL, K.; BERROCAL-LOBO, M.; JINRONG, S.K.; EDWARD, H. STACEY, G.; SOMERVILLER, S. Loss-of-function mutations in chitin responsive genes show increased susceptibility to the powdery mildew pathogen *Erysiphe cichoracearum*. **Plant Physiology**, v. 138, p. 1027-1036. Jun. 2005.

RIECHMANN, J. L.; HEARD, J.; MARTIN, G.; REUBER, L.; JIANG, C.Z.; KEDDIE, J.; ADAM, L.; PINEDA, O.; RATCLIFFE, O. J.; SAMAHA, R. R.; CREELMAN, R.; PILGRIM, M.; BROUN, P.; ZHANG, J.Z.; GHANDEHARI, D.; SHERMAN, B. K.; YU, G. L. *Arabidopsis* Transcription Factors: Genome-wide comparative analysis among eukaryotes. **Science**, v. 290, p. 2105-2109, Dez. 2000.

ROMERO, I.; FUERTES, A.; BENITO, M.J.; MALPICA, J.M.; LEYVA, A.; PAZ-ARES, J. More than 80R2R3-MYB regulatory genes in the genome of *Arabidopsis thaliana*. **The Plant Journal**, v.14, p. 273-284. Maio 1998.

SEO, P. J.; PARK, C. M. MYB96-mediated abscisic acid signals induce pathogen resistance response by promoting salicylic acid biosynthesis in *Arabidopsis*. **New Phytologist**, v. 186, p. 471-483. Abr. 2010.

SEVERIN, A. J. WOODY, J. L.; BOLON, Y. T.; JOSEPH, B.; DIERS, B. W.; FARMER, A. D.; MUEHLBAUER, G. J.; NELSON, R. T.; GRANT, D.; SPECHT, J. E.; GRAHAM, M.; CANNON, S. B.; MAY, G. D.; VANCE, C. P.; SHOEMAKER, R. C. RNA-Seq Atlas of *Glycine max*: A guide to the soybean transcriptome. **BMC: Plant Biology**, v. 10, p. 1-16. Ago. 2010

SHNEIDER, K.T.; MORTEL, M.; BANCROFT, T.J.; BRAUN, E.; NETTLETON, D.; NELSON, R.T.; FREDERICK, D.; BAUM, T.J.; GRAHAN, M.A.; WHITIHAN, S.A. Biphasic gene expression changes elicited by *Phakopsora pachyrhizi* in soybean correlate with fungal penetration and haustoria formation. **Plant Physiology**, v.157, p.355-371, set. 2011.

SHIBUYA, N.; MINAMI, E. Oligosaccharide signaling for defense responses in plant. **Physiological and Molecular Plant Pathology**, v. 59, p. 223-233. 2001.

SILVEIRA, A. B. **Análise funcional do regulador de transcrição do tipo bZIP *AtbZIP9* de *Arabidopsis thaliana* através da superexpressão de seus genes alvos.** 2007. Dissertação (Mestrado em Genética e Biologia Molecular). Universidade Estadual de Campinas, Campinas. 2007.

SKADHAUGE, B.; THOMSEN, K.K.; WETTSTEIN, D.V. The role of the barley testa layer and its flavonoid content in resistance to *Fusarium* infection. **Hereditas**, v.126, p.147-160. Jun. 1997.

SORIA-GUERRA, R.E.; ROSALES-MENDOZA, S.; CHANG, S.; HAUNDENSHEILD, J.S.; PADMANABAN, A.; RODRIGUEZ-ZAS, S.; HARTMAN, G.L.; GHABRIAL, S.A.; KORBAN, S.S. Transcriptome analysis of resistance and susceptible genotypes of *Glycine tomentella* during *Phakopsora pachyrhizi* infection reveals novel rust resistance genes. **Theoretical Applied Genetics**, v.120, p.1315-1333, Maio 2010.

SPOEL, S.H.; JOHNSON, J.S.; DONG, X. Regulation of tradeoffs between plant defenses against pathogen with different lifestyle. **Proceedings of the National Academy of Science**, v.104, n.47, p. 18842-18847, Nov. 2007

STASWICK PE, SU W, HOWELL SH. Methyl jasmonate inhibition of root growth and induction of a leaf protein are decreased in an *Arabidopsis thaliana* mutant. **PNAS**. v. 89, n. 15, p. 6837– 6840, Ago.1998.

STRACKE, R.; WEBER, M.; WEISSHAAR. The R2R3-MYB gene family in *Arabidopsis thaliana*. **Plant Biology**, v. 4, p. 447-456. Out. 2001.

SUGIMOTO, K.; TAKEDA, S.; HIROSHIKA, H. MYB-related transcription factor NtMYB2 induced by wounding and elicitors is a regulator of the Tobacco

retrotransposon *Tto1* and Defense-related genes. **The Plant Cell**, v. 12, n. 12, p. 2511-2527. Dez. 2000.

TAIZ, L.; ZEIGER, E. **Fisiologia vegetal**. 3. ed. Porto Alegre: Artmed, p.719, 2004.

THOMMA, B.P.H.J.; EGGERMONT, K.; PENNINGCKX, I.A.M.A.; MAUCH-MANI, B.; VOGELSANG, R.; CAMMUE, B.P.A.; BROEKAERT, W.F. Separate jasmonate-dependent and salicylate-dependent defense-response pathways in *Arabidopsis* are essential for resistance to distinct microbial pathogens. **Plant Biology**, v. 95, p. 15107-15111, Dez. 1998.

TREMBLAY, A.; HOSSEINI, P.; ALKHAROUF, N.W.; LI, S.; MATTHEWSB.F. Transcriptoma analysis of a compatible response by *Glycine max* to *Phakopsora pachyrhizi* infection. **Plant Science**, v. 179, p. 183-193, Set. 2010.

TREUTTER, D. Significance of flavonoids in plant resistance: a review. **Environmental Chemistry Letters**, v. 4, p. 147-157. Ago. 2006.

TSUKAHARA, R.Y.; HIKISHIMA, M.; CANTERI. M.G. Relações entre o clima e o progresso da ferrugem asiática em duas microrregiões do Estado do Paraná. **Semina**. v. 29, n.1, p. 47-52, 2008.

USDA, 2013. Soybean: World Supply and Distribution. Disponível em > <http://www.fas.usda.gov/psdonline/psdreport.aspx?hidReportRetrievalName=BVS&hidReportRetrievalID=706&hidReportRetrievalTemplateID=8>. Acesso em 14 de Fevereiro de 2013.

USDA. World Agricultural Supply and Demand Estimates, n. 515, 8 Fev. 2013

VAILLEAU, F.; DANIEL, X.; TRONCHET, M.; MONTILLE, J.L.; TRIANTAPHYLIDES, C.; ROBY, D. A R2R3-MYB gene. *AtMYB30*, act as a positive regulator of gene of the hypersensitive cell death program in plants in response to pathogen attack. **Plant Biology**, v. 99, p. 10179-10184. Jul. 2002.

MORTEL, RECKNOR J.C.; GRAHAN, M.A.; NETTLETON, D.; DITTMAN, J.D.; NELSON, R.T.; GODOY, C.V.; ABDELNOOR, R.V.; ALMEIDA, A.M.R.; BAUM, T.J.; WHITHAM, S.A. Distinct biphasic mRNA changes in response to Asian soybean rust infection. **Molecular Plant Microbe Interaction**, v. 20, n. 8, p. 887–899, Ago. 2007.

VAKHRUSHEVA, O.A.; NEDOSPASOV, S.A. System of innate immunity in plants. **Molecular Biology**, v. 45, p. 16-23, Fev. 2011.

VIJAYAN, P.; SHOCKEY, J.; LÉVESQUE. C. A.; COOK, R. J.; BROWSE, J. A role for jasmonate in pathogen defense of *Arabidopsis*. **Plant Biology**, v. 95, n. 12, p. 7209-7214, Jun.1998.

WANG, Z.; LIBAUT, M.; JOSHI, T.; VALLIYODAN, B.; NGUYEN, H. T.; XU, D.; STACEY, G.; CHENG, J. SoyDB: a knowledge database of soybean transcription factors. **BMC: Plant Biology**, v. 10, p. 1-12, Jan. 2010.

XU, L.; ZHU, L.; TU, L.; LIU, L.; YUAN, D.; JIN, L.; LONG, L.;ZHANG, X. Lignin metabolism has a central role in the resistance of cotton to the wilt fungus *Verticillium*

dahlia as revealed by RNA-Seq-dependent transcriptional analysis and histochemistry. **Journal of Experimental Botany**, v. 62, p. 5607-5621. Nov. 2011.

ZAMBALA, G.; ZOU, J.; TUJELA, J.; GONZALEZ, D.O.; CLOUGH, S. J.; VODKIN, L. O. Transcriptome changes in the phenylpropanoid pathway of *Glycine max* in response to *Pseudomonas syringae* infection. **BMC Plant Biology**, v. 6, p. 1-18. Nov. 2006.

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

ZHANG, B.; RAMONELL, K.; SOMERVILLER, S.; STACEY, G. Characterization of early, chitin-induced gene expression in *Arabidopsis*. **Molecular Plant Microbe Interaction**, v. 15, n. 9, p. 963-970. Set. 2002

ZHANG, C.; YANG, C.; WHITHAM SA.; HILL, H.H. Development and Use of an Efficient DNA-Based Viral Gene Silencing Vector for Soybean. **Molecular Plant Microbe Interaction**, v.22, n. 2, p. 123–131, Fev. 2009.

ZHONG, R.; YE, Z-H. The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in *Arabidopsis*. **The Plant Cell**, v. 19, n. 9, p. 2776-2792. Set. 2007.

ZHONG, R.; YE, Z-H. Transcriptional regulation of lignin biosynthesis. **Plant Signaling & Behavior**, v. 4, p. 1028-1034. Nov. 2009.

ZHOU, J.; LEE, C.; ZHONG, R.; YE, Z.H. MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. **The Plant Cell**, v. 21, p. 248-266. Jan. 2009

ZHOU, B.; JIA, Z.; TIAN, S.; WANG, X.; GOU, Z.; LÜ, B.; DONG, H. AtMYB44 positively modulates disease resistance to *Pseudomonas syringae* through the salicylic acid signaling pathway in *Arabidopsis*. **Functional Plant Biology**, v. 40, p. 304-313. Nov. 2013.

ANEXOS

ANEXO A

GmMYBs com domínio R2R3, identificados através de análise das sequências proteicas nas ferramentas SMART, PFAM e MEME:

GENE ID	ANÁLISE DE DOMÍNIO			
	SMART	PFAM	MEME	R2R3 COMPLETO ¹
Glyma0041s00310.1	✓		✓	+
Glyma0076s00200.1				
Glyma01g00810.1				
Glyma01g02070.1	✓	✓	✓	+
Glyma01g05190.1	✓			
Glyma01g05270.1				
Glyma01g05980.1	✓		✓	+
Glyma01g06190.1				
Glyma01g06220.1	✓	✓	✓	+
Glyma01g09280.1	✓	✓	✓	+
Glyma01g20440.1				
Glyma01g21900.1				
Glyma01g23460.1				
Glyma01g24780.1				
Glyma01g26650.1		✓	✓	+
Glyma01g27720.1				
Glyma01g28790.1				
Glyma01g39740.1	✓	✓	✓	+
Glyma01g40410.1	✓		✓	+
Glyma01g41610.1	✓	✓	✓	+
Glyma01g42050.1	✓	✓	✓	+
Glyma01g42650.1	✓	✓	✓	+
Glyma01g43120.1	✓	✓	✓	+
Glyma01g44370.1	✓	✓	✓	+
Glyma02g00820.1	✓	✓	✓	+
Glyma02g00960.1	✓	✓	✓	+
Glyma02g01300.1	✓	✓	✓	+
Glyma02g01740.1	✓	✓	✓	+
Glyma02g02310.1	✓			
Glyma02g12100.1	✓		✓	+
Glyma02g12240.1	✓	✓	✓	+
Glyma02g12250.1	✓	✓	✓	+
Glyma02g12260.1	✓	✓	✓	+
Glyma02g13770.1	✓	✓	✓	+
Glyma02g26540.1				
Glyma02g26980.1				
Glyma02g30820.1				
Glyma02g30920.1				
Glyma02g34010.1				
Glyma02g39000.1				
Glyma02g39070.1	✓		✓	+
Glyma02g40930.1				
Glyma02g41180.1	✓	✓	✓	+
Glyma02g41260.1				
Glyma02g41440.1	✓	✓	✓	+
Glyma02g42030.1	✓	✓	✓	+
Glyma02g42140.1				
Glyma02g42140.2				
Glyma02g43280.1	✓	✓	✓	+
Glyma02g44190.1				
Glyma03g00890.1	✓	✓	✓	+
Glyma03g00980.1	✓		✓	+
Glyma03g01540.1	✓	✓	✓	+

ANEXO A

(continuação)

GENE ID	ANÁLISE DE DOMÍNIO			
	SMART	PFAM	MEME	R2R3 COMPLETO ¹
Glyma03g03660.1				
Glyma03g06230.1				
Glyma03g07840.1				
Glyma03g08710.1				
Glyma03g13550.1				
Glyma03g14440.1				
Glyma03g15810.1	✓		✓	+
Glyma03g15870.1	✓		✓	+
Glyma03g15930.1	✓		✓	+
Glyma03g19030.1			✓	+
Glyma03g19470.1	✓	✓	✓	+
Glyma03g22590.1				
Glyma03g26210.1				
Glyma03g26830.1				
Glyma03g31980.1		✓	✓	+
Glyma03g34110.1			✓	+
Glyma03g37640.1	✓	✓	✓	+
Glyma03g38040.1	✓	✓	✓	+
Glyma03g38070.1				
Glyma03g38410.1	✓	✓	✓	+
Glyma03g38550.1				
Glyma03g38660.1	✓	✓	✓	+
Glyma03g41100.1	✓	✓	✓	+
Glyma04g00550.1	✓	✓	✓	+
Glyma04g00550.2				
Glyma04g03910.1	✓	✓	✓	+
Glyma04g04490.1	✓		✓	+
Glyma04g04850.1				
Glyma04g05170.1			✓	+
Glyma04g08550.1	✓	✓	✓	+
Glyma04g11040.1				
Glyma04g15150.1	✓	✓	✓	+
Glyma04g26650.1	✓	✓	✓	+
Glyma04g33210.1	✓	✓	✓	+
Glyma04g33720.1	✓	✓	✓	+
Glyma04g34630.1				
Glyma04g35720.1				
Glyma04g36110.1	✓	✓	✓	+
Glyma04g38240.1	✓	✓	✓	+
Glyma04g42110.1	✓		✓	+
Glyma05g01080.1	✓	✓	✓	+
Glyma05g02170.1				
Glyma05g02300.1	✓		✓	+
Glyma05g02550.1	✓	✓	✓	+
Glyma05g03780.1	✓	✓	✓	+
Glyma05g04900.1	✓	✓	✓	+
Glyma05g06410.1		✓	✓	+
Glyma05g07980.1				
Glyma05g08690.1	✓	✓	✓	+
Glyma05g08760.1		✓		
Glyma05g15860.1				
Glyma05g18140.1				
Glyma05g18820.1				
Glyma05g21220.1	✓	✓	✓	+
Glyma05g23080.1	✓		✓	+
Glyma05g31400.1				
Glyma05g32270.1				
Glyma05g33210.1			✓	+
Glyma05g35050.1	✓	✓	✓	+
Glyma05g36120.1				
Glyma05g37460.1	✓	✓	✓	+
Glyma05g37570.1				
Glyma05g37570.2				
Glyma06g00630.1	✓	✓	✓	+
Glyma06g00630.2				

ANEXO A

(continuação)

GENE ID	ANÁLISE DE DOMÍNIO			
	SMART	PFAM	MEME	R2R3 COMPLETO ¹
Glyma06g04010.1	✓	✓	✓	+
Glyma06g04950.1				
Glyma06g05260.1	✓		✓	+
Glyma06g08660.1	✓	✓	✓	+
Glyma06g10840.1	✓	✓	✓	+
Glyma06g12690.1	✓		✓	+
Glyma06g16820.1	✓	✓	✓	+
Glyma06g18830.1	✓	✓	✓	+
Glyma06g19280.1	✓		✓	+
Glyma06g20020.1				
Glyma06g20800.1	✓	✓	✓	+
Glyma06g21040.1	✓	✓	✓	+
Glyma06g22680.1				
Glyma06g38340.1	✓	✓	✓	+
Glyma06g44120.1				
Glyma06g45460.1	✓	✓	✓	+
Glyma06g45520.1				
Glyma06g45530.1				
Glyma06g45540.1		✓	✓	+
Glyma06g45550.1		✓	✓	+
Glyma06g45560.1				
Glyma06g45570.1	✓	✓		+
Glyma06g46590.1				
Glyma06g47000.1	✓	✓	✓	+
Glyma07g01050.1	✓	✓	✓	+
Glyma07g04210.1	✓	✓	✓	+
Glyma07g04240.1	✓	✓	✓	+
Glyma07g05960.1	✓	✓	✓	+
Glyma07g07960.1	✓	✓	✓	+
Glyma07g08080.1				
Glyma07g10320.1	✓	✓	✓	+
Glyma07g11330.1	✓	✓		+
Glyma07g11330.2	✓	✓		+
Glyma07g14480.1	✓		✓	+
Glyma07g15250.1				
Glyma07g15820.1	✓		✓	+
Glyma07g15820.2	✓		✓	+
Glyma07g15820.3	✓	✓	✓	+
Glyma07g15850.1			✓	+
Glyma07g16980.1	✓	✓	✓	+
Glyma07g17660.1				
Glyma07g28310.1				
Glyma07g28590.1				
Glyma07g29680.1				
Glyma07g30860.1	✓	✓	✓	+
Glyma07g33960.1	✓	✓	✓	+
Glyma07g34030.1				
Glyma07g34030.2				
Glyma07g35560.1	✓	✓	✓	+
Glyma07g35580.1				
Glyma07g36430.1	✓	✓	✓	+
Glyma07g37140.1	✓	✓	✓	+
Glyma08g00810.1	✓	✓	✓	+
Glyma08g02080.1	✓	✓	✓	+
Glyma08g03530.1				
Glyma08g04670.1	✓	✓	✓	+
Glyma08g06440.1	✓	✓	✓	+
Glyma08g17370.1	✓	✓	✓	+
Glyma08g17860.1	✓	✓	✓	+
Glyma08g20150.1				
Glyma08g20150.2				
Glyma08g20440.1	✓	✓	✓	+
Glyma08g26890.2				
Glyma08g27660.1	✓	✓	✓	+
Glyma08g27660.2				

ANEXO A

(continuação)

GENE ID	ANÁLISE DE DOMÍNIO			
	SMART	PFAM	MEME	R2R3 COMPLETO ¹
Glyma08g40020.1				
Glyma08g40950.1	✓			
Glyma08g42920.1				
Glyma08g42960.1	✓	✓	✓	+
Glyma08g43000.1				
Glyma08g44350.1				
Glyma08g44950.1	✓	✓	✓	+
Glyma09g00370.1	✓	✓	✓	+
Glyma09g03690.1	✓	✓	✓	+
Glyma09g04370.1	✓	✓	✓	+
Glyma09g12170.1				
Glyma09g12230.1				
Glyma09g25460.1				
Glyma09g25590.1	✓	✓	✓	+
Glyma09g29240.1				
Glyma09g29940.1			✓	+
Glyma09g30900.1		✓	✓	+
Glyma09g31570.1	✓	✓	✓	+
Glyma09g33870.1	✓	✓	✓	+
Glyma09g36970.1	✓	✓	✓	+
Glyma09g36980.1				
Glyma09g36990.1	✓	✓	✓	+
Glyma09g37010.1	✓		✓	+
Glyma09g37040.1	✓	✓	✓	+
Glyma09g37340.1	✓	✓	✓	+
Glyma09g39720.1	✓	✓	✓	+
Glyma10g00930.1	✓	✓	✓	+
Glyma10g01330.1	✓	✓	✓	+
Glyma10g01340.1	✓	✓	✓	+
Glyma10g01800.1				
Glyma10g01810.1				
Glyma10g02110.1				
Glyma10g04250.1				
Glyma10g04660.1				
Glyma10g06190.1	✓		✓	+
Glyma10g06680.1	✓	✓	✓	+
Glyma10g06930.1				
Glyma10g14830.1				
Glyma10g22770.1				
Glyma10g24130.1				
Glyma10g26680.1	✓	✓	✓	+
Glyma10g27940.1	✓	✓	✓	+
Glyma10g28250.1	✓	✓	✓	+
Glyma10g30860.1	✓	✓	✓	+
Glyma10g30870.1	✓			
Glyma10g32410.1	✓	✓	✓	+
Glyma10g33450.1	✓	✓	✓	+
Glyma10g35050.1	✓	✓	✓	+
Glyma10g35060.1				
Glyma10g35670.1				
Glyma10g38090.1	✓	✓	✓	+
Glyma10g38110.1	✓	✓	✓	+
Glyma10g41930.1	✓	✓	✓	+
Glyma11g01150.1	✓	✓	✓	+
Glyma11g02400.1	✓	✓	✓	+
Glyma11g02580.3				
Glyma11g02780.1				
Glyma11g03300.1	✓	✓	✓	+
Glyma11g03770.1	✓	✓	✓	+
Glyma11g04880.1				
Glyma11g05550.1	✓		✓	+
Glyma11g08760.1				
Glyma11g11450.1	✓	✓	✓	+
Glyma11g11570.1	✓	✓	✓	+
Glyma11g14200.1		✓	✓	+

ANEXO A

(continuação)

GENE ID	ANÁLISE DE DOMÍNIO			
	SMART	PFAM	MEME	R2R3 COMPLETO ¹
Glyma11g15180.1	✓	✓	✓	+
Glyma11g15180.2				
Glyma11g15180.3				
Glyma11g19980.1	✓		✓	+
Glyma11g22960.1				
Glyma11g27600.1				
Glyma11g33620.1	✓	✓	✓	+
Glyma11g34930.1				
Glyma11g34930.2				
Glyma11g34930.3				
Glyma11g34930.4				
Glyma11g38210.1				
Glyma12g01960.1	✓	✓	✓	+
Glyma12g03600.1	✓	✓	✓	+
Glyma12g06180.1		✓	✓	+
Glyma12g07110.1				
Glyma12g07110.2				
Glyma12g08480.1	✓		✓	+
Glyma12g10790.1				
Glyma12g11330.1	✓	✓		
Glyma12g11340.1		✓	✓	+
Glyma12g11390.1	✓	✓	✓	+
Glyma12g11490.1	✓	✓		
Glyma12g11600.1				
Glyma12g12990.1				
Glyma12g15290.1				
Glyma12g17020.1				
Glyma12g29030.1				
Glyma12g30140.1	✓		✓	+
Glyma12g31950.1	✓	✓	✓	+
Glyma12g32130.1				
Glyma12g32130.2				
Glyma12g32530.1	✓	✓		
Glyma12g32540.1				
Glyma12g32610.1	✓	✓	✓	+
Glyma12g34650.1	✓	✓	✓	+
Glyma12g36630.1	✓	✓	✓	+
Glyma12g37030.1	✓	✓	✓	+
Glyma13g01200.1	✓		✓	+
Glyma13g03740.1				
Glyma13g04030.1	✓	✓	✓	+
Glyma13g04920.1	✓	✓	✓	+
Glyma13g05370.1	✓	✓	✓	+
Glyma13g05550.1	✓	✓	✓	+
Glyma13g07020.1				
Glyma13g09010.1	✓	✓	✓	+
Glyma13g09090.1				
Glyma13g09980.1	✓	✓	✓	+
Glyma13g16890.1	✓	✓	✓	+
Glyma13g20510.1	✓		✓	+
Glyma13g20760.1				
Glyma13g20880.1	✓	✓	✓	+
Glyma13g21370.1				
Glyma13g22590.1				
Glyma13g25720.1				
Glyma13g27310.1	✓	✓	✓	+
Glyma13g29890.1				
Glyma13g32090.1	✓	✓	✓	+
Glyma13g35810.1	✓	✓	✓	+
Glyma13g36370.1				
Glyma13g36380.1				
Glyma13g37820.1	✓	✓	✓	+
Glyma13g37900.1				
Glyma13g37910.1				
Glyma13g37920.1				
Glyma13g38520.1		✓		

ANEXO A

(continuação)

GENE ID	ANÁLISE DE DOMÍNIO			
	SMART	PFAM	MEME	R2R3 COMPLETO ¹
Glyma13g39760.1	✓		✓	+
Glyma13g41470.1				
Glyma13g42430.1	✓	✓	✓	+
Glyma14g04370.1				
Glyma14g04590.1				
Glyma14g06320.1	✓	✓	✓	+
Glyma14g06750.1				
Glyma14g06870.1	✓	✓	✓	+
Glyma14g07510.1				
Glyma14g08090.2				
Glyma14g09540.1	✓		✓	+
Glyma14g10340.1	✓		✓	+
Glyma14g10480.1				
Glyma14g21490.1				
Glyma14g24500.1	✓	✓	✓	+
Glyma14g27020.1				
Glyma14g27260.1				
Glyma14g35490.1				
Glyma14g37050.1	✓			
Glyma14g37140.1	✓		✓	+
Glyma14g39530.1	✓	✓	✓	+
Glyma15g02950.1	✓	✓	✓	+
Glyma15g03920.1	✓	✓	✓	+
Glyma15g04620.1	✓	✓	✓	+
Glyma15g04620.2				
Glyma15g04620.3				
Glyma15g04620.4				
Glyma15g07230.1	✓	✓	✓	+
Glyma15g14190.1	✓	✓	✓	+
Glyma15g14620.1	✓	✓	✓	+
Glyma15g14620.2				
Glyma15g15400.1	✓	✓	✓	+
Glyma15g17710.1				
Glyma15g19350.1				
Glyma15g19360.1		✓	✓	+
Glyma15g19360.2		✓	✓	+
Glyma15g19930.1				
Glyma15g20630.1				
Glyma15g33040.1				
Glyma15g33700.1				
Glyma15g35670.1				
Glyma15g35860.1	✓	✓	✓	+
Glyma15g41250.1	✓	✓	✓	+
Glyma15g41810.1			✓	+
Glyma16g00920.1	✓	✓	✓	+
Glyma16g00930.1				
Glyma16g02570.1	✓	✓	✓	+
Glyma16g06900.1		✓	✓	+
Glyma16g07930.1				
Glyma16g07960.1	✓	✓	✓	+
Glyma16g13440.1	✓	✓	✓	+
Glyma16g16270.1				
Glyma16g16800.1				
Glyma16g31280.1	✓	✓	✓	+
Glyma16g31280.2				
Glyma16g32100.1				
Glyma16g34490.1			✓	+
Glyma17g03480.1		✓	✓	+
Glyma17g04170.1		✓	✓	+
Glyma17g05830.1		✓	✓	+
Glyma17g07330.1	✓		✓	+
Glyma17g09310.1				
Glyma17g09640.1	✓	✓	✓	+
Glyma17g10820.1	✓	✓	✓	+
Glyma17g12820.1				
Glyma17g13010.1				

ANEXO A

(continuação)

GENE ID	ANÁLISE DE DOMÍNIO			
	SMART	PFAM	MEME	R2R3 COMPLETO ¹
Glyma17g14290.1	✓	✓	✓	+
Glyma17g14290.2	✓	✓	✓	+
Glyma17g15270.1	✓	✓	✓	+
Glyma17g16980.1	✓		✓	+
Glyma17g17560.1	✓	✓	✓	+
Glyma17g22860.1				
Glyma17g26240.1	✓	✓	✓	+
Glyma17g35020.1				
Glyma17g35620.1	✓	✓	✓	+
Glyma17g36370.1	✓	✓	✓	+
Glyma18g03430.1				
Glyma18g04580.1	✓	✓	✓	+
Glyma18g07250.1				
Glyma18g07360.1	✓		✓	+
Glyma18g07960.1	✓	✓	✓	+
Glyma18g10920.1	✓	✓	✓	+
Glyma18g16040.1	✓			
Glyma18g18140.1				
Glyma18g26600.1			✓	+
Glyma18g32460.1				
Glyma18g37640.1				
Glyma18g39740.1	✓	✓	✓	+
Glyma18g39760.1			✓	+
Glyma18g39760.2			✓	+
Glyma18g40790.1				
Glyma18g41520.1	✓	✓	✓	+
Glyma18g42530.1				
Glyma18g46480.1	✓	✓	✓	+
Glyma18g49360.1	✓	✓	✓	+
Glyma18g49630.1	✓	✓	✓	+
Glyma18g49670.1			✓	+
Glyma18g49690.1	✓	✓	✓	+
Glyma18g50880.1				
Glyma18g50890.1	✓	✓	✓	+
Glyma19g00930.1		✓	✓	+
Glyma19g02090.1	✓	✓	✓	+
Glyma19g02600.1	✓	✓	✓	+
Glyma19g02890.1	✓	✓	✓	+
Glyma19g02980.1	✓	✓	✓	+
Glyma19g05080.1	✓	✓	✓	+
Glyma19g07830.1		✓	✓	+
Glyma19g13990.1				
Glyma19g14230.1	✓	✓	✓	+
Glyma19g14270.1	✓	✓	✓	+
Glyma19g24450.1				
Glyma19g24530.1				
Glyma19g24770.1				
Glyma19g25790.1				
Glyma19g27750.1				
Glyma19g29670.1	✓		✓	+
Glyma19g29750.1	✓	✓	✓	+
Glyma19g34740.1	✓	✓	✓	+
Glyma19g36830.1	✓		✓	+
Glyma19g40250.1	✓	✓	✓	+
Glyma19g40650.1				
Glyma19g40670.1				
Glyma19g41010.1	✓	✓	✓	+
Glyma19g41150.1				
Glyma19g41250.1	✓	✓	✓	+
Glyma19g43740.1	✓	✓	✓	+
Glyma19g44660.1	✓	✓	✓	+
Glyma20g01450.1				
Glyma20g01610.1	✓	✓	✓	+
Glyma20g03940.1				
Glyma20g04240.1	✓	✓	✓	+
Glyma20g04510.1				

ANEXO A

(continuação)

GENE ID	ANÁLISE DE DOMÍNIO			
	SMART	PFAM	MEME	R2R3 COMPLETO ¹
Glyma20g08730.1				
Glyma20g09140.1				
Glyma20g11040.1	✓	✓	✓	+
Glyma20g11110.1				
Glyma20g12290.1				
Glyma20g20980.1	✓	✓	✓	+
Glyma20g21670.1				
Glyma20g21680.1				
Glyma20g22230.1	✓	✓	✓	+
Glyma20g24600.1				
Glyma20g25110.1	✓	✓	✓	+
Glyma20g29710.1	✓	✓	✓	+
Glyma20g29730.1	✓	✓	✓	+
Glyma20g32500.1	✓	✓	✓	+
Glyma20g32510.1	✓	✓	✓	+
Glyma20g34140.1	✓	✓	✓	+
Glyma20g35180.1	✓	✓	✓	+
Glyma20g36600.1	✓			
Glyma20g36600.2	✓			
Gnl ug gma#s342734 66				
Gnl ug gma#s393129 23	✓	✓	✓	+
Gnl ug gma#s530865 90;put-169a- glycine_max-63505	✓		✓	+
Glyma14g21490.1				
Put-169a-glycine_max- 102940	✓	✓	✓	+
Put-169a-glycine_max- 179128;gnl ug gma#s 53090519	✓	✓	✓	+
Put-169a-glycine_max- 81255;gnl ug gma#s4 8315927			✓	+

¹ – Foi utilizado como critério de identificação de sequências R2R3, aquelas que apresentaram o domínio identificado em ao menos duas ferramentas, com $e\text{-value} \leq 1e-10$, ou por inspeção manual direta. Sequências excluídas por não apresentarem o domínio característico de MYB estão sombreadas em cinza

ANEXO B

GmMYBs induzidos após a infecção por *P. pachyrhizi* em soja

GENE ID	R2R3 COMPLETO	SOYBASE	ANÁLISE DE EXPRESSÃO				OUTROS	EXPRESSÃO CONFIRMADA
			GENOSOJA - LGE		Genevestigator (Fold-chage)			
			SUPERSAGE (Fold-change)	SSH (FPKM)	REAÇÃO COMPATÍVEL	REAÇÃO INCOMPATÍVEL		
			PI561356 MOCKXINOC 12-24-48HAI	PI561356 MOCKXINOC 12-192HAI	PI462312 X TAIWAN80-2	PI462312 X HAWAII94-1		
Glyma0041s00310.1	+	✓					+	
Glyma0076s00200.1		✓					+	
Glyma01g00810.1		✓					+	
Glyma01g02070.1*	+				(f) 1.92		+	
Glyma01g05190.1					(a) 1.18		+	
Glyma01g05270.1		✓					+	
Glyma01g05980.1	+	✓					+	
Glyma01g06190.1		✓					+	
Glyma01g06220.1	+	✓					+	
Glyma01g09280.1	+							
Glyma01g20440.1		✓					+	
Glyma01g21900.1		✓					+	
Glyma01g23460.1		✓					+	
Glyma01g24780.1		✓					+	
Glyma01g26650.1*	+					(3C) 2.14	+	
Glyma01g27720.1						(3B) 1.30	+	
Glyma01g28790.1		✓					+	
Glyma01g39740.1	+	✓					+	
Glyma01g40410.1	+	✓					+	
Glyma01g41610.1	+	✓					+	
Glyma01g42050.1	+	✓					+	
Glyma01g42650.1	+	✓					+	
Glyma01g43120.1	+	✓					+	
Glyma01g44370.1	+							
Glyma02g00820.1*	+				(a) 20.70	(a) 32.02	+	
Glyma02g00960.1	+							
Glyma02g01300.1*	+				(a) 6.21	(a) 5.97	+	
Glyma02g01740.1*	+						(2C) 3.19/(3C) 2.92	
Glyma02g02310.1		✓					+	
Glyma02g12100.1	+	✓					+	
Glyma02g12240.1	+	✓					+	
Glyma02g12250.1	+	✓					+	
Glyma02g12260.1	+	✓					+	
Glyma02g13770.1	+	✓					+	
Glyma02g26540.1								
Glyma02g26980.1								
Glyma02g30820.1								
Glyma02g30920.1								
Glyma02g34010.1							+	
Glyma02g39000.1		✓					+	
Glyma02g39070.1	+						+	
Glyma02g40930.1		✓					+	
Glyma02g41180.1	+	✓					+	
Glyma02g41260.1						(2B) 2.61 (2C) 2.84	+	
Glyma02g41440.1	+	✓					+	
Glyma02g42030.1	+	✓					+	
Glyma02g42140.1		✓					+	
Glyma02g42140.2						(b) 1.52	+	

Glyma02g43280.1*	+					(2A) 1.39	+
Glyma02g44190.1						(3C) 1.16	+
Glyma03g00980.1*	+					(3A) 1.16	+
Glyma03g01540.1	+	✓					+
Glyma03g03660.1							
Glyma03g06230.1							
Glyma03g07840.1							
Glyma03g08710.1							
Glyma03g13550.1							
Glyma03g14440.1						3.62 (2C)	+
Glyma03g15810.1	+	✓					+
Glyma03g15870.1	+	✓					+
Glyma03g15930.1	+	✓					+
Glyma03g19030.1*	+					(3C) 1.33	+
Glyma03g19470.1	+	✓					+
Glyma03g22590.1							
Glyma03g26210.1				(e) 1.18			+
Glyma03g26830.1							
Glyma03g31980.1	+	✓					+
Glyma03g34110.1*	+					(2A) 5.99/(3A) 4.51	+
Glyma03g37640.1*	+					(3C) 1.20	+
Glyma03g38040.1*	+				(a)2.76(f) 1.96		+
Glyma03g38070.1		✓					+
Glyma03g38410.1	+					(3B) 1.55	+
Glyma03g38550.1			1.76	(A) 11.16			+
Glyma03g38660.1*	+					(5A) 1.15	+
Glyma03g41100.1*	+				(a) 1.91		+
Glyma04g00550.1	+	✓			(f) -1.69		+
Glyma04g00550.2							+
Glyma04g03910.1*	+					(a) 4.36	+
Glyma04g04490.1*	+					(f)5.23(a) 3.00	+
Glyma04g04850.1		✓					+
Glyma04g05170.1	+	✓					+
Glyma04g08550.1*	+					(4C) 1.13	+
Glyma04g11040.1		✓					+
Glyma04g15150.1	+	✓					+
Glyma04g26650.1	+						
Glyma04g33210.1*	+					(2C) 2.77	+
Glyma04g33720.1*	+					(5B) 1.26	+
Glyma04g34630.1							
Glyma04g35720.1							
Glyma04g36110.1	+	✓					+
Glyma04g38240.1*	+				(e) 1.42		+
Glyma04g42110.1	+	✓					+
Glyma05g01080.1	+	✓					+
Glyma05g02170.1							
Glyma05g02300.1	+						
Glyma05g02550.1*	+				(f) 1.10		+
Glyma05g03780.1	+	✓					+
Glyma05g04900.1	+	✓					+
Glyma05g06410.1*	+				(f) 1.12	(4B) -1.11	+
Glyma05g07980.1		✓					+
Glyma05g08690.1*	+					(6)1.12(3C)1.14	+

ANEXO B

GmMYBs induzidos após a infecção por *P. pachyrhizi* em soja (continuação)

ANÁLISE DE EXPRESSÃO								
GENE ID	R2R3 COMPLETO	SOYBASE	GENOSOJA - LGE		Genevestigador (Fold-chage)		OUTROS	EXPRESSÃO CONFIRMADA
			SUPERSAGE (Fold-change)	SSH (FPKM)	REAÇÃO COMPATÍVEL	REAÇÃO INCOMPATÍVEL		
			PI561356 <i>MOCKXINOC</i>	PI561356 <i>MOCKXINOC</i>	PI462312 X TAIWAN80-2	PI462312 X HAWAII94-1		
			12-24-48HAI	12-192HAI				
Glyma05g08760.1		✓						+
Glyma05g15860.1								
Glyma05g18140.1								
Glyma05g18820.1								
Glyma05g21220.1	+	✓						+
Glyma05g23080.1	+	✓						+
Glyma05g31400.1						(f) 20.14		+
Glyma05g32270.1							(2B) 3.36	+
Glyma05g33210.1	+							+
Glyma05g35050.1*	+					(d) 4.62		+
Glyma05g36120.1		✓						+
Glyma05g37460.1	+	✓						+
Glyma05g37570.1			5.051					+
Glyma05g37570.2			5.051					+
Glyma06g00630.1	+						(2B) -4.78	+
Glyma06g00630.2							(2B) -4.40	+
Glyma06g04010.1*	+					(a) 2.12		+
Glyma06g04950.1							(2A) -1.28	+
Glyma06g05260.1*	+					(A) 3.04		+
Glyma06g08660.1	+	✓						+
Glyma06g10840.1	+	✓						+
Glyma06g12690.1*	+					(f) 2.03		+
Glyma06g16820.1*	+						(3C) 3.67	+
Glyma06g18830.1	+					(f) 1.23		+
Glyma06g19280.1	+							
Glyma06g20020.1								
Glyma06g20800.1*	+						(3C) 2.67	+
Glyma06g21040.1	+	✓						+
Glyma06g22680.1								
Glyma06g38340.1	+							
Glyma06g44120.1								
Glyma06g45460.1	+							
Glyma06g45520.1*						3.23 (c)		+
Glyma06g45530.1								
Glyma06g45540.1*	+						(2C) 6.86	+
Glyma06g45550.1	+	✓						+
Glyma06g45560.1								
Glyma06g45570.1	+							
Glyma06g46590.1		✓						+
Glyma06g47000.1*	+						(b) 1.68	+
Glyma07g01050.1	+	✓						+
Glyma07g04210.1*	+						(c) 1.14	+
Glyma07g04240.1	+							
Glyma07g05960.1*	+					(a) 7.69	(a) 14.34	+
Glyma07g07960.1	+	✓						+
Glyma07g08080.1								
Glyma07g10320.1	+	✓						+

ANEXO B

GmMYBs induzidos após a infecção por *P. pachyrhizi* em soja (continuação)

GENE ID	R2R3 COMPLETO	SOYBASE	ANÁLISE DE EXPRESSÃO				OUTROS	EXPRESSÃO CONFIRMADA
			GENOSOJA - LGE		Genevestigator (Fold-chage)			
			SUPERSAGE (Fold-change)	SSH (FPKM)	REAÇÃO COMPATÍVEL	REAÇÃO INCOMPATÍVEL		
			PI561356 <i>MOCKXINOC</i> 12-24-48HAI	PI561356 <i>MOCKXINOC</i> 12-192HAI	PI462312 X TAIWAN80-2	PI462312 X HAWAII94-1		
Glyma09g36990.1	+							
Glyma09g37010.1	+							
Glyma09g37040.1	+				(f) -9.47		+	
Glyma09g37340.1	+				(a) -1.58		+	
Glyma09g39720.1	+	✓					+	
Glyma10g00930.1	+							
Glyma10g01330.1*	+				(d) 7.25 (e) 10.46(f) 4.94		+	
Glyma10g01340.1*	+				(a) 2.09	(3B) 2.44	+	
Glyma10g01800.1		✓					+	
Glyma10g01810.1						(2A) 4.14	+	
Glyma10g02110.1								
Glyma10g04250.1								
Glyma10g04660.1		✓					+	
Glyma10g06190.1	+	✓					+	
Glyma10g06680.1	+							
Glyma10g06930.1								
Glyma10g14830.1								
Glyma10g22770.1								
Glyma10g24130.1								
Glyma10g26680.1	+							
Glyma10g27940.1	+	✓					+	
Glyma10g28250.1*	+				(e) -3.71		+	
Glyma10g30860.1	+							
Glyma10g30870.1	+					(2B) 1.53	+	
Glyma10g32410.1	+	✓					+	
Glyma10g33450.1	+							
Glyma10g33500.1	+							
Glyma10g35060.1		✓					+	
Glyma10g35670.1		✓					+	
Glyma10g38090.1		✓					+	
Glyma10g38110.1	+	✓					+	
Glyma10g41930.1*	+				(f) 7.50		+	
Glyma11g01150.1	+	✓					+	
Glyma11g02400.1	+					(3B) -1.63	+	
Glyma11g02580.3		✓					+	
Glyma11g02780.1								
Glyma11g03300.1	+				(f) -2.99		+	
Glyma11g03770.1	+							
Glyma11g04880.1						(5A) 1.07	+	
Glyma11g05550.1*	+		5.51			(2C) -4.14 (3A) 1.29	+	
Glyma11g08760.1						(3A) 1.86	+	
Glyma11g11450.1*	+					(3A) 3.28	+	
Glyma11g11570.1	+							
Glyma11g14200.1	+	✓					+	
Glyma11g15180.1	+	✓					+	
Glyma11g15180.2		✓					+	
Glyma11g15180.3		✓					+	

Glyma11g19980.1*	+	✓			(f) 1.62	+
Glyma11g22960.1						+
			<u>12-24-48HAI</u>	<u>12-192HAI</u>		
Glyma11g27600.1						
Glyma11g33620.1*	+	✓			(2B) 2.98	+
Glyma11g34930.1		✓				+
Glyma11g34930.2		✓				+
Glyma11g34930.3		✓				+
Glyma11g34930.4		✓			(b) 1.88	+
Glyma11g38210.1		✓				+
Glyma12g01960.1*	+				(4A) 1.07	+
Glyma12g03600.1*	+		4.59			+
Glyma12g06180.1	+	✓				+
Glyma12g07110.1		✓				+
Glyma12g07110.2		✓			(a) -4.95	+
Glyma12g08480.1	+	✓				+
Glyma12g10790.1		✓				+
Glyma12g11330.1		✓				+
Glyma12g11340.1	+	✓				+
Glyma12g11390.1*	+			(A) 5.94(B) 26.49(C)7,94		
Glyma12g11490.1		✓				+
Glyma12g11600.1						
Glyma12g12990.1						
Glyma12g15290.1						
Glyma12g17020.1						
Glyma12g29030.1						+
Glyma12g30140.1	+	✓				+
Glyma12g31950.1	+	✓				+
Glyma12g32130.1		✓				+
Glyma12g32130.2		✓				+
Glyma12g32530.1						+
Glyma12g32540.1						+
Glyma12g32610.1	+					
Glyma12g34650.1	+	✓				+
Glyma12g36630.1	+	✓				+
Glyma12g37030.1	+					
Glyma13g01200.1*	+				(d) 1.13	+
Glyma13g03740.1		✓				+
Glyma13g04030.1*	+				(e) 1.49	+
Glyma13g04920.1	+	✓				+
Glyma13g05370.1	+	✓				+
Glyma13g05550.1	+	✓				+
Glyma13g07020.1		✓				+
Glyma13g09010.1*	+				(5B) 1.08	+
Glyma13g09090.1						
Glyma13g09980.1	+	✓				+
Glyma13g16890.1	+			(a) -1.24		+
Glyma13g20510.1*	+				(a) 6.49	+
Glyma13g20760.1		✓				+
Glyma13g20880.1	+					
Glyma13g21370.1		✓				+
Glyma13g22590.1						

ANEXO B

GmMYBs induzidos após a infecção por *P. pachyrhizi* em soja (continuação)

GENE ID	R2R3 COMPLETO	SOYBASE	ANÁLISE DE EXPRESSÃO				OUTROS	EXPRESSIONO CONFIRMADA
			GENOSOJA - LGE		Genevestigator (Fold-chage)			
			SUPERSAGE (Fold-change)	SSH (FPKM)	REAÇÃO COMPATÍVEL	REAÇÃO INCOMPATÍVEL		
			PI561356 MOCKKINOC 12-24-48HAI	PI561356 MOCKKINOC 12-192HAI	PI462312 X TAIWAN80-2	PI462312 X HAWAII94-1		
Glyma13g25720.1						(c) 1.42	+	
Glyma13g27310.1	+	✓					+	
Glyma13g29890.1		✓					+	
Glyma13g32090.1	+	✓					+	
Glyma13g35810.1	+	✓					+	
Glyma13g36370.1						(c) -1.13	+	
Glyma13g36380.1		✓					+	
Glyma13g37820.1	+							
Glyma13g37900.1		✓					+	
Glyma13g37910.1		✓					+	
Glyma13g37920.1		✓					+	
Glyma13g38520.1		✓					+	
Glyma13g39760.1	+							
Glyma13g41470.1								
Glyma13g42430.1	+	✓					+	
Glyma14g04370.1								
Glyma14g04590.1				(C) 2,34		(f) -2.92	+	
Glyma14g06320.1	+	✓					+	
Glyma14g06750.1			9.18				+	
Glyma14g06870.1	+	✓					+	
Glyma14g07510.1		✓					+	
Glyma14g08090.2		✓					+	
Glyma14g09540.1*	+					(f) 1.59	+	
Glyma14g10340.1	+	✓					+	
Glyma14g10480.1		✓					+	
Glyma14g21490.1						(3C) 1.81	+	
Glyma14g24500.1	+	✓					+	
Glyma14g27020.1								
Glyma14g27260.1		✓					+	
Glyma14g35490.1								
Glyma14g37050.1						(f) -6.21	+	
Glyma14g37140.1	+							
Glyma14g39530.1	+	✓					+	
Glyma15g02950.1	+	✓					+	
Glyma15g03920.1*	+				(f) 1.21		+	
Glyma15g04620.1*	+		8.87				+	
Glyma15g04620.2			8.87				+	
Glyma15g04620.3			8.87				+	
Glyma15g04620.4			8.87				+	
Glyma15g07230.1	+				(f) -4.25		+	
Glyma15g14190.1	+							
Glyma15g14620.1	+	✓					+	
Glyma15g14620.2						(f) 9.16	+	
Glyma15g15400.1	+	✓					+	
Glyma15g17710.1		✓					+	
Glyma15g19350.1								
Glyma15g19360.1	+	✓					+	
Glyma15g19360.2	+					(3A) -15.64	+	

Glyma15g19930.1								
Glyma15g20630.1								+
Glyma15g33040.1								+
Glyma15g33700.1								+
Glyma15g35670.1								+
Glyma15g35860.1	+	✓						+
Glyma15g41250.1*	+					(f) 1.17		+
Glyma15g41810.1	+	✓						+
Glyma16g00920.1	+	✓						+
Glyma16g00930.1								+
Glyma16g02570.1*	+				(a) 23.34			+
Glyma16g06900.1	+	✓						+
Glyma16g07930.1		✓						+
Glyma16g07960.1	+							+
Glyma16g13440.1	+							+
Glyma16g16270.1								+
Glyma16g16800.1								+
Glyma16g31280.1	+	✓						+
Glyma16g31280.2		✓						+
Glyma16g32100.1								+
Glyma16g34490.1	+	✓						+
Glyma17g03480.1	+							+
Glyma17g04170.1*	+					(a) 3.95		+
Glyma17g05830.1	+	✓						+
Glyma17g07330.1	+	✓						+
Glyma17g09310.1		✓						+
Glyma17g09640.1	+							+
Glyma17g10820.1	+			-2.44		(e) -1.73		+
Glyma17g12820.1								+
Glyma17g13010.1		✓						+
Glyma17g14290.1	+	✓						+
Glyma17g14290.2*	+					(b) 1.27		+
Glyma17g15270.1	+	✓						+
Glyma17g16980.1*	+						(2B) 1.55	+
Glyma17g17560.1	+							+
Glyma17g22860.1								+
Glyma17g26240.1	+			-1.08		(e) -1.14		+
Glyma17g35020.1		✓						+
Glyma17g35620.1	+			-10.88		(b) -2.52		+
Glyma17g36370.1*	+					(a) 6.21		+
Glyma18g03430.1		✓						+
Glyma18g04580.1	+	✓						+
Glyma18g07250.1		✓					(3A) 2.99	+
Glyma18g07360.1	+							+
Glyma18g07960.1	+	✓						+
Glyma18g10920.1	+	✓						+
Glyma18g16040.1						(a) 5.08		+
Glyma18g18140.1						(c) 1.71		+
Glyma18g26600.1	+					(f) -10.21		+
Glyma18g32460.1								+
Glyma18g37640.1								+
Glyma18g39740.1	+					(c) -1.70		+
Glyma18g39760.1	+	✓						+
Glyma18g39760.2	+	✓						+
Glyma18g40790.1								+
Glyma18g41520.1	+	✓						+
Glyma18g42530.1						(d)) -3.20		+
Glyma18g46480.1	+	✓						+
Glyma18g49360.1*	+					(f) 1.51		+
Glyma18g49630.1	+	✓						+
Glyma18g49670.1	+							+

Glyma18g49690.1	+	✓							+
Glyma18g50890.1	+								
Glyma19g00930.1	+	✓							+
Glyma19g02090.1	+								
Glyma19g02600.1	+							(f) -8.10	+
Glyma19g02890.1	+							(a) -1.19	+
Glyma19g02980.1	+	✓							+
Glyma19g05080.1	+								+
Glyma19g07830.1	+	✓							+
Glyma19g13990.1		✓							+
Glyma19g14230.1	+								+
Glyma19g14270.1	+	✓							+
Glyma19g24450.1									
Glyma19g24530.1									
Glyma19g24770.1									
Glyma19g25790.1									
Glyma19g27750.1		✓							+
Glyma19g29670.1	+								+
Glyma19g29750.1	+							(f) -2.83	+
Glyma19g34740.1	+	✓							+
Glyma19g36830.1*	+							(3B) 1.63	+
Glyma19g40250.1*	+								+
Glyma19g40650.1		✓							+
Glyma19g40670.1		✓							+
Glyma19g41010.1	+							(f) -1.26	+
Glyma19g41150.1									+
Glyma19g41250.1	+								+
Glyma19g43740.1	+								+
Glyma19g44660.1*	+							(f) 13.96	+
Glyma20g01450.1		✓							+
Glyma20g01610.1	+	✓							+
Glyma20g03940.1		✓							+
Glyma20g04240.1	+	✓							+
Glyma20g04510.1		✓							+
Glyma20g08730.1		✓							+
Glyma20g09140.1									
Glyma20g11040.1	+	✓							+
Glyma20g11110.1		✓							+
Glyma20g12290.1								(f) 6.77	+
Glyma20g20980.1	+								+
Glyma20g21670.1									
Glyma20g21680.1									
Glyma20g22230.1*	+								+
Glyma20g24600.1								(a) 5.74	+
Glyma20g25110.1	+	✓							+
Glyma20g29710.1	+	✓							+
Glyma20g29730.1*	+							(e) 1.20	+
Glyma20g32500.1	+	✓							+
Glyma20g32510.1	+	✓							+
Glyma20g34140.1	+								+
Glyma20g35180.1*	+							(a) 26.51	+

ANEXO B

GmMYBs induzidos após a infecção por *P. pachyrhizi* em soja (continuação)

GENE ID	R2R3 COMPLETO	SOYBASE	ANÁLISE DE EXPRESSÃO				OUTROS	EXPRESSÃO CONFIRMADA
			GENOSOJA - LGE		Genevestigator (Fold-chage)			
			SUPERSAGE (Fold-change)	SSH (FPKM)	REAÇÃO COMPATÍVEL	REAÇÃO INCOMPATÍVEL		
			PI561356 <i>MOCKKINOC</i>	PI561356 <i>MOCKKINOC</i>	PI462312 X	PI462312 X		
		12-24-48HAI	12-192HAI	TAIWAN80-2	HAWAII94-1			
Glyma20g36600.1		✓						+
Glyma20g36600.2		✓						+
Gnl ug gma#s34273466								
Gnl ug gma#s39312923	+							
Gnl ug gma#s53086590;put-169a-glycine_max-63505	+							
Glyma14g21490.1								+
Put-169a-glycine_max-102940	+							
Put-169a-glycine_max-179128;gnl ug gma#s53090519	+							
Put-169a-glycine_max-81255;gnl ug gma#s48315927	+							

SSH (LINO, 2011): 12, 24 e 48 HAI (A); 72 e 96 HAI (B) e 192 HAI (C); SuperSAGE-LGE: biblioteca SuperSAGE criada a partir ensaios de folhas de soja PI561356 (resistente) infectadas com *P. pachyrhizi*, e controles compostos de folhas sem infecção. Ambos disponíveis em www.lge.ibi.unicamp.br/soja/.

Genevestigator: Ensaio formado por plantas de soja (V3) PI462312 (Rpp3) em três tratamentos: com cepa virulenta de *P. pachyrhizi* (TW80-2), com cepas avirulentas (HW94-1) e controles (mock). As análises foram realizadas em 6 pontos: 12hpi (a), 24hpi (b), 72hpi (c), 144hpi (d), 216hpi (e), 288hpi (f). Dados do experimento disponíveis em: http://www.plexdb.org/modules/PD_browse/experiment_browser.php?experiment=GM36

P. pachyrhizi study 2 e 3: Análise do perfil de expressão de soja PI200492 (Rpp1) com dois tratamentos (HW94-1 ou TW72-1) em diferentes tempos: P. pachyrhizi study 2 (early, 6-12hpi)/non-infected leaf samples (PI200492)(2a), P. pachyrhizi study 2 (intermediate-24hpi)/non-infected leaf samples (PI200492)(2b), P. pachyrhizi study 2 (late48hs)/non-infected leaf samples (PI200492)(2c), P. pachyrhizi study 3 (early, 6-12hpi)/non-infected leaf samples (PI200492)(3a), P. pachyrhizi study 3 (intermediate – 24hpi)/non-infected leaf samples (PI200492)(3b), P. pachyrhizi study 3 (late-48hpi)/non-infected leaf samples (PI200492)(3c). Disponível em <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE8432>

P. pachyrhizi study 4 e 5: Ensaio com folhas de soja PI230970 (Rpp2) e Embrapa-48, infectadas com FAS, em: P. pachyrhizi study 4 (early, 6-24hpi)/mock inoculated PI230970 leaf samples (4a), P. pachyrhizi study 4(intermediate, 36-72hpi)/mock inoculated PI230970 leaf samples(4b), P. pachyrhizi study 4 (late, 96-168hpi)/mock inoculated PI230970 (4c), P. pachyrhizi study 5 (early, 6-24hpi)/mock inoculated embrapa48 leaf samples (early)(5a);P. pachyrhizi study 5 (intermediate, 36-72hpi)/mock inoculated embrapa48 leaf samples (5b), P. pachyrhizi study 5 (late, 96-168hpi)/mock inoculated embrapa48 leaf samples (5c), P. pachyrhizi/non-infected leaf samples (5601t)_ (6). Disponível em: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE33410>

Gene ID acompanhado de: “*”- apresentaram expressão em resposta a *P. pachyrhizi* (Fold-change $\geq 1,0$; ratio $\geq 0,5$); Sequências sombreadas (cinza) apresentam R2R3, contudo sem nenhuma expressão;

Sequências sem domínio R2R3 (+) e sem expressão em algum dos bancos de expressão consultados (ultima coluna: +) representam sequências excluídas

ANEXO C

GmMYBs induzidos em soja durante ataque por *Aphis glycines*, *Heterodera glycines* e *Phytophthora sojae*.

Gene ID	R2R3	<i>Aphis glycines</i> x soja (V3)		<i>Heterodera glycines</i> X WH82			<i>P. pachyrhizi</i>	<i>Phytophthora sojae</i>								
		SD01-76R (S)	LD05-16060 (R)	2dpi	5dpi	10dpi	soja (5601T)	athow	conrad	general	ox20-8	PI291327	sloan	v71-370	Williams	Harosoy
Glyma0041s00310.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma0076s00200.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g00810.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g02070.1*a	+	-	-	-	-	-	-	1,95 a	1,65*/1,54(4)	1,57a	2,22	-	1,33(4)	1,52a/1,3(4)	1,77a/1,36b	-
Glyma01g05190.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g05270.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g05980.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g06190.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g06220.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g09280.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g20440.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g21900.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g23460.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g24780.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g26650.1*e	+	-	-	1,77a	-	-	-	1,28b	-	-	-	1,63b	-	-	-	-
Glyma01g27720.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g28790.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g39740.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g40410.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g41610.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g42050.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g42650.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g43120.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma01g44370.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g00820.1*g	+	-	3,26a	2,26a	1,70a	-	-	8,46b	19,80a/1,48(3)/ 41,12(4)/23,88(5)/9,8(6)/	19,00a	29,51a	20,23a/ 30b	19,89a/13,19b/ 40,39(4)/24,91(5)/9,8(6)/ 6,4(7)/6,32(8)/ 7,0(9)	17,51/18,4/ 2,14(3)/24(4)/11, 49(5)/10(6)/6,9(7)/ 6,1(8)/5,5(9)	16,29b	3,07b
Glyma02g00960.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g01300.1*g	+	1,79a	-	1,23a	-	-	-	1,63b	5,03a/1,62(3)/5 ,49(4)/4,59(5)/2 ,3(6)	4,43a/ 3,44b	9,08a/ 11,69b	2,3a/ 2,39b	6,56a/5,58b/6, 94(4)/5,27(5)/3 ,63(6)/3,3(7)/2, 7(8)/2,7(9)	5,02a/3,64b/4,07(4)/3,13(5)/2(6)/3(7)/2,5(8)/2,9(9)	6,53a/4,67b	3,78b/ 7,2c
Glyma02g01740.1*d	+	-	1,65a	-	-	-	-	1,14a	-	1,28	-	-	-	-	-	-
Glyma02g02310.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g12100.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g12240.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g12250.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g12260.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g13770.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g26540.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g26980.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g30820.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g30920.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g34010.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g39000.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g39070.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g40930.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma02g41180.1a	+	-	-	-	-	-	-	1,42a/1 ,68b	1,69b/1,28(6)	1,37a/1,3 5b	1,53a/1 ,93b	1,38a	1,76a/2,05b/1, 25(3)/1,6(6)/1,	1,55(7)/1,3(8)/1,2 (9)	1,45a/1,68b	-

												9(7)/1,67(8)/1,65(9)				
Glyma02g41260.1a	-	-	-	-	-	-	1,25a	1,12a/1,19b/ 1,17(5)/1,17(6)	1,35	1,41b	1,27a/1,35b	1,34a/1,48/1,29(5)/1,08(7)/1,198/1,3(9)	1,25a/1,14(6)/1,13(7)	1,46b	-	
Glyma02g41440.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma02g42030.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma02g42140.1d	-	1,16b	-	-	-	-	1,60b/1,62b	1,54b	1,25a/1,62b	1,59b	-	1,73b/1,26(6)	1,13a/1,32a 1,66b	-	-	
Glyma02g42140.2a	-	-	-	-	-	-	-	-	-	-	-	1,55(7)/1,3(8)/ 1,3(9)	-	-	4,86c	
Glyma02g43280.1a	+	-	-	-	-	-	-	-	-	-	-	-	1,07(5)	-	-	
Glyma02g44190.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g00890.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g00980.1d	+	1,11a	-	-	-	-	1,84a/2,73b	3,38a/3,72b/ 1,94(5)/4,12(6)	2,17- 2,28a/2,61b	4,99- 5,70a/3,37- 6,56b	4,63- 4,70a/5,47- 6,56b	5,34a/9,18b/1,98(5)/4(6)/5,1(7)/4,3(8)/4,8(9)	1,62a/2,73b/1,41(5)/2,46(6)/3,4(7)/ 2,7(8)/2,4(9)	2,85-3,18b	-	
Glyma03g01540.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g03660.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g06230.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g07840.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g08710.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g13550.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g14440.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g15810.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g15870.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g15930.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g19030.1g	+	-	1,22/1,09b	2,01a	1,15a	-	1,05	1,05	1,01	1,07b	1,05	1,06/1,11b	1/1,10b	1,03	1,06	-
Glyma03g19470.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g22590.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g26210.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g26830.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g31980.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Glyma03g34110.1*g	+	1,89a/2,8b	-	1,48a/1,16b	1,06b	-	-	-	1,08(3)/1,31(4)	-	-	-	1,0(3)/1,35(4)	1,05(3)/1,12(4)	-	
Glyma03g37640.1g	+	-	1,11a/1,19b	1,03a	-	-	-	1,13b	1,03	1,06/1,14b	1,03/1,10b	1,12b/1,06(6)/1,0(7)/1,0(8)	1,05/1(3)/1(4)/ 1,08(5)/1,0(7)/1,0	1,07	1,03a/1,04b	

Glyma20g22230.1*	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma20g24600.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma20g25110.1a	+	-	-	-	-	-	-	-	1,48a/ 1,79b	-	-	1,57a/1,65b/1, 48(7)/1,65(8)/1, ,52(9)	-	-	-	-	-	-	-
Glyma20g29710.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma20g29730.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma20g32500.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma20g32510.1a	+	-	-	-	-	-	-	1,84b	-	-	1,42b	1,65a	-	-	-	1,78a/2,0b	64,85c	-	-
Glyma20g34140.1	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma20g35180.1*a	+	-	-	-	-	-	-	3,51a/ 3,43b	3,75a/5,0b/2,6 7(4)/3,05(5)/3, 25(6)	3,74a/5,6 3b	3,09a/ 7,19b	4,0a/7,57 b	3,2a/5,58b/ 1,94(4)/2,61(5) / 4,1(6)/2,4(7)/ 2,0(8)/2,0(9)	4,63a/5,69b/ 2,87(4)3,24(5)/ 3,11(6)/3,4(7/3,0(8)/2,9(9)	3,52a/5,88b	-	-	-	-
Glyma20g36600.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma20g36600.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gma#s34273466	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gma#s39312923	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gma#s53086590;put-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glyma14g21490.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Put-169a-glycine_max-102940	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gma#s53090519	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Put-169a-glycine_max-81255;gnl ug gma#s48315927	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Todos os dados foram obtidos através de consulta ao Genevestigator (<https://www.genevestigator.com/gv/plant.jsp>), utilizando os seguintes parâmetros para a análise dos resultados: $p < 0,05$; $\text{ratio} \geq 0,5$; $\text{fold-change} \geq 1,0$. Gene IDs acompanhados de "*" representam os 69 modelos com indução na interação soja-FAS; "a" – *P. sojae*; "b" – *A. glycines*; "c" – *H. glycines*; "d" – *P. sojae* + *A. glycines*; "e" – *P. sojae* + *H. glycines*; "f" – *H. glycines* + *A. glycines*; "g" – *P. sojae* + *A. glycines* + *H. glycines*;

Aphis glycines x soja (V3): Análise da resposta transcricional de plantas de soja à infestação de pulgões de soja resistentes e suscetíveis. Experimento com três fatores: variedade de soja em V3 (susceptível SD01-76R, resistente LD05-16060), o tratamento de pulgões (controle, pulgões), e duração de infestação (1 dia e 7 dias). Dados disponíveis em: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE35427>. (a): 1dpi, (b): 7dpi (estudo 2);

Heterodera glycines: Análise de perfis de transcrição de raízes de soja (Williams 82), via microarranjo, em três pontos de tempo de infecção (2,5,10 dias após inóculo - dpi) com nematódeos císticos de soja. Dados disponíveis em: <http://www.ebi.ac.uk/arrayexpress/browse.html?keywords=%09+E-MEXP-808>. (a): estudo 1, (b): estudo 2.

P. pachyrhizi x soja 5601T: Análise de expressão gênica por microarranjo em plantas de soja infectadas com o fungo *Phakopsora pachyrhizi* (72hpi) na fase de crescimento V2. Dados disponíveis em: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE7108>.

Phytophthora sojae: Identificação de genes diferencialmente expressos durante infecção pelo patógeno *P. sojae* em 8 cultivares de soja que diferem na resistência quantitativa (Athow, Conrad, General, Ox20-8, PI291327, Sloan, V71-370, Williams), através de microarranjo. As análises foram realizadas em tempos: 3dpi (a) e 5dpi (b). Dados disponíveis em: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE7124>; Padrões de expressão em diferentes tempos, durante a infecção por *P. sojae* (isolado PT2004C2.S1), de cultivares de soja que diferem na resistência quantitativa (resistentes a cepa – CONRAD e V71-370, suscetíveis – SLOAN e VP-RIL9): estudos (3) (24hpi), (4) (48hpi), (5) (72hpi), (6) (120hpi), disponível em: http://www.plexdb.org/modules/PD_browse/experiment_browser.php?experiment=GM3 e estudos (7)(24hpi), (8) (48hpi), (9) (72hpi), disponível em: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE11611>. Estudo 10 – Amostras de hipocótilo de soja Harosoy (susceptível) inoculados com a cepa P6497 (raça 2) de *P. sojae*, foram analisados através de microarranjo em 12 (a), 24 (b), 48 (c) horas após o inóculo (hpi). Todos os resultados foram obtidos a partir de ensaios com microarranjo. Expressão em *Fold-change* para todos os ensaios.

ANEXO D

Sequências anotadas como MYBs excluídas

Glyma	Expressão		Domínio R2R3
	Soybase	Geneinvestigador (Fold-change)	
Glyma01g20440.1	-		
Glyma01g21900.1	+		
Glyma01g23460.1	-		
Glyma01g24780.1	-		
Glyma01g27720.1		+	
Glyma01g28790.1	-		
Glyma02g26540.1	-		
Glyma02g26980.1	-		
Glyma02g30820.1	-		
Glyma02g30920.1	-		
Glyma02g34010.1	-		
Glyma02g39000.1	+		
Glyma02g40930.1	+		
Glyma02g41260.1		+	
Glyma02g42140.1	+		
Glyma02g42140.2		+	
Glyma02g44190.1		+	Sem R2R3
Glyma03g03660.1	-		
Glyma03g06230.1	-		
Glyma03g08710.1	-		
Glyma03g13550.1	-		
Glyma03g14440.1		+	
Glyma03g22590.1	-		
Glyma03g26210.1	-		
Glyma03g26830.1	-		
Glyma03g38550.1		+	
Glyma04g04850.1	+		
Glyma05g02170.1	-		
Glyma05g07980.1	+		
Glyma05g15860.1	-		
Glyma05g18820.1	-		
Glyma05g31400.1		+	
Glyma05g32270.1		+	

1		
Glyma05g37570.		+
1		
Glyma05g37570.		+
2		
Glyma06g04950.		+
1		
Glyma06g44120.	-	
1		
Glyma06g46590.	+	
1		
Glyma07g08080.	-	
1		
Glyma07g17660.	+	
1		
Glyma07g28310.		+
1		
Glyma07g28590.	-	
1		
Glyma07g29680.	-	
1		
Glyma07g34030.	-	
1		
Glyma07g34030.		+
2		
Glyma08g03530.	-	
1		
Glyma08g20150.	+	
1		
Glyma08g20150.		+
2		
Glyma08g26890.	+	
2		
Glyma08g27660.	-	
2		
Glyma08g40020.	-	
1		
Glyma08g43000.	+	
1		
Glyma08g44350.	-	
1		
Glyma09g12230.	-	
1		
Glyma09g25460.	-	
1		
Glyma09g29240.	-	
1		
Glyma09g36980.	-	
1		
Glyma10g01810.		+
1		
Glyma10g02110.	-	
1		
Glyma10g04660.	+	
1		
Glyma10g06930.	-	
1		
Glyma10g14830.	-	
1		
Glyma10g22770.	-	
1		
Glyma10g24130.	-	
1		
Glyma10g35670.	+	
1		
Glyma11g02580.	+	
3		
Glyma11g02780.	-	
1		

ANEXO D
(continuação)

Glyma	Expressão		Domínio R2R3
	Soybase	Genevestigat or	
Glyma11g0488 0.1		+	
Glyma11g0876 0.1		+	
Glyma11g1518 0.2	+		
Glyma11g1518 0.3	+		
Glyma11g2296 0.1	+		
Glyma11g2760 0.1	-		
Glyma11g3493 0.1	+		
Glyma11g3493 0.2	+		
Glyma11g3493 0.3	+		
Glyma11g3493 0.4		+	
Glyma11g3821 0.1	+		
Glyma12g0711 0.1	-		
Glyma12g0711 0.2		+	
Glyma12g1079 0.1	+		
Glyma12g1160 0.1	-		
Glyma12g1299 0.1	-		
Glyma12g1529 0.1	-		Sem R2R3
Glyma12g1702 0.1	-		
Glyma12g2903 0.1	-		
Glyma12g3213 0.1	+		
Glyma12g3213 0.2	+		
Glyma13g0374 0.1	+		
Glyma13g0702 0.1	+		
Glyma13g0909 0.1	-		
Glyma13g2076 0.1	+		
Glyma13g2137 0.1	+		
Glyma13g2259 0.1	-		
Glyma13g2572 0.1		+	
Glyma13g2989 0.1	-		
Glyma13g3637 0.1		+	
Glyma13g3638 0.1	+		
Glyma13g3790 0.1	+		
Glyma13g3791 0.1	+		

Glyma14g0459 0.1		+
Glyma14g0675 0.1		+
Glyma14g0809 0.2	+	
Glyma14g1048 0.1	+	
Glyma14g2702 0.1	-	
Glyma14g2726 0.1	+	
Glyma14g3549 0.1	-	
Glyma14g3705 0.1		+
Glyma15g0462 0.2		+
Glyma15g0462 0.3		+
Glyma15g0462 0.4		+
Glyma15g1771 0.1	+	
Glyma15g1993 0.1	-	
Glyma15g2063 0.1	-	
Glyma15g3304 0.1	-	
Glyma15g3370 0.1	-	
Glyma15g3567 0.1	-	
Glyma16g0093 0.1	-	
Glyma17g1282 0.1	-	
Glyma17g1301 0.1	+	
Glyma17g2286 0.1	-	
Glyma18g0343 0.1	-	
Glyma18g0725 0.1		+
Glyma18g1814 0.1		+
Glyma18g3246 0.1	-	
Glyma18g3764 0.1	-	
Glyma18g4079 0.1	-	
Glyma18g4253 0.1		+
Glyma19g2453 0.1	-	
Glyma19g2579 0.1	-	
Glyma19g2775 0.1	+	
Glyma19g4115 0.1	-	
Glyma20g0145 0.1	+	
Glyma20g0394 0.1	+	
Glyma20g0873 0.1	+	
Glyma20g0914 0.1	-	
Glyma20g1111 0.1	-	

ANEXO D
(continuação)

Glyma	Expressão		Domínio R2R3
	Soybase	Genevestigat or	
Glyma20g1229 0.1		+	
Glyma20g2167 0.1	-		
Glyma20g2168 0.1	-		
Glyma20g2460 0.1		+	Sem R2R3
gnl UG Gma#S 34273466	-		
Glyma14g2149 0.1	-		
Glyma	Soybase	Genevestigat or	Apenas R2
Glyma01g0081 0.1	+		MGRAPCCDKSNV KRGPWSPDE DATLKNYLEKHGTGGNWIALPKKAG LKRCGKSCRLRWLNLYLRPHIKLGGFTEEDKIICTLYDIGSRQVVLYSS STSREN RQ
Glyma01g0619 0.1	-		PCCEKVGLKKG PWTPEEDKKL VAYVEKHDHGNWHSVPAKAGFKLRL QIHFRFVMILDIVGHC RQM QSIDKVAIVVAMSSKMWKELQAEVDQLPQ ARYKKRKLQHGG RPYHYSTSCSSWKQ
Glyma03g0784 0.1	-		NNIKSNLRKGLWSPEEDEKLIRYMITKGQRCWSDIARNGLQRCGKSC QLRWINYFRPDLKRG
Glyma03g3807 0.1	-		
Glyma04g0055 0.2		+	AWTKEEDHRLISYIRAHGEGCWRSLPKAAGLLRCGKSCRLRWINYLRP DLKRGNF SLEEDQLIIKLHSLGNKLPGRTDNEIKNYWNTHIRRKL LSRG ID
Glyma05g1814 0.1	-		MGRSPCCNENGLKKG PWTPEEDQKL VQHIQKHGHSWRALPKQAGL NRCGKSCRLRW TNYLRPDIKRGKFSQEEEQ TILNLHSILGNX
Glyma05g3612 0.1	-		MGRAPCCDKANV KRGRWSPEEDELKNYLKKHATPGNWITLPQKAGL KRCGKSCRLRWLNLYLRPHIKHGGFTHEEDQFICSLYATIGTRQIECFLF CLFIPYINKWSLIAAQLPGRTDNDVKNHWNTK [W]-x(20)-[W]-x(19)-[W]-x(12)-[F]-
Glyma06g0063 0.2		+	WTKEEDHRLISYIRAHGEGCWRSLPKAAGLLRCGKSCRLRWINYLRPD LKRGNF SLL
Glyma06g4553 0.1	-		WSREEDEKLTAYV KRHGHSNWRQLPKFAGLARC GSKSCRLRW
Glyma06g4556 0.1	-		WTPEEDKKLVEYITRYGHWNWRLLPKFAGLARC GSKSCRLRW
Glyma07g1525 0.1	-		MGRAPCCDKSNV KRGPWSPDE DATLKNYLEKHGTGGNWIALPKKAG LKRCGKSCRLRWLNLYLRPHIKLGGFTEEDNIICTLYDIIGSRQLTAQLP GRTDNDVKNHWNTK LKKKFLAGNTSSSIATTS [W]-x(20)-[W]-x(19)-[W]-x(12)-[F]-
Glyma07g3558 0.1	-		WRPAEDSKL KELVALYGPQNWNLIAEKLEGRSGKSCRLRW
Glyma09g1217 0.1	-		WSTEEVQILIRLVERYDPQNWSLISRYIKGRSNKLC LLRW
Glyma10g0180 0.1	-		WTAEEDETLAKYIQTNGEGSWRSLPKNAGLLRCGKSCRLRW
Glyma10g0425 0.1	-		WTQEEDLILIHINTHGHKNWRALPKQAGLLRCGKSCRLRW
Glyma10g3506 0.1	+		WSAEEDKILMNDVQVHGERKWRRELSKRAGLKRCGKSCRLRW
Glyma12g3254 0.1	-		RTPSSSHKSGLNKGTWTPPEEAKLIAYITRYGSWNWRQLPRFAGLAR CGKSCRLRWLNLYLRPNIKRGNYTKEEEEIIRLHEKLGNYALTFQAKLN FPQHCTASKARQGN YKIFPYFNTYRPMHLAX
Glyma13g3792 0.1	-		WTPEEDRKLIDYVTKYGHWNWRLLPKFAGLARC GSKSCRLRW
Glyma14g2149 0.1		+	WTREEELALIH YHQIFGNKWAELSKVIPGRTDNAIKNH W
Glyma15g1935 0.1	-		WSYEE DKIPMKYVQVHGERNWRNFLERASLKRCGESCKHPW
Glyma17g0931 0.1	+		WSPEEDEKLFNYITRFVGVCSWSSVPKLAGLQRCGKSCRLRWINYLRP DLKRGMF
Glyma18g5088	-		WIGEEDKLLCEYVSLHGEGRWSSVAKLTGLNRSGKSCRLRWVNYLRP

0.1			GLKKAQL
Glyma19g2445	-		WTTEEEELTKYIMANGEGSWRSLPKNTGLLRGNSCRLRWINYLRAD LKRGTG
Glyma0076s00	-		VWRKGPWTAEDRLLVEYVRLHGEGRWNSVARLAX
Glyma	Soybase	Genevestigat or	Apenas R3
Glyma04g1104	+		MGRSPCCDENGLKKGPWTPPEEDQKLQVHIQKHGHGLNRCGKSCRLR WTNYLRPDIKRGKFSQEEEQTLHLHLSILGNKWSSIATHLPGRTDNEIK NFW
Glyma04g3572	-		WTDEEDKVLIEVHDEVGNKWAEIATRLPGRTEINSIKNHWNATK
Glyma06g2002	-		WTNYPRPDLKDDNFTTQEEDLIKLHAAIGSRWSIAAQLPGRDTDVK NYW
Glyma19g2477	-		WFNQLDPRINRNPFTEEEEERLLASHRIHGNRWAVIARHFPGRTDNAV KNHW
Glyma	Soybase	Genevestigat or	R2 Incompleto
Glyma05g0876	+		LVELVNQFGLKKWSQIAKLLPGRIGKQCRERWHNHLKPNIKKDSWTLE EDMILIKAHKKVGNKWSEIAKRLPGRPENTIKNHWNTRKRRQNCRRHK HTIYEGSLLHGYYIKK
Glyma05g3321	-		RCGKSCRLRWNTYQRPDVKKGKFTTEESNLIIHLHSLGNKNEHKPSS LYFHLLNCLILSILREIKVDKNNLQKTCIWSQMARSPLGRDTDNKIKNYW [W]-x(19)-[]-x(19)-[]-x(12)-[F]-x(18)-[W]-x(18)-[W]
Glyma13g3852	+		NDDETAAKEEEEEADILMEYVKKHGEENWNSVQKNSGLLRGKSCRLR WANHLRPNLKKGAFSPPEEQVIIDLHSLGNKWARMAAQLPGRTDNEI KNFWNTRMKRRRQAGLPIYPPE []-x(19)-[W]-x(19)-[W]-x(12)-[F]-x(18)-[W]-x(18)-[W]
Glyma16g0793	+		RVLVQLVERFGFKKWSHIAIRLLNGRVGKQCRERWHNHLRPNIRKESW NEEEDRILIEAHKEVGNRWAEIARRMPGRTEINTIKNHWNATKRRRLNAK RLRNKRRGLKGPPTLLESYIRQVTANEETQKEMIKNSMSNMNLKGESX
Glyma16g1627	-		
Glyma16g1680	-		
Glyma06g4552	0.1	+	WTPEEDKKLIAYITRYGHWNNLLPKFAGLERCGKSCRLRWLNLYLRPNIKRGN ^Y T QEEDETIIMVQRLGNRWSLIAAQLPGRTDNEIKNYWHTNLKK [W]-x(19)-[W]-x(19)-[W]-x(12)-[L]-x(18)-[W]-x(18)-[W]
lyma06g45570.	1	-	WTPEEDKKLIAYITRYGHWNNLLPKFAGLERCGKSCRLRWLNLYLRPNIKRGN ^Y T QEEDETIIMVQRLGNRWSLIAAQLPGRTDNEIKNYWHTNLKK [W]-x(19)-[W]-x(19)-[W]-x(12)-[L]-x(18)-[W]-x(18)-[W]
Glyma08g4292	0.1	-	WTTAEDVILMDYVTNKGEENWNAVQRNTGLNRCGKSCRHRWANHLRPNLKKG AFSPEEEKLIVDLHAQFGNKW
Glyma09g3701	0.1	-	VRKGAWSKCEDDLLRACVQLYGEKWHLVPQRAGLNRCRKSRLRWLNLYLKPN IKRGDFSEDEVDMMIRMHKLGNRHLFRDLKEDQGLSKKDDQGNKWS [W]-x(19)-[W]-x(19)-[W]-x(12)-[F]-x(18)-[W]-x(18)-[W]
Glyma10g3087	0.1	+	MRIMIKGGVWKNTEDEILKAAVMKYGKNQWARISSLLVRKSAKQCKARWYEWLD PSIKKTEWTREDEKLLHLAKLMPTQWRTIPIVGRTPSQCLERYEKLLDVACVKD ENYEPGDDPRKLRPGEIDPNPESKPARPDPVDMDEDEKEMLSEARARLANTKGG KAKRKAREKQLEEARRLASLQKKRELKAAGIDIRQRKRKRKIDYNAEIPFEKRPP PGFFDVTDEDPRVEQPQFPTTIEELGKRRVDVEAQLRKQDIANKIAQRQDAPS AILHAN
Glyma12g1133	0.1	-	WTPEEDKKLIAYITRYGHWNNLLPKFAGLERCGKSCRLRWLNLYLRPNIKRGN ^Y T QEEDETIIMVQRLGNRWSLIAAQLPGRTDNEIKNYWHTNLKK [W]-x(19)-[W]-x(19)-[W]-x(12)-[L]-x(18)-[W]-x(18)-[W]
Glyma12g1149	0.1	-	WTPEEDKKLIAYITRYGHWNNLLPKFAGLERCGKSCRLRWLNLYLRPNIKRGN ^Y T QEEDETIIMVQRLGNRWSLIAAQLPGRTDNEIKNYWHTNLKK [W]-x(19)-[W]-x(19)-[W]-x(12)-[L]-x(18)-[W]-x(18)-[W]
Glyma12g3253	0.1	-	WTPEEDKKLIAYITRYGHWNNLLPKFAGLERCGKSCRLRWLNLYLRPNIKRGN ^Y T QEEDETIIMVQRLGNRWSLIAAQLPGRTDNEIKNYWHTNLKK [W]-x(19)-[W]-x(19)-[W]-x(12)-[L]-x(18)-[W]-x(18)-[W]
Glyma15g1462	0.2	+	WTVDEDLALINYIANHGEGRWNSLARSAGLKRTGKSCRLRWLNLYLRPDRRGIN
Glyma20g3660	0.1	+	MRIMIKGGVWKNTEDEILKAAVMKYGKNQWARISSLLVRKSAKQCKARWYEWLD PSIKKTEWTREDEKLLHLAKLMPTQWRTIPIVGRTPSQCLERYEKLLDAACVKD ENYEPGDDPRKLRPGEIDPNPESKPARPDPVDMDEDEKEMLSEARARLANTKGG KAKRKAREKQLEEARRLASLQKKRELKAAGIDIRQRKRKRKIDYNAEIPFEKRPP PGFFDVTDEDPRVEQPQFPTTIEELGKRRVDVEAQLRKQDIANKIAQRQDAPS AILHAN
Glyma20g3660	0.2	+	WKNTEDEILKAAVMKYGKNQWARISSLLVRKSAKQCKARWYEWLDPSIKKTEWT REDEKLLHLAKLMPTQW
Glyma	Soybase	Genevestigat or	Substituição de "W"

Glyma02g0231 0.1	+		MVNEESKKKERHIVT W TQEEDDILREQIGVHGTE W AIASKFKDKTTRQCRRRW YTYLNSDFKKG W SPEEDMLLCEAQKIFGNRWTEIAKVVSGRTDNAVKNR F STL CKKRAKY
Glyma08g4095 0.1	+		KKKERHIVT W TQEEDDILREQIGIHGTE W AIASKFKDKTTRQCRRRW W YTYLNSD FKKGG W SAEEDMLLCEAQKVFGNRWTEIAKVVSGRTDNAVKNR F STLCRKKQK YAALAKENSTSYINSNNKRMMLQHCNNMDTTSESGVPIKNLRRADIADDAEKIKFE DRSHLRNGTPINQQPRAPLAVLAQNCHNSNNLDQHHLCNPKFSSSAQNNKIQQ TFLKKDDPKISALMQQAELSSALKVDTENMDQSLSENAWKVLQEFNLRSKESDIP GQKVDP
Glyma15g1419 0.1	-		MGWGVIEEGWRKGP W TSEEDRLLLYVKFHGEGR W NSAARLAGLKRNGKSCR LR W VNYLRPDLEKQITPQEESIILEHARWGNRWSTIARSLPGRTDNEIKNY C RT HFKK [W]-x(19)-[W]-x(19)-[W]-x(12)-[I]-x(18)-[W]-x(18)-[C]
Glyma18g1604 0.1		+	KKKERHIVT W TQEEDDILREEIGIHGTE W AIASKFKDKTTRQCRRRW W YTYLNSD FKKGG W SAEEDILLCEAQKIFGNRWTEIAKVVSGRTDNAVKNR F STLCRKKQKYE ALAKENSTSYINSNNKRVMFQHCNNMDTTSESGVPIKNLRRTHIHYNAEKIKFEDR LHLRNETPINQQPRAPLAVLAQNSHNSNNLPDQHHVCNPKFSSLAQNYKIQQGFL KKDYPKISALMQQAELSSALKVDAGNMDQSFENAWKVLQEFLLKRTKESNIPGQ KIPD
Glyma01g0519 0.1		+	MVNEESKKKERHIVT W SQEEDDILREQIGVHGTE W AIASKFKDKTTRQCRRRW W YTYLNSDFKKG W SPEEDKLLCEAQKIFGNRWTEIAKVVSGRTDNAVKNR F STLC KKRAKY [W]-x(19)-[W]-x(18)-[W]-x(12)-[F]-x(18)-[W]-x(18)-[F]
Glyma	Soybase	Genevestigat or	Com variação no espaçamento entre "W"
Glyma04g3463 0.1	-		WTTEEDTKLAFGSKHRSGNWTSPKKSRLKRCGKSCRLRWNTNYPRPDLKDDN FTTQEDLIMKLHAAIGSR W SIVAQQLLGRTDNDVKNY W [W]-x(19)-[W]-x(19)-[W]-x(12)-[F]-x(17)-[W]-x(18)-[W]
Glyma18g3976 0.2	+		W RAEEDALLRAYVKQYGP W NLVSQRMNTPLNRDAKSCLER W KNYLKPGIHK GSLTEEEQRLVINLQATHGN W KKIAAQVPGRTAKRLGK W [W]-x(19)-[W]-x(21)-[W]-x(12)-[L]-x(18)-[W]-x(18)-[W]
Glyma19g4065 0.1	-		RSLSGSSEEEIMSITKGPWTEEEDSVLFNYITVHGELKRTGKSCRL W LNYLRPNV RRGNITLEEQLLIDLHSRWGNRW W SKIAEHLPGRTDNEIKNY W [W]-x(30)-[W]-x(19)-[W]-x(12)-[I]-x(18)-[W]-x(18)-[W]
Glyma19g4067 0.1	-		KKARSIVLDCEEEIRKGPWSVEEDTILQNHVATHGDGLKRSKGKSCRL W LNYLRP DVRNGNITLQEQITILELHSRWGNRW W SKIAEHLPGRTDNEIKNY W [W]-x(30)-[W]-x(19)-[W]-x(12)-[I]-x(18)-[W]-x(18)-[W]
Glyma17g3502 0.1	-		MGRAPCCDKANVKRGP W SPEEDATLKNYVETHGTGLRRCGKSCRL W LNYLRP DIKHGG F TEEDDNIICLYAQMGSR W SAIASKLPGRTDNDVKNY W NTLKKKIMA RKVTLKTLTHNDTLPSTSTP -[W]-x(30)-[W]-x(12)-[F]-x(18)-[W]-x(18)-[W]

ANEXO E

Comparação dos subgrupos MYB de soja com os identificados por Du et al., (2012)

Du et al., (2012)		Diferenças observadas
Subgrupo	AtMYB/GmMYB	
C1 (S6)	AtMYB75, AtMYB114, AtMYB113, AtMYB90 GmMYB96, GmMYB80, GmMYB227, GmMYB8, GmMYB241, GMYB155, GmMYB192	Também agrupou Glyma9g36970
C2 (S15)	AtMYB82, AtMYB66, AtMYB0, AtMYB23	Sem AtMYB82, como observado em DUBOS et al., 2010
C3	AtMYB5	Agrupou também Glyma14g24500 e Glyma13g09980
C4	GmMYB47, GmMYB115, GmMYB58, GmMYB127	Também agrupou GmMYB33 e GmMYB195
C5	GmMYB45, GmMYB24, GmMYB103	Sem diferenças
C6 (S5)	AtMYB123 GmMYB191, GmMYB193, GmMYB205	Apenas AtMYB123 foi agrupado, os demais GmMYB foram agrupados em ?1
C7	GmMYB33, GmMYB195, GmMYB136, GmMYB175, GmMYB178, GmMYB194	GmMYB33 e GmMYB195 foram agrupados em C4
C8	GmMYB170, GmMYB203, GmMYB177, GmMYB211	Sem diferenças
C9 (S4)	AtMYB3, AtMYB8, AtMYB7, AtMYB4, AtMYB32 GmMYB6, GmMYB141, GmMYB2, GmMYB5, GmMYB42	Também agrupou AtMYB6, Glyma4g00550 e Glyma8g00820
C10 (S7)	AtMYB11, AtMYB12, AtMYB111 GmMYB176, GmMYB190, GmMYB200, GmMYB174, GmMYB204, GmMYB197, GmMYB198	Também agrupou Gmas53090519
C11	AtMYB20, AtMYB43, AtMYB85, AtMYB42 GmMYB98, GmMYB122, GmMYB116, GmMYB129, GmMYB97, GmMYB114, GmMYB101, GmMYB126	Agrupo também AtMYB99, AtMYB40, GmMYB124, GmMYB238, GmMYB239
C12	AtMYB40 GmMYB124, GmMYB238, GmMYB239	Foram agrupados em C11
C13	AtMYB103 GmMYB163, GmMYB232, GmMYB165, GmMYB209	AtMYB103 foi agrupado em C32
C14 (S13)	AtMYB55, AtMYB50, AtMYB61, AtMYB86 GmMYB185, GmMYB213, GmMYB184, GmMYB9, GmMYB187, GmMYB94, GmMYB39, GmMYB23, GmMYB214, GmMYB189, GmMYB186, GmMYB57, GmMYB162	Sem GmMYB94 e GmMYB189 (sem domínio R2R3)
C15	AtMYB26, AtMYB6 GmMYB107, GmMYB121, GmMYB102, GmMYB65, GmMYB85	Também agrupou Glyma15g02950
C16	AtMYB46, AtMYB83 GmMYB20, GmMYB188, GmMYB179, GmMYB202, GmMYB15, GmMYB182, GmMYB118, GmMYB180	Sem GmMYB188 e GmMYB15 (sem domínio R2R3)
C17 (S2)	AtMYB13, AtMYB14, AtMYB15 GmMYB132, GmMYB201, GmMYB158, GmMYB159, GmMYB215, GmMYB223	Também agrupou Glyma10g30860, Glyma19g43740, Glyma3g41100, Glyma6g45540, Glyma12g11390, Glyma6g45550, Glyma12g11340, Glyma6g45570

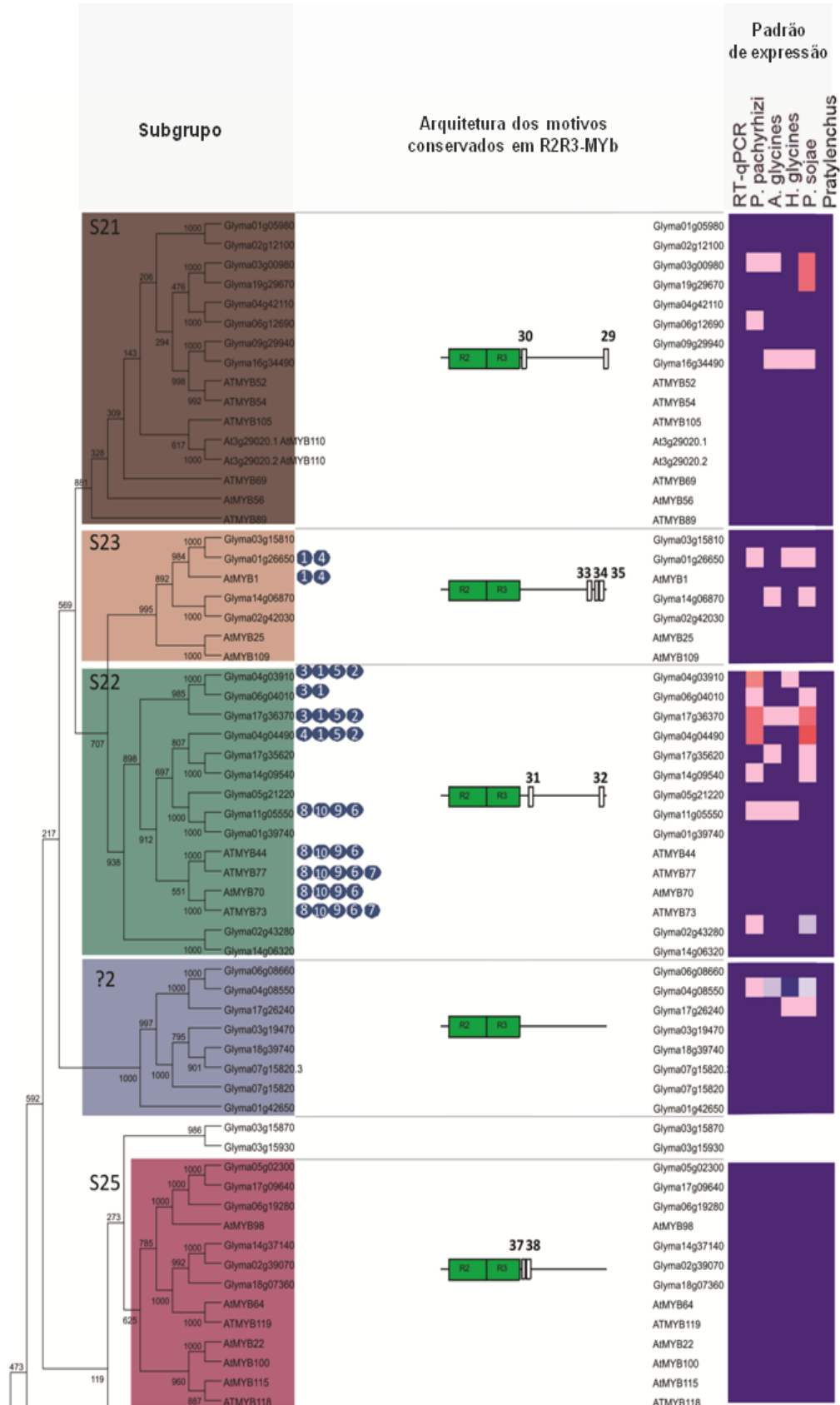
ANEXO E
(continuação)

Du et al., (2012)		
Subgrupo	AtMYB/GmMYB	Diferenças observadas
C18	GmMYB199, GmMYB22, GmMYB134	Todos agrupados em C17
C19 (S3)	AtMYB72, AtMYB10, AtMYB58, AtMYB63 GmMYB66, GmMYB123	Sem diferenças
C20 (S16)	AtMYB18, AtMYB19, AtMYB45 GmMYB25, GmMYB148, GmMYB145, GmMYB152	Sem diferenças
C21	GmMYB233, GmMYB142, GmMYB149, GmMYB150, GmMYB62, GmMYB34, GmMYB35	Agrupado em C17
C22	AtMYB17 GmMYB71, GmMYB95, GmMYB196, GmMYB206, GmMYB169, GmMYB212	Agrupados em C23
C23 (S9)	AtMYB16, AtMYB106 GmMYB77, GmMYB70, GmMYB93, GmMYB92, GmMYB63, GmMYB52	Também agrupou AtMYB17 como em Dubos et al., 2010
C24 (S12)	AtMYB28, AtMYB76, AtMYB29, AtMYB34, AtMYB51, AtMYB122	Sem diferenças
C25	GmMYB157, GmMYB171	Também agrupou AtMYB47 e AtMYB95
C26 (S11)	AtMYB102, AtMYB41, AtMYB74 GmMYB17, GmMYB83, GmMYB210, GmMYB64, GmMYB61, GmMYB76, GmMYB79, GmMYB242, GmMYB117	Também agrupou AtMYB49
C27 (S24)	AtMYB94, AtMYB96, AtMYB92, AtMYB53 GmMYB54, GmMYB59, GmMYB48, GmMYB53	Sem GmMYB54 (sem domínio R2R3)
C28	AtMYB9 e AtMYB107	Agrupado em C29
C29 (S10)	GmMYB10, GmMYB69, GmMYB86	Também agrupou GmMYB29, GmMYB73, GmMYB50, GmMYB72
C30	GmMYB29, GmMYB73, GmMYB50, GmMYB72 AtMYB30, AtMYB96, AtMYB94, AtMYB60, AtMYB31	Agrupados em C29
C31 (S1)	GmMYB4, GmMYB46, GmMYB153, GmMYB207, GmMYB172, GmMYB173, GmMYB222, GmMYB160, GmMYB7	Também agrupou Glyma9g37340
C32	AtMYB35, AtMYB80 GmMYB119, GmMYB131, GmMYB89, GmMYB144	Também agrupou AtMYB103
C33 (S14)	AtMYB36, AtMYB37, AtMYB38, AtMYB87, AtMYB68, AtMYB84 GmMYB164, GmMYB218, GmMYB13, GmMYB220, GmMYB137, GmMYB219, GmMYB168, GmMYB208, GmMYB161, GmMYB40, GmMYB130, GmMYB224, GmMYB225, GmMYB221, GmMYB166, GmMYB216, GmMYB167, GmMYB217, GmMYB18, GmMYB38	GmMYB224, GmMYB225, GmMYB221 não apresentam R2R3; GmMYB130 foi agrupado em C37; Também agrupou Glyma1g40410 e GmaS53086590
C34	AtMYB79, AtMYB71, AtMYB121 GmMYB243, GmMYB244, GmMYB100, GmMYB183, GmMYB113, GmMYB126, GmMYB135, GmMYB108	GmMYB126 foi agrupado em C11
C35	AtMYB48, AtMYB59 GmMYB74, GmMYB37, GmMYB82	Sem GmMYB37 (Sem domínio R2R3)
C36	AtMYB27 GmMYB120, GmMYB146	Sem diferenças
C37 (S19)	AtMYB21, AtMYB24, AtMYB57 GmMYB27, GmMYB130, GmMYB43, GmMYB26,	Sem diferenças

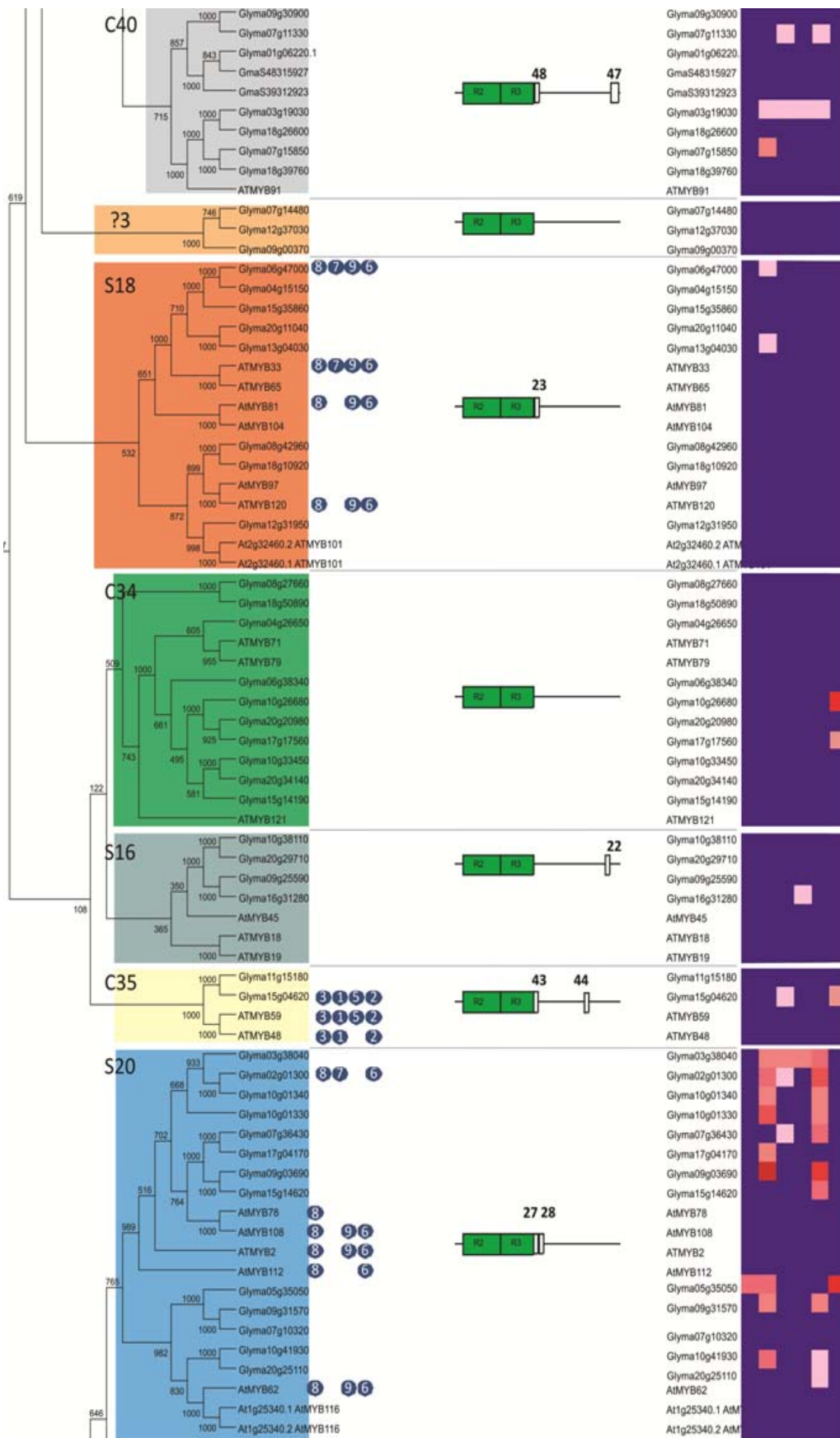
GmMYB125		
C38 (S20)	AtMYB62, AtMYB116, AtMYB112, AtMYB2, AtMYB78, AtMYB108	Sem GmMYB226 e GmMYB1 (sem domínio R2R3)
	GmMYB31, GmMYB133, GmMYB226, GmMYB28, GmMYB110, GmMYB14, GmMYB105, GmMYB21, GmMYB44, GmMYB104, GmMYB111, GmMYB1, GmMYB99, GmMYB16, GmMYB112	
C39 (S18)	AtMYB81, AtMYB104, AtMYB125, AtMYB33, AtMYB97, AtMYB120, AtMYB101, AtMYB65	GmMYB143 foi agrupado em ?_creme; sem GmMYB181 (sem R2R3); Também agrupou Glyma13g04030, Glyma20g11040, Glyma6g47000
	GmMYB55, GmMYB84, GmMYB143, GmMYB75, GmMYB181, GmMYB90, GmMYB109	
C40	AtMYB91	GmMYB140 foi agrupado em C9; Também foram agrupados Glyma1g06220, GmaS48315927, GmaS39312923
	GmMYB140, GmMYB151, GmMYB240, GmMYB30, GmMYB154	
C41 (S23)	AtMYB1, AtMYB25, AtMYB109	Também agrupou Glyma3g15810 e Glyma2g42030
	GmMYB81, gmMYB49	
C42 (S22)	AtMYB44, AtMYB77, AtMYB70, AtMYB73	Também agrupou Glyma17g36370
	GmMYB41, GmMYB3, GmMYB32, GmMYB36, GmMYB12, GmMYB11, GmMYB19, GmMYB231	
C43 (S21)	AtMYB52, AtMYB69, AtMYB89, AtMYB110, AtMYB56, AtMYB105, AtMYB117	Sem diferenças
	GmMYB68, GmMYB87, GmMYB56, GmMYB60, GmMYB230, GmMYB236, GmMYB228	
C44 (S25)	AtMYB64, AtMYB119, AtMYB118, AtMYB98, AtMYB115, AtMYB22, AtMYB100	Sem diferenças
	GmMYB229, GmMYB235, GmMYB88, GmMYB234	
C45	-	-
C46	-	-
C47	-	-

ANEXO F

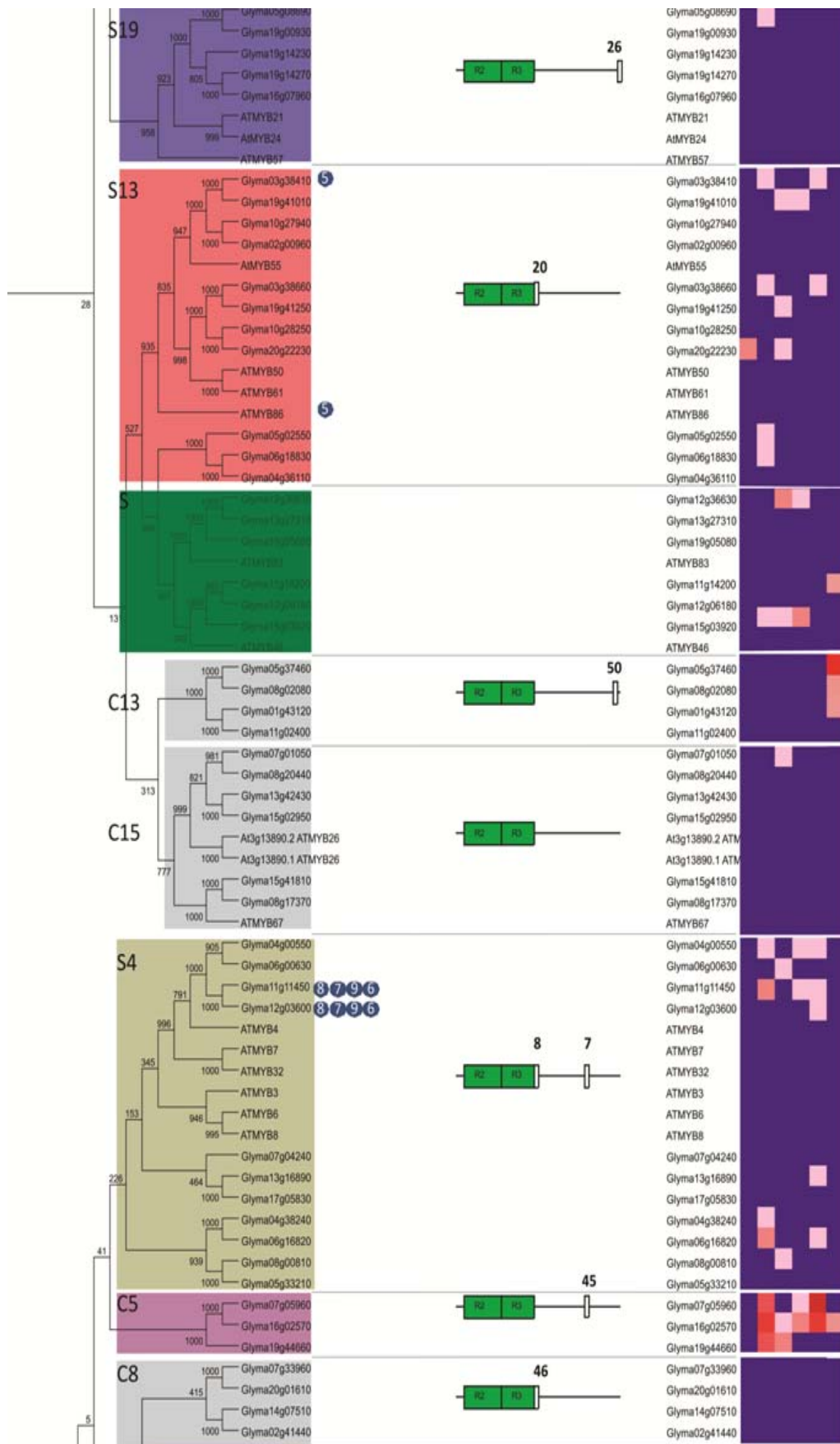
Versão completa da Figura 8: Relações Filogenéticas, classificação em subgrupos e motivos da região C-terminal de proteínas GmMYB e AtMYB.



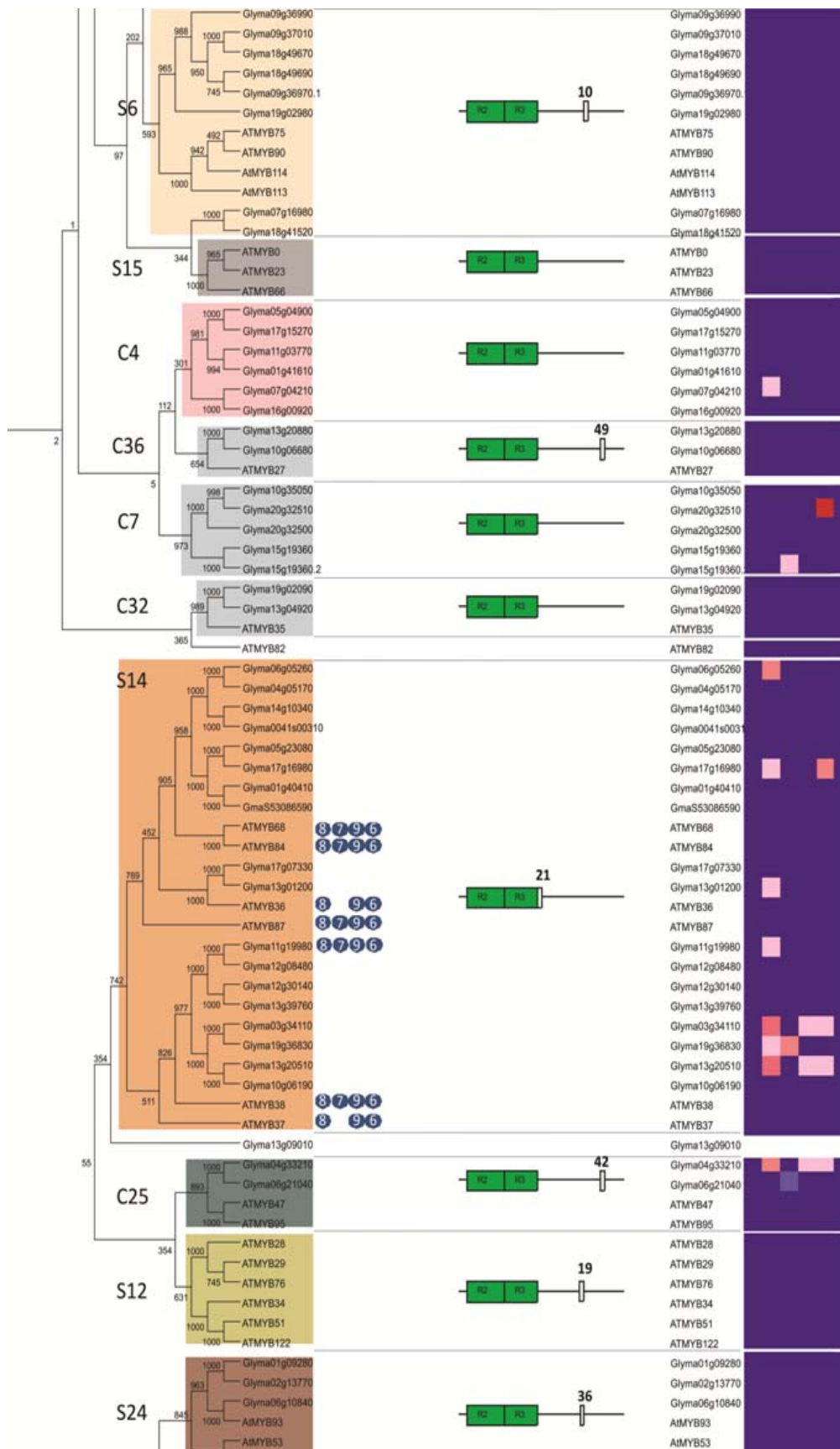
ANEXO F (Continuação)



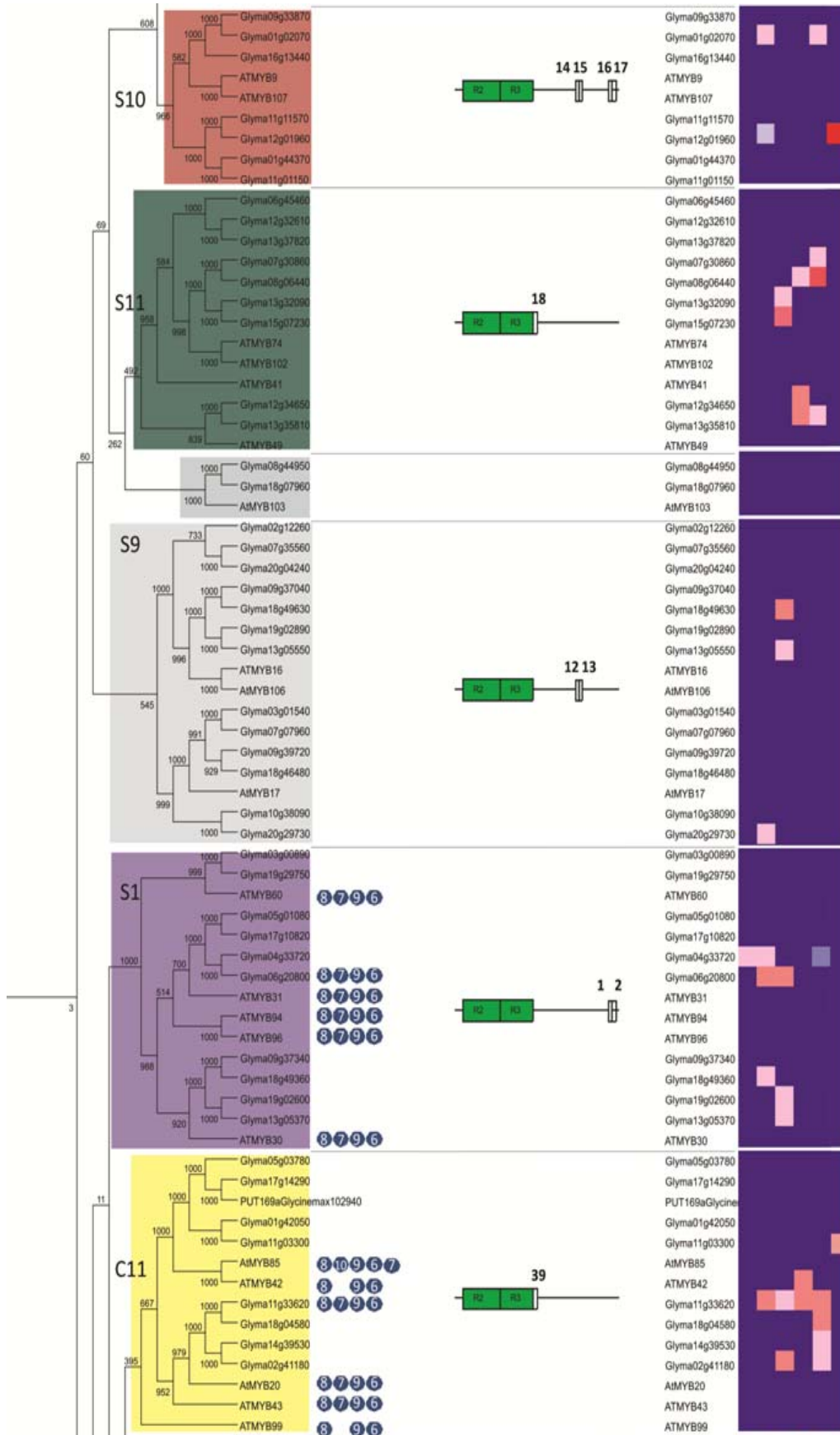
ANEXO F
(Continuação)



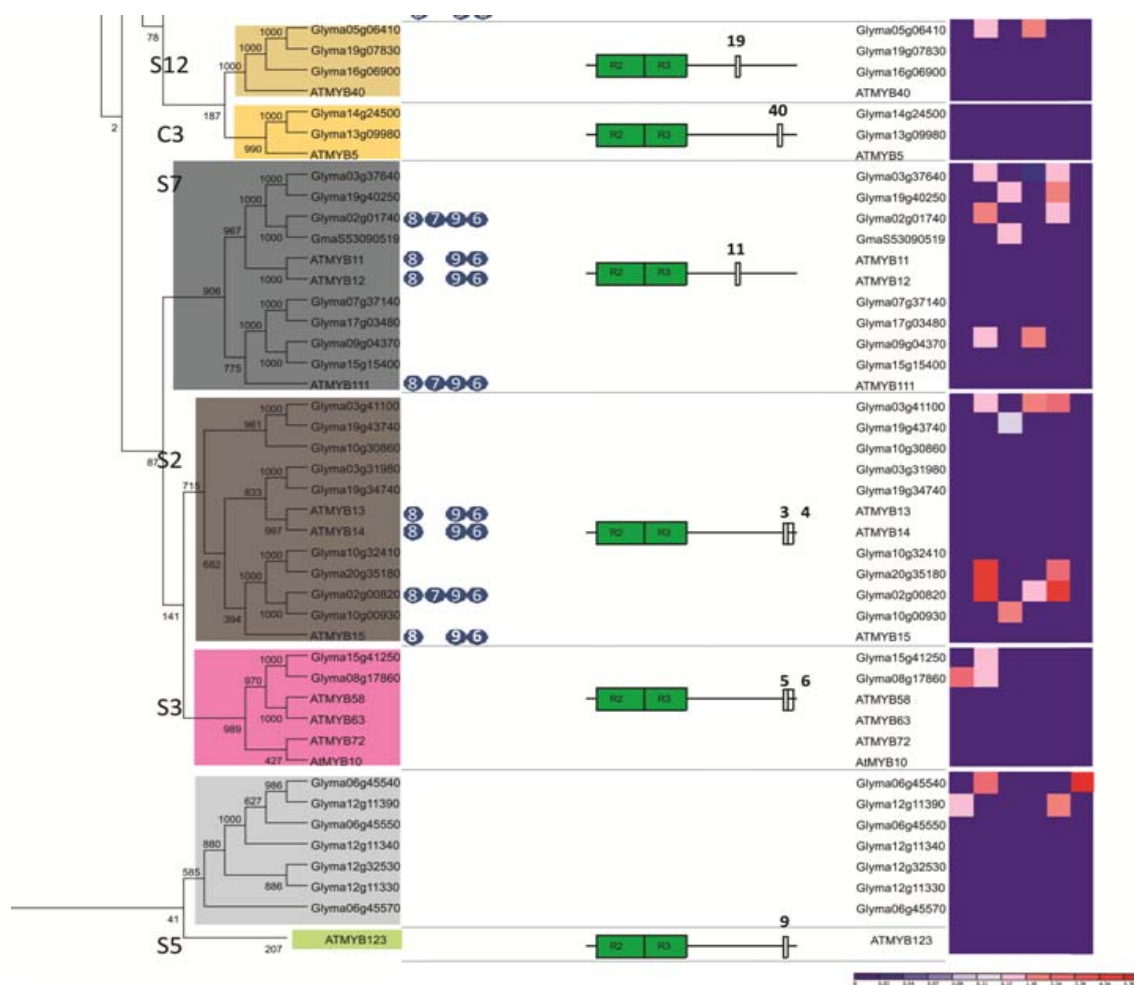
ANEXO F (Continuação)



ANEXO F
(Continuação)



ANEXO F (Continuação)



Primeira coluna: Árvore filogenética obtida pelo método *Neighbor-joining*, representa 89 sequências AtMYB (MYB-R2R3 de *Arabidopsis*) e 264 sequências de GmMYB (MYB-R2R3 de *G. max*). As sequências foram agrupadas em 43 subgrupos (designados pelas letras S, conhecidos em *Arabidopsis*, e C, identificados em soja). Os números dos ramos representam os valores de *bootstrap* de 1000 repetições. **Segunda coluna:** Estrutura dos R2R3-MYBs de soja e *Arabidopsis* (caixas verdes R2 e R3). Círculos azuis indicam os *cis*-motivos encontrados na região promotora (tabela 5). Caixas brancas representam os motivos conservados, detectados pelo programa MEME (Tabela 3). **Terceira coluna:** padrões de expressão dos GmMYBs em ensaios com os patógenos *P. pachyrhizi*, *A. glycines*, *H. glycines*, *P. sojae* e *P. brachyurus*. A primeira coluna do *Heatmap* representa ensaios com *P. pachyrhizi* (validação através de RT-qPCR), de GmMYBs selecionados (Tabela 7). Os demais valores de expressão provém dos maiores níveis de expressão obtidos a partir dos bancos de expressão Genevestigator e Genosoja (Anexo 2 e 3).