

SILVANA PINHEIRO DADALTO

**ESTUDO DA FAMÍLIA DE FATORES DE TRANSCRIÇÃO WHIRLY EM
SOJA**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica Aplicada, para obtenção do título de *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS-BRASIL
2016

**Ficha catalográfica preparada pela Biblioteca Central da
Universidade Federal de Viçosa - Câmpus Viçosa**

T

D121e Dadalto, Silvana Pinheiro, 1985-
2016 Estudo da família de fatores de transcrição Whirly em
soja / Silvana Pinheiro Dadalto. - Viçosa, MG, 2016.
viii, 44f. : il. (algumas color.) ; 29 cm.

Inclui anexo.

Orientador : Luciano Gomes Fietto.

Tese (doutorado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. Soja - Estresse. 2. Soja - Genes. 3. Soja - Aspectos
bioquímicos. I. Universidade Federal de Viçosa.
Departamento de Bioquímica e Biologia Molecular.
Programa de Pós-graduação em Bioquímica Aplicada.
II. Título.

CDD 22 ed. 633.34

SILVANA PINHEIRO DADALTO

**ESTUDO DA FAMÍLIA DE FATORES DE TRANSCRIÇÃO WHIRLY EM
SOJA**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica Aplicada, para obtenção do título de *Doctor Scientiae*.

APROVADA: 24 de novembro de 2016.

Lanna Clícia Carrijo

Leandro Licursi de Oliveira

Murilo Siqueira Alves
(Coorientador)

Pedro Augusto Braga do Reis

Luciano Gomes Fietto
(Orientador)

A todos aqueles que escolheram a ciência como propósito e passam suas vidas plantando sementes para o futuro.

Dedico.

AGRADECIMENTOS

Agradeço a Deus por ter me guiado e protegido, e por me colocar sempre no caminho de me melhorar a cada dia;

À Universidade Federal de Viçosa e ao departamento de Bioquímica e Biologia Molecular e à School of Biology – University of Leeds, pela infraestrutura que permitiu a realização deste trabalho, e a todos os que trabalham aqui, funcionários e professores;

À Coordenação de Aperfeiçoamento de Pessoal de nível Superior (CAPES) por ambas as bolsas concedidas, tanto no Brasil quanto pelo Programa de Doutorado Sanduíche no Exterior (PDSE);

Ao Luciano Gomes Fietto por ter sido um ótimo orientador desde a graduação, sempre ético, prático, inteligente e paciente; obrigada por saber como lidar com meu excesso de preocupação e por estar sempre comigo durante os desafios que o doutorado me trouxe;

Thank you to Christine Foyer, for accepted to supervise me during my time in Leeds;

Ao meu coorientador Murilo Siqueira Alves, por tudo o que me ensinou, por caminhar ao meu lado e por ser presente mesmo quando eu estava em outro país;

Aos professores que fizeram parte da banca pelas sugestões e participação;

Aos meus pais, Leila e Luiz, e irmãos, Tati, Juli, Artur, Renata, Drei, Bruno e Júnior, estrelas iluminadas colocadas em meu caminho; em especial à minha irmã Juliana, que sempre acreditou e apostou em meus sonhos, sem ela essa caminhada não seria possível;

Aos amigos do LBM pelo apoio e compreensão e pelos momentos de distração, em especial a Vanessa, Patrícia e Amanda, minhas “coorientadoras

não oficiais”, pelo carinho, amizade e todos os bons momentos de trabalho em equipe.

A todos os meus amigos, especialmente a Fernanda Prieto Bruckner, Juliana Prieto Bruckner, e Ana Cristina Amorim, pessoas especiais que me ajudaram imensamente. Sem vocês eu não teria conseguido!

I would like to say thank you to my friends in Leeds, for all professional and personal support everyone gave me during a difficult but rich moment of my life. I miss the “Brazilian Mafia” lunch time, “Travelers around the world” and “Noisy group”. For Robbie all my gratitude for everything he taught me and for taking care of me so well.

I also thank some special people who helped me at Leeds: Barbara, Daria, James, Danny, Stefanía, Ambra, Teresa, Yat and Blanca. Thank you for being serious and working without any noise, in a very quiet environment. In this crazy place I learned so much and I am glad that I could meet you.

Thank you very much! Muito obrigada!

BIOGRAFIA

Silvana Pinheiro Dadalto, filha de Luiz Roberto Dadalto e Leila Gean Costa Pinheiro Dadalto, nasceu na cidade de Eunápolis, Bahia, em 19 de setembro de 1985. Em dezembro de 2002 completou o ensino médio no Instituto Associados Saber e Cultura – IASC em Eunápolis-BA, após estudar por dois anos no Centro Federal de Educação Tecnológica - CEFET-BA, Unidade descentralizada de Eunápolis.

Em 2003 iniciou os estudos em Desenho Industrial na Universidade do Estado de Minas Gerais - UEMG, tendo interrompido o curso em 2004.

Em março de 2005 ingressou na Universidade Federal de Viçosa, concluindo o curso de bacharelado em Bioquímica em julho de 2010. Em agosto do mesmo ano, ingressou no Programa de Pós-Graduação em Bioquímica Agrícola, em nível de mestrado, na Universidade Federal de Viçosa com defesa de dissertação em 23 de julho de 2012. Em agosto de 2012 iniciou o doutorado em Bioquímica Aplicada na Universidade Federal de Viçosa. Participou do Programa de Doutorado Sanduíche no Exterior – PDSE, fomentado pela CAPES, de julho de 2015 até maio de 2016 na School of Biology, University of Leeds, Leeds-UK. Submeteu-se à defesa de tese em 24 de novembro de 2016.

SUMÁRIO

RESUMO	iv
ABSTRACT	v
INTRODUÇÃO	1
Capítulo I: Do cloroplasto ao núcleo: papel das proteínas Whirly durante a resposta a estresse biótico em plantas.	3
StWHY1 e AtWHY1 – atividade de ligação ao DNA em resposta a patógenos	8
Família Whirly em cloroplastos e mitocôndrias	9
Proteínas Whirly nucleares e alterações transcricionais.....	13
REFERÊNCIAS	20
Chapter II: Characterisation of the WHIRLY family in soybean: differential regulation of <i>GmWHY1A</i> mRNAs expression by environmental stresses and phytohormones.	24
ABSTRACT	25
Keywords	25
INTRODUCTION.....	26
CONCLUSION	39
ACKNOWLEDGMENTS	39
REFERENCES.....	40
CONCLUSÕES GERAIS	43
Anexo I	44
Tabela suplementar do capítulo 2.	44

RESUMO

DADALTO, Silvana Pinheiro, D.Sc., Universidade Federal de Viçosa, novembro de 2016. **Estudo da família de fatores de transcrição Whirly em soja.** Orientador: Luciano Gomes Fietto. Coorientador: Murilo Siqueira Alves.

Plantas submetidas a estresse possuem complexos mecanismos de defesa, que resultam na expressão de proteínas relacionadas à defesa ou a um estado de adaptação. Esta mudança ocorre com a participação de fatores de transcrição, proteínas que regulam a transcrição de determinados genes. A família Whirly compreende fatores de transcrição que possuem a capacidade de se ligar ao ssDNA e apresentam-se na forma de tetrameros. Esta família tem papel na proteção ao DNA e formação de nucleóides em cloroplastos e mitocôndrias. Outra característica importante é o transporte das proteínas Whirly do cloroplasto para núcleo, em um processo conhecido como sinalização retrógrada, que está relacionado à resposta a estresses como excesso de luz e ataque por patógenos. Em soja não há, até o momento, nenhum estudo sobre esta importante família. Assim este trabalho visa o estudo desta família em soja, e traz no capítulo 1 uma revisão sobre a família Whirly com ênfase a resposta a estresses bióticos. No capítulo dois a família Whirly de soja é descrita e a expressão de GmWHY1A é analisada durante resfriamento e após infecção por *Phakopsora pachyrhizi*. A família Whirly em soja é composta por 18 membros, codificados por sete *loci*. Os membros foram nomeados com base na similaridade com proteínas Whirly de *Arabidopsis thaliana* e em análises *in silico*. Foram divididos em três grupos: 1- com localização predita para o cloroplasto; 2- preditas como mitocondriais, e 4- formado por proteínas com localização indeterminada. Cada gene teve sua expressão mensurada em raízes, caule, cotilédones, folhas e trifólios. GmWHY1A foi escolhido para estudos posteriores pelas suas características de expressão e pela similaridade com AtWNHY1. GmWHY1 foi reprimido após tratamento com etefon e foi positivamente regulado após tratamento com Ácido salicílico e jasmonato. Durante o estresse por frio e após a infecção por *Phakopsora pachyrhizi*, GmWHY1A foi reprimido. Tomados em conjunto esses dados sugerem que a expressão de Whirly pode ser regulada através do contato com o fungo.

ABSTRACT

DADALTO, Silvana Pinheiro, D.Sc., Universidade Federal de Viçosa, November, 2016. **Study of transfactor family Whirly in soybean.** Advisor: Luciano Gomes Fietto. Co-advisor: Murilo Siqueira Alves.

When submitted to stress, plants triggers complex defense pathways that result in expression of defense-related proteins. This change occurs with the participation of transcription factors, proteins that regulates the transcription of specific genes. The Whirly family comprises transcription factors which have the ability to bind single-stranded DNA and remains as tetramers. This family has a role in DNA protection and formation of nucleoids in chloroplasts and mitochondria. Another important feature is the transport of the Whirly proteins from chloroplast to nucleus in a process known as retrograde signaling, which is related to the response to stresses such as high light and pathogen contact. So far there is no study of this important family in soybean. Thus, this work aims to study this family in soybean. The chapter 1 is a review on the Whirly family with emphasis on biotic stress response. In chapter two, the soybean Whirly family is described and the expression of GmWHY1A is analyzed during chilling stress and after infection by *Phakopsora pachyrhizi*. The Whirly family in soybean comprises 18 members, coded by seven *loci*. Members were named based on its similarity to Whirly proteins of *Arabidopsis thaliana* and *in silico* analyses. They were divided into three groups: 1- with chloroplast predicted location; 2 - predicted as mitochondrial, and 4 - formed by proteins with undetermined location. Each gene had its expression measured in roots, stem, cotyledons, leaves and trifoliolate leaves. GmWHY1A was chosen for further studies for its expression characteristics and its similarity to AtWNHY1. GmWHY1 was repressed after treatment with ethephon and was upregulated following treatment with salicylic acid and jasmonic acid. During chilling stress and after infection by *Phakopsora pachyrhizi*, GMWHY1A was suppressed. Taken together, these data suggest that Whirly expression can be regulated by contact with this fungus.

INTRODUÇÃO

Os estresses bióticos e abióticos são os principais fatores que afetam a produção de importantes culturas agrícolas. Um bom exemplo do dano que um estresse pode causar é a infecção da soja pelo fungo *Phakopsora pachyrhizi*, que pode levar a 80% de perdas durante uma epidemia severa (Cabreira *et al.*, 2013).

Durante uma situação de estresse são ativadas vias de sinalização que resultam numa reprogramação da expressão gênica, favorecendo a expressão de genes de defesa. Essas vias são complexas e interdependentes, a ativação de uma via pode influenciar outras vias de defesa ou de desenvolvimento das plantas. É característica deste processo a presença de hormônios vegetais, como auxinas, ácido salicílico, ácido abscísico, etileno e ácido jasmônico (Dersken *et al.*, 2013). Em resposta aos estresses bióticos, é comum haver uma cascata de sinalização que resulta na produção de proteínas PR (*Pathogen Related*) e envolvem diversas organelas celulares, como as mitocôndrias e os cloroplastos (Moore *et al.*, 2011).

Os cloroplastos funcionam como sensores ambientais, que guiam a planta a um estado de adaptação. As flutuações das condições ambientais são percebidas através de importantes “sensores moleculares”. Um dos mais importantes “sensores” é o estado redox da cadeia fotossintética de transporte de elétrons, que pode ser medida pela concentração de espécies reativas de oxigênio (ROS) produzidas em uma dada condição celular (Huner *et al.*, 2012).

Em plantas, ROS agem como moléculas sinalizadoras, principalmente em condições de estresse; durante estresses a produção de ROS é intensificada. O aumento de ROS na célula pode acarretar oxidação e peroxidação de componentes celulares, e inclusive pode levar à morte celular (Triantaphylidés *et al.*, 2008, Torres, 2010).

Durante a resposta ocorre uma comunicação entre os cloroplastos e o núcleo, chamada de sinalização retrógrada, que leva à transcrição, no núcleo, de genes que estão relacionados à defesa ou a adaptação das plantas ao estresse. Esta regulação é possível graças à presença de moléculas chamadas de fatores de transcrição (Xiao *et al.*, 2013).

Os fatores de transcrição que estão diretamente envolvidos na resposta a estresses regulam de maneira positiva ou negativa genes que participam de vias de sinalização que levam à adaptação a essas condições. A existência de fatores de transcrição móveis envolvidos na sinalização retrógrada confirma a complexidade dos mecanismos de regulação genica. (Estavillo *et al.*, 2013).

Os principais fatores de transcrição ligados à resposta de plantas ao ataque de patógenos estão agrupados principalmente nas famílias bZIPs, AP2/ERF, MYB, WRKY e Whirly (WHY) (Alves *et al.*, 2013). Destas, a família Whirly é a menos estudada neste contexto. Este trabalho traz primeiramente uma revisão sobre as informações disponíveis no momento sobre a família Whirly, no contexto da resposta a estresse biótico em plantas. No capítulo 2, a família Whirly em soja é descrita e a expressão de seu principal membro, GmWHY1A, é analisada durante resfriamento, após infecção por *Phakopsora pachyrhizi* e após tratamento com ácido salicílico, ácido jasmonico e etefon.

Capítulo I: Do cloroplasto ao núcleo: papel das proteínas Whirly durante a resposta a estresse biótico em plantas.

Fatores de transcrição são proteínas que possuem a capacidade de regular, de maneira positiva ou negativa, genes específicos. Estão diretamente envolvidos na resposta a estresses e participam de vias de sinalização que levam à adaptação a essas condições. As principais famílias de fatores de transcrição ligados à resposta de plantas ao ataque de patógenos são: bZIPs, AP2/ERF, MYB, WRKY e Whirly (WHY) (Alves et al., 2013). Destas, a família Whirly é a menos estudada.

Proteínas Whirly são fatores de transcrição que se ligam ao ssDNA. A primeira proteína da família Whirly a ser caracterizada foi StWHY1, inicialmente chamada de PBF-2. Em batata, PBF-2 (PR10 α Binding Fator) regula a transcrição do gene PR10a (Matton *et al.*, 1993, Desveaux *et al.*, 2000), que é expresso após injúria e em resposta ao tratamento com elicitores ou pelo contato com o patógeno *Phytophthora infestans* (Matton e Brisson, 1989).

Solanum tuberosum expressam duas proteínas Whirly: StWHY1 e StWHY2. Sequências similares as de StWHY1 foram encontradas em diversas espécies de plantas, incluindo arroz, milho, *Arabidopsis thaliana*, cevada e soja (Krause e Krupinska, 2009). O domínio SSB (single strand binding) é a região mais conservada, e foi denominado de domínio Whirly por Desveaux *et al.* (2004). *A. thaliana* possuem três proteínas Whirly: AtWHY1 e AtWHY3, que são mais similares entre si e estão presentes nos cloroplastos, e AtWHY2, presente em mitocôndrias. A comparação de suas estruturas com as proteínas Whirly de batatas revela que StWHY1 é similar a AtWHY1 e AtWHY3, enquanto AtWHY2 é similar a StWHY2. Resíduos envolvidos no domínio SSB são conservados ou substituídos por resíduos com características semelhantes (Capadócia *et al.*, 2013, Krause *et al.*, 2005).

StWHY1 é encontrada em forma tetramérica em extratos nucleares de tubérculos tratados e não tratados com elicitores em quantidades semelhantes, porém, em extratos não tratados, a ligação ao DNA só ocorre após cromatografia de troca aniônica, sugerindo que há a presença de um regulador negativo em tubérculos não tratados (Desveaux *et al.*, 2000).

A ativação da expressão de PR10a por StWHY1 em batatas se dá de maneira dependente do cis-elemento ERE (*Elicitor Response Element*), presente na região promotora de PR10a, e o processo é regulado por uma

proteína cinase C (Subramanian *et al.*, 1997). Foi verificado que cada tetrâmero liga-se a uma única molécula de ssDNA no motivo ERE (Desveaux *et al.*, 2002). O elemento GTCAAAAAT, chamado de PB core (*PBF-2 binding core*) é suficiente para a atividade ótima do cis-elemento (Desveaux *et al.*, 2004).

O tetrâmero possui simetria C₄. Cada protômero consiste em duas superfícies formadas por quatro folhas beta cada, e três alfas-hélices. A região C-terminal contém um motivo hélice-volta-hélice, região que apresenta aminoácidos que estão em contato com outros protômeros. As alfas hélices convergem para o centro do tetrâmero, formando um poro hidrofílico contra o qual as folhas-beta desviam-se. Assim, o tetrâmero forma uma estrutura quaternária que se parece com um catavento (*Wirligig*, em inglês), daí a sugestão do nome Whirly para esta família de proteínas (Figura1). A porção N-terminal compreende uma região aparentemente desordenada que se estende da ponta da primeira folha beta, ate o topo das alfas hélices centrais. (Desveaux *et al.*, 2002).

A região entre os aminoácidos 86 e 109 contém uma sequência KGKAAL que tem importância para a ligação ao ssDNA (domínio de ligação ao ssDNA). A mudança dessa sequência para QGQGGV não afeta a tetramerização, mas abole a ligação ao DNA. Na porção N-terminal do domínio de ligação ao ssDNA há uma região que contém onze glutaminas (Q) consecutivas. In vivo, o tetrâmero pode se ligar ao dsDNA dissociado, e ativar a expressão de PR10a através da região poliglutamínica. A simetria de PBF-2 reorienta o DNA em 90°, para o protômero adjacente, e isso pode favorecer o aparecimento de uma bolha de DNA dissociado (Figura 2) (Desveaux *et al.*, 2002).

Ao se ligar ao DNA, StWHY2 não sofre grandes mudanças conformacionais. A presença de rotações abruptas indica que a ligação acontece somente ao ssDNA. A interação ocorre entre as bases e resíduos hidrofóbicos e envolve sete ligações de hidrogênio e uma interação mediada por água. Poucas interações sequencia-específica entre StWHY2 e ERE são observadas. As estruturas observadas para StWHY1, StWHY2, AtWHY1, AtWHY2 e AtWHY3 são similares e aminoácidos-chave para a ligação ao ssDNA também são conservados. Isso suporta a suposição de que estas

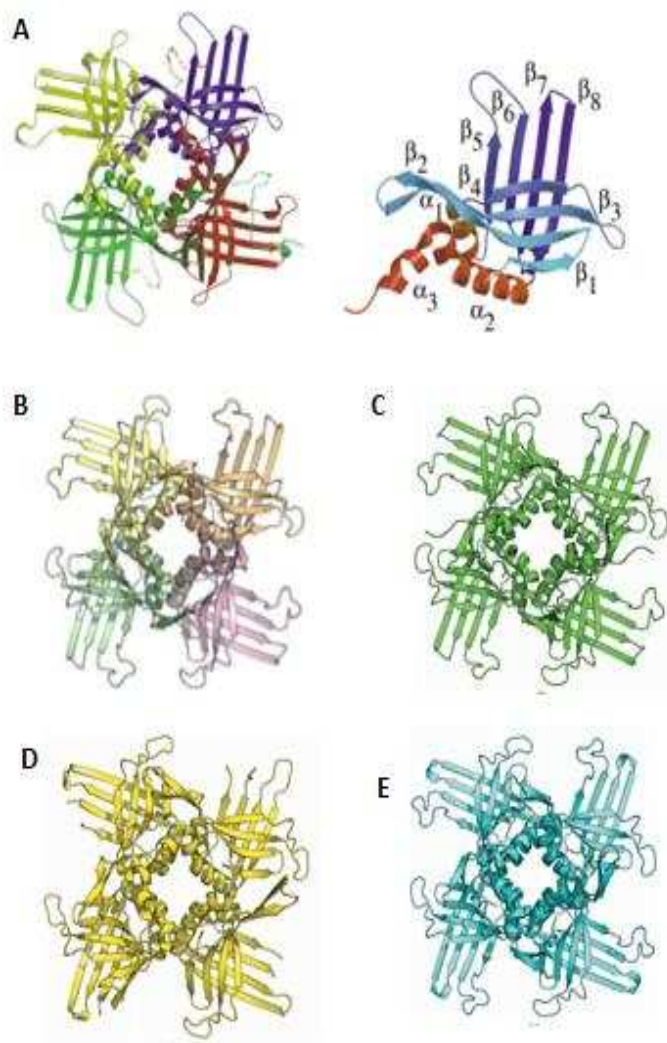


Figura 1: Estruturas obtidas para proteínas Whirly. A- StWHY1 (tetramérica e monomérica) (Deseveaux *et al.*, 2002a), B-StWHY2 (Cappadocia *et al.*, 2010), C- AtWHY1, D-AtWHY2, E-AtWHY3 (Cappadocia *et al.*, 2013).

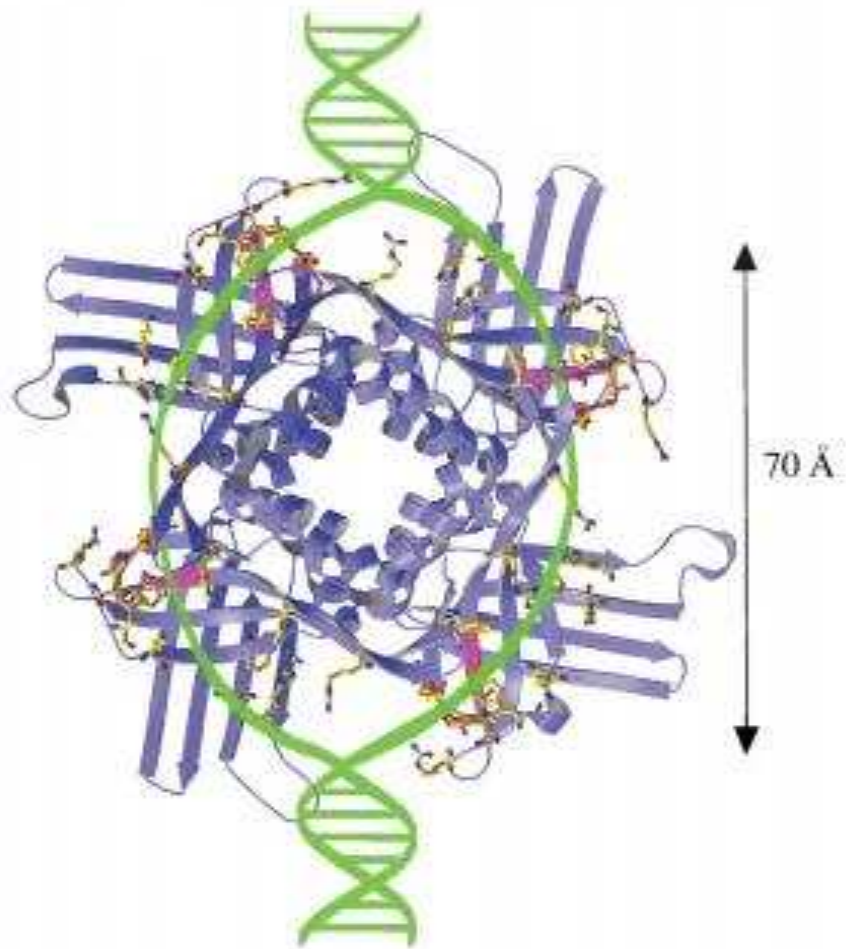


Figura 2: Ligação de tetrâmero de Whirly ao DNA (Desveaux *et al.*, 2002). A ligação ao ssDNA não resulta em grandes alterações conformacionais. Uma rotação de 90° propicia o aparecimento de uma bolha de DNA dissociado.

proteínas são intimamente relacionadas e podem ter função comum em cada organela. (Cappadocia *et al.*, 2010).

StWHY1 e AtWHY1 – atividade de ligação ao DNA em resposta a patógenos

StWHY1 apresenta maior atividade de ligação ao promotor de PR10a após o tratamento com elicitor ácido araquidônico que o observado na injúria. Porém, não foi observado ligação quando proteínas foram extraídas de tubérculos não tratados (Desveaux *et al.*, 2004). A menor ligação de StWHY1 ao DNA observada durante injúria corresponde aos menores níveis de RNA obtidos para o gene PR10a após esses tratamentos, o que sugere o controle de PR10a por StWHY1 (Matton e Brisson, 1989).

Maleck *et al.* (2000) realizaram uma busca pelo PB core utilizando dados de microarranjo de Arabidopsis em promotores de genes relacionados à defesa contra patógenos. Os autores verificaram um enriquecimento de 3,6 vezes em 11/26 genes associados a SAR. Considerando que este elemento pode ter papel na regulação destes genes, Desveaux *et al.* (2004) analisaram a atividade de ligação de AtWHY1 ao DNA, e a relação desta proteína com a via de SA.

Extratos nucleares de dois Mutantes obtidos através de TILLING (*Targeting Induced Local Lesions IN Genomes*, McCallum *et al.*, 2000) demonstraram menor atividade de ligação ao DNA em relação ao extrato selvagem. O mutante *atwhy1.1* carrega a mutação Pro183Ser, que resulta numa falha no motivo Folha- β que margeia o domínio de ligação ao ssDNA, e assim afeta a interação ao DNA. O mutante *atwhy1.2* possui a mudança Gly148Glu, localizada na alfa hélice central, que resulta na abolição da formação dos tetrâmeros, uma vez que essa porção é importante para a interação entre os monômeros (Desveaux *et al.*, 2004).

O tratamento com AS induz a atividade de ligação ao DNA da proteína AtWHY1 (Desveaux *et al.*, 2004). Além disso, AtWHY1 e AtWHY3 são 3 e 4 vezes mais expressos após tratamento com SA (Xiong *et al.*, 2009). O tratamento com SA em mutantes *atwhy1.2* resultou em menor atividade de ligação ao DNA. A quantificação de ligação de AtWHY1 ao DNA em plantas *npr1-1* (deletadas para o gene NPR1, que codifica uma proteína-chave

durante a sinalização por ácido salicílico) foi similar àquela obtida para planta selvagem (Desveaux *et al.*, 2004). Assim, presume-se que a atividade de ligação ao DNA da proteína AtWHY1 ocorre de maneira independente de NPR1.

A expressão do gene PR1 é pouco induzida em plantas *atwhy1.1* e não detectada em plantas *atwhy1.2* e em plantas *npr1-1*. Os autores demonstraram que para a completa indução da expressão de PR1 tanto NPR1 quanto AtWHY1 são necessários (Desveaux *et al.*, 2004).

A infecção de *Arabidopsis* com *Peronospora parasitica* isolado *Emoy2* (incompatível), e isolado *Noco2* (compatível), implica em um aumento da suscetibilidade das plantas *atwhy1.1* e *atwhy1.2* ao isolado *Noco2*. Durante a infecção pelo isolado incompatível *Emoy2*, *atwhy1.2* apresenta um fenótipo intermediário, com traços de necrose, alto crescimento nas hifas dos fungos, além de um aumento na contagem de esporangióforos. Quando a infecção ocorre após tratamento com SA, a resistência ao isolado *Emoy2* é completamente atingida para plantas selvagens, enquanto continua comprometida em plantas *atwhy1.2*. Estes dados demonstram a importância de AtWHY1 para o completo estabelecimento da SAR (Desveaux *et al.*, 2004).

Assim, AtWHY1 é importante para a expressão de PR1 e está relacionado à via de sinalização por ácido salicílico de maneira independente de NPR-1. Proteínas AtWHY1 não funcionais resultam em diminuição da ligação ao DNA e aumento da suscetibilidade a patógenos. Seria interessante determinar se a superexpressão de Whirly aumenta a resistência da planta a patógenos, e quais patógenos teriam sua infecção diminuída (Desveaux *et al.* 2005).

Família Whirly em cloroplastos e mitocôndrias

A maior parte dos genes codificados no DNA plastidial codificam proteínas relacionadas à fotossíntese (Green, 2011). Assim, danos ocorridos no ptDNA resultam em perda da eficiência na cadeia transportadora de elétrons, e conseqüentemente no aumento da produção de ROS. Dessa forma, a manutenção do genoma plastidial é de suma importância para a

célula, e vários estudos sobre a estabilidade do genoma plastidial vêm sendo realizados (Lepage *et al.*, 2013).

Um tipo de dano que ocorre no ptDNA são os DSBs (*Double strand breaks*) quebras na fita dupla, que podem resultar em pareamentos inadequados como consequência de colapsos na forquilha de replicação, ou pelo tratamento com ciprofloxacina (CIP) (Parent *et al.*, 2011). Em cloroplastos, PollB, uma DNA polimerase tipo I, possui papel no reparo do ptDNA, assim como proteínas Whirly, que atuam na estabilização do ptDNA, a partir da ligação ao ptDNA fita simples, limitando a ocorrência de pareamentos incorretos e consequentemente diminuindo a ocorrência de uma via não conservativa de reparo chamada MMBIR (*microhomology-mediated breakinduced replication*), replicação induzida pela quebra e mediada por microhomologia (Parent *et al.*, 2011, Marechal *et al.*, 2009).

Duplos *knok-out* para os genes AtWHY1 e AtWHY3 submetidos ao tratamento com indutores de DSB se mostraram mais sensíveis ao tratamento, que resulta em plantas variegadas e estioladas, e apresentaram maior número de rearranjos no ptDNA causados por MMBIR que plantas selvagens em *A. thaliana* e *Zea mays* (Marechalet *et al.*, 2009, Cappadocia *et al.*, 2010).

A instabilidade do genoma plastidial resulta em impacto negativo na eficiência fotossintética e leva a geração de ROS no cloroplasto. Mutantes *why1why3* e *why1why3pollb-1* resultam em fenótipo variegado branco em *Arabidopsis*, mas os mutantes *why1why3pollb-1* também apresentam, além da variegação branca, um fenótipo variegado amarelo, menor capacidade fotossintética, e são mais suscetíveis ao tratamento com CIP (Lepage *et al.* 2013). Os mutantes *why1why3pollb-1* apresentam produção de ROS cerca de duas vezes maior que plantas selvagens, e o tratamento com CIP intensifica o fenótipo. O crescimento em condições de luz amena não restabelece a eficiência fotossintética. Porém, o aumento de ROS não está relacionado a um aumento na morte celular, e plantas *why1why3pollb-1* apresentam cloroplastos com aparência normal, e funcionais (Lepage *et al.* 2013).

A ausência da proteína mitocondrial WHY2 também resulta em aumento de rearranjos no DNA mitocondrial após a indução de DSB. A ausência de

proteínas Whirly na mitocôndria não altera o DNA plastidial e viceversa. As alterações no mtDNA e ptDNA são predominantemente do tipo MHMR, rearranjos de DNA mediados por microhomologias (Maréchal *et al.*, 2008, Cappadocia *et al.*, 2010). Em plantas de couve-flor, proteínas mitocondriais WHY2 são encontradas em complexos que contem ODB1 (proteína relacionada à estabilidade mitocondrial). A interação é mediada por DNA ou RNA e está relacionada à prevenção de DSB (Janicka *et al.*, 2012).

Experimentos *in vitro* mostram que WHY2 em contato com a dsDNA pode deslocar o equilíbrio no sentido da formação de ssDNA. Isso pode ser importante durante o reparo de DSB, que implica numa etapa onde as bases de uma das fitas do DNA danificado são retiradas por exonucleases, e o DNA permanece parcialmente como ssDNA, ligado a WHY2. Além da capacidade de ligação ao ssDNA, foi verificado que a presença de WHY2 efetivamente protege o DNA de degradação por nucleases (Cappadocia *et al.*, 2010).

A superexpressão de AtWHY2 está relacionada com a diminuição da expressão do mtDNA, comprometendo a função mitocondrial, resultando em um fenótipo caracterizado por crescimento diminuído e senescência em folhas. A associação entre AtWHY2 e o mtDNA parece ocorrer não de forma sítio específica, e sim em todo o DNA mitocondrial. Essa associação é conservada em *Arabidopsis thaliana* e *Zea mays* (Maréchal *et al.*, 2008). Considerando o fenótipo semelhante entre plantas *why1* e plantas superexpressando AtWHY2, discute-se se AtWHY1 e AtWHY2 não poderiam ter funções antagônicas em cloroplastos e mitocôndrias (Miao *et al.*, 2013). Porém, fenótipos diferentes foram observados para os mutantes *why1*. Yoo *et al.* (2007), por exemplo, não verificou alteração do fenótipo após a mutação, enquanto Miao *et al.* (2013) obteve fenótipo de senescência para as mesmas linhagens de mutantes. Outros fenótipos observados foram os variegados, nesse caso com a deleção para as proteínas AtWHY1 e AtWHY3 (Lepage *et al.*, 2013) e o aumento da sensibilidade de sementes ao hormônio ácido abscísico (ABA) promovido apenas por AtWHY1 plastidial (Isemer *et al.*, 2012).

Aparentemente deve haver um equilíbrio na quantidade de StWHY2 mitocondrial, uma vez que tanto a deleção quanto a superexpressão levam a efeitos deletérios (danos ao mtDNA ou o comprometimento da função da

organela). Estudos com StWHY2 mitocondriais mostram que esta proteína possui a capacidade de formar hexâmeros dos tetrâmeros, ou seja, StWHY2 liga-se ao DNA sob a forma de 24-mêros, em que um resíduo de lisina, K67, é fundamental para esta interação ocorrer (Capadocia *et al.*, 2012). Esta pode ser uma das maneiras pela qual a presença desta proteína protege o mtDNA de danos, como DSB, e ao mesmo tempo sua superexpressão poderia aumentar a formação desses 24-meros, levando a uma condensação no mtDNA e ao comprometimento da função mitocondrial. Essa formação também é observada para proteínas AtWHY1, em que a mutação K91A resulta em fenótipo variegado e estiolado, o que indica a importância dessa formação para os cloroplastos (Capadocia *et al.* 2012).

Em plastídios de cevada, proteínas HvWHY1 compactam o DNA em subpopulações de nucleóides. A morfologia de nucleóides em cloroplastos de plantas de cevada *Knock-down* para o gene HvWHY1 é mais heterogênea considerando forma e tamanho. Essa alta compactação do DNA interfere na expressão do DNA plastidial: plantas que não expressavam o gene tiveram aumentados os níveis de expressão de um possível HvPoll-Like (*Barley Organelle DNA Polymerase*), enquanto *E.coli* superexpressando a proteína HvWHY1 tiveram seu crescimento reduzido, e apresentaram nucleóides mais condensados. A expressão foi feita excetuando-se apenas a sequência relacionada à localização plastidial da proteína, PTP (*plastid transit peptide*) (Krupinska *et al.*, 2014). As mudanças na expressão de HvWHY1 relacionadas ao desenvolvimento do cloroplasto acompanham a mudança na expressão do gene SVR4, proteína importante para a organização dos nucleóides em plastídios (Powikrowska *et al.*, 2014a). No entanto, a mudança no acúmulo de HvWHY1 e HvSRV4 em cloroplastos não acompanha a mudança dos níveis de mRNA das proteínas. Enquanto os níveis proteicos de HvSVR4 vão aumentando durante a maturação do cloroplasto, os níveis de HvWHY1 vão decrescendo durante o processo. Isso indica que HvSVR4 e HvWHY1 possuem papéis diferentes na compactação do ptDNA durante a maturação (Krupinska *et al.*, 2014).

Nos cloroplastos, a fração transcricionalmente ativa do ptDNA está organizada em nucleóides e esta fração está relacionada à membrana tilacoide. AtWHY1 foi detectado no proteoma da fração transcricionalmente

ativa obtida a partir de cloroplastos de *Arabidopsis*. Em cevada foi verificado que uma parte das proteínas HvWHY1 podem estar associadas à essa membrana (Grabowski *et al.*, 2008).

As proteínas Whirly de cevada interagem com a região intrônica em RNA plastidial, e podem ter função em processos de splicing (Melonek *et al.*, 2010). Proteínas Whirly em milho também se associam a RNA e DNA plastidiais *in vitro* e *in vivo*. ZmWHY1 possui papel na biogênese plastidial a partir da maturação de RNA de *atpF* e de rRNA 23S, mas mutantes *Zmwhy1* não estão diretamente envolvidos na replicação ou na transcrição global de plastídios (Prikryl *et al.*, 2008).

Além disso, AtWHY1 também foi encontrada em frações de proteínas que se ligam ao telômero, sendo que a ausência da proteína resultou em mutantes que possuíam telômeros maiores, o que se leva a discussão de que Whirly1 pode ser um inibidor da telomerase (Yoo *et al.*, 2007).

Proteínas Whirly nucleares e alterações transcricionais

As proteínas Whirly de cevada foram encontradas no núcleo e nos cloroplastos de uma mesma célula (Grabowski *et al.*, 2008). Ensaio de complementação bimolecular de fluorescência (BiF-C) demonstraram a capacidade de proteínas HvWHY1 e AtWHY1 de formar homooligômeros no núcleo. No entanto, não há a mesma interação nos cloroplastos (Grabowski *et al.*, 2008). Posteriormente, a expressão heteróloga da proteína AtWHY1 inserida no DNA plastidial e fusionadas a uma tag HA, em *Arabidopsis*, resultou no acúmulo destas proteínas tanto em cloroplastos como no núcleo, demonstrando que há um transporte entre os compartimentos (Isemer *et al.*, 2011).

Uma vez no núcleo, estas proteínas podem modular a transcrição gênica de proteínas relacionadas a estresses, como AtPK1 (Xiong *et al.*, 2009), PR1 e PR2, (Isemer *et al.*, 2011) WRK53 (Miao *et al.*, 2013). AtWHY1 e AtWHY3, juntos, são repressores transcricionais de AtKP1, um gene relacionado a morfogênese do fuso, durante a meiose em *Arabidopsis* (Xiong *et al.*, 2009). A fusão HAWHY1 localiza-se no núcleo em regiões de cromatina densa, e resulta no aumento da transcrição de genes relacionados à defesa contra

patógenos, como PR1 e PR2 em 700 e 70 vezes, respectivamente (Isemer *et al.*, 2011). A forma nuclear de WHY1 também tem papel na supressão de WRK53 durante a senescência foliar, ao ligar-se a sequência GNNNAATT presente no promotor de WRK53. A deleção da proteína resultou em fenótipo de senescência em *A. thaliana* (Miao *et al.*, 2013). Knockdown de HvWHY1 resulta em modificações de histonas de HvS40 e nos níveis de fatores da família WRKY e contribui para tolerância a seca (Janack *et al.*, 2016). Através de RNA-sequencing (RNA-Seq) observou-se que há uma reprogramação genética em plantas *why1why3pollb-1*, alterando a expressão genética de maneira semelhante ao observado na resposta a ROS e a estresses abióticos, inclusive promovendo adaptação à alta luminosidade. Essa reprogramação não é observada em plantas crescidas em condições de baixa luminosidade. Nessas condições a produção de ROS é menor, o que sugere que a expressão de genes em resposta aos rearranjos no ptDNA são dependentes de ROS como moléculas sinalizadoras. (Lepage *et al.*, 2013)

O transporte de proteínas dos plastídios para o núcleo e a regulação da expressão de genes relacionados à defesa a partir da atividade de proteínas Whirly continua sendo um mistério a ser desvendado. Considerando sua dupla localização, AtWHY1 é um excelente candidato para realizar a sinalização retrógrada entre plastídios e núcleo.

Foyer *et al.*, 2014, propõem um mecanismo de sinalização retrógrada durante a resposta mediada por ácido salicílico: uma mudança no estado redox da célula resultaria na dissociação de AtWHY1, proteína inicialmente encontrada sob a forma de oligômeros e intimamente associada à membrana tilacóide. Reduções em pontes dissulfeto, nitrosilação de proteínas, e proteínas cinases provavelmente estão envolvidos no processo de mudança conformacional. Sob a forma de monômeros, AtWHY1 é translocado para o núcleo, onde pode ter papel na reprogramação da expressão genica ocorrida após o estresse. Essa sinalização retrógrada pode ser importante para a regulação da ligação ao DNA por proteínas Whirly em plantas submetidas a estresse biótico, e para o crosstalk entre respostas a estresses bióticos e estresse oxidativo.

No entanto, esse processo parece ocorrer não somente quando a planta está numa situação de estresse. Além disso, apesar de haver a possibilidade

de formação de 24-meros no citoplasma e uma possível associação desses à membrana tilacóide, em sua maioria as proteínas AtWHY1 estariam no estroma do citoplasma (Grabowski *et al.*, 2008), em forma monomérica, de maneira que a monomerização de AtWHY1 não parece ser a forma de regulação da localização da proteína. Existem também alguns indícios de que não há formação de 24-meros nos cloroplastos, pois a presença de quimeras HvWHY1-GFP e HvWHY1-CFP não resultou em sobreposição da fluorescência, além de não haver complementação em ensaios de BiF-C (Grabowski *et al.*, 2008). No entanto, ao fusionar HvWHY1 à proteína GFP, a proteína foi exclusivamente enviada para os plastídeos. Discute-se que isso pode ocorrer em função do tamanho da quimera (68kD). Assim, a associação de HvWHY1 com alguma outra proteína pode ser o ponto de regulação do transporte ao núcleo ou permanência de HvWHY1 nos plastídeos.

A associação de WHY1 com outras proteínas pode também ser a chave para entender a regulação da transcrição gênica por estas proteínas. Essa regulação provavelmente depende da sinalização desencadeada após o estresse sofrido pelas plantas em mecanismos que resultariam na ativação de WHY1 por fosforilação, na dissociação de reguladores negativos, ou ainda na associação com proteínas relacionadas à regulação positiva da transcrição. Assim, conhecer as proteínas que são capazes de interagir com WHY1 pode ser fundamental para a elucidação do processo de transporte.

Os estudos relacionados às proteínas Whirly plastidiais e mitocondriais mostram uma diversidade de funções relacionadas à capacidade de ligação ao DNA ou RNA, mostrando que estas proteínas são importantes para a estabilidade das organelas as quais pertencem. As proteínas Whirly são transportadas dos plastídios para o núcleo, onde podem se ligar ao DNA e aparentemente só possuem papel como reguladores de transcrição no núcleo, onde podem regular a expressão de genes relacionados à defesa contra patógenos, senescência e adaptação a estresses abióticos. Essas características fazem com que as proteínas Whirly sejam boas candidatas para estudos envolvendo a sinalização retrograda.

Algumas possibilidades sobre como ocorre a regulação da ligação de proteínas Whirly ao DNA são inferidas através do conhecimento que se tem

acerca da proteína e do cis-elemento ao qual esta se liga. O PB core, por exemplo, possui sobreposições com cis-elementos de outras famílias, como Wbox, reconhecidos por WRKY, e o elemento TGACG, reconhecido por bZIP, porém, a família Whirly liga-se preferencialmente ao ssDNA, enquanto WRKY e bZIP ligam-se ao dsDNA. Poderia estar aí uma fonte de regulação da atividade de ligação ao DNA de proteínas Whirly in vivo, pois em situações onde a planta não está submetida a estresse proteínas destas famílias poderiam estar ligadas ao dsDNA, impedindo a ligação de proteínas Whirly na região promotora e vice-versa (figura 3)(Desveaux *et al.*, 2004). O contato com o patógeno poderia ativar vias que resultam na dissociação de proteínas de outras famílias do promotor (Figura 3). Proteínas WRKY e Whirly poderiam, ainda, estar associados e promover, em conjunto, a regulação da transcrição; nesse caso um dos fatores estaria atuando como proteína modeladora e promovendo a associação de outras proteínas responsáveis pela transcrição. Mais uma vez a chave da atuação de whirly pode estar no conhecimento de proteínas que interagem com WHY1.

A recuperação da atividade de ligação ao DNA em tubérculos após cromatografia de troca aniônica é um indício da interação de um regulador negativo com proteínas Whirly no núcleo de células antes destas serem submetidas ao estresse (Figura3) (Desveaux *et al.*, 2004). Estudos de interação podem ser úteis para a identificação deste(s) possível (is) repressor (es). Experimentos como duplo híbrido e co-imunoprecipitação podem gerar uma grande gama de conhecimento sobre os processos e mecanismos aos quais as proteínas Whirly podem estar inseridas.

A formação tetramérica seria outro ponto de regulação destas proteínas (Figura 3), ao se considerar a formação de homo e heterotetrâmeros. Seria interessante determinar a especificidade de ligação de várias espécies de heterotetrâmeros e examinar como o seu papel biológico pode variar (Desveaux *et al.*, 2002). Proteínas nucleares, mitocondriais e plastidiais possuem a mesma massa (Grabowski *et al.*, 2008), mas a massa para os monômeros maduros (após clivagem do PTP) são similares nas diferentes espécies como observado para os pares StWHY1 e StWHY2, que possuem 24,6 e 24,0 kDA, respectivamente, HvWHY1 (21,9) e HvWHY2 (20,9), e AtWHY1 (24,0), AtWHY2 (23,0), AtWHY3 (21,3), além de apresentarem

predição semelhante para seus respectivos pontos isoelétricos (Krause e Krupinska, 2009). A expressão de uma proteína fusionada ao “tag” HA resulta em acúmulo desta tanto em plastídios como no núcleo (Isemer *et al.*, 2011), mas AtWHY3 foi descrita como repressor do gene AtPK1, junto com AtWHY1, o que indica a formação de heterodímeros no núcleo (Xiong *et al.*, 2009) (Fig3).

Além disso, a regulação pode ser diferente para cada tipo de organismo estudado. Em soja, por exemplo, existem 18 transcritos, que por sua vez são codificados por 7 genes (dados não publicados). O gene Glyma08g40850, chamado de WHYRLY1A, ou GmWHY1A é reprimido durante a infecção pelo patógeno *Phakopsora pachyrhizi*, e possui quatro possíveis transcritos. Neste caso os transcritos alternativos poderiam ter importância para a regulação da atividade do transfator (dados não publicados).

Uma forma truncada dos fatores de transcrição, formada a partir do splicing alternativo, pode atuar como um peptídeo de interferência, siPEP, onde não há a formação da estrutura correta para a ativação da transcrição. Esses peptídeos podem não apresentar o domínio de ligação ou de ativação, por exemplo, e então, a partir da competição entre as proteínas (completas ou truncadas) pelo sítio de ligação ao DNA poderia haver uma diminuição da expressão do gene controlado por estes fatores. Outra possibilidade é a interação entre o fator de transcrição completo e truncado, que poderia resultar em formação de uma estrutura incorreta, incapaz de realizar a transcrição, ou ainda uma interação que resultaria na retenção do fator de transcrição no citoplasma ou outras organelas, e sua ausência no núcleo. As formas alternativas poderiam ainda não apresentar sítios de ligação a cofatores e reguladores, resultando numa regulação negativa ou positiva do gene controlado por estes fatores (SEO *et al.* 2011). O splicing alternativo de fatores de transcrição pode ocorrer em função de estresses ambientais, onde transcritos alternativos podem ter função como regulador do próprio transcrito principal (SEO *et al.*, 2013).

O estudo dos padrões de transcrição para os diversos genes Whirly em soja pode ajudar a esclarecer a forma de regulação desta proteína, mesmo que nesta planta a família seja muito maior do que o observado para outras espécies. Nesse ínterim, seria de suma importância verificar a localização

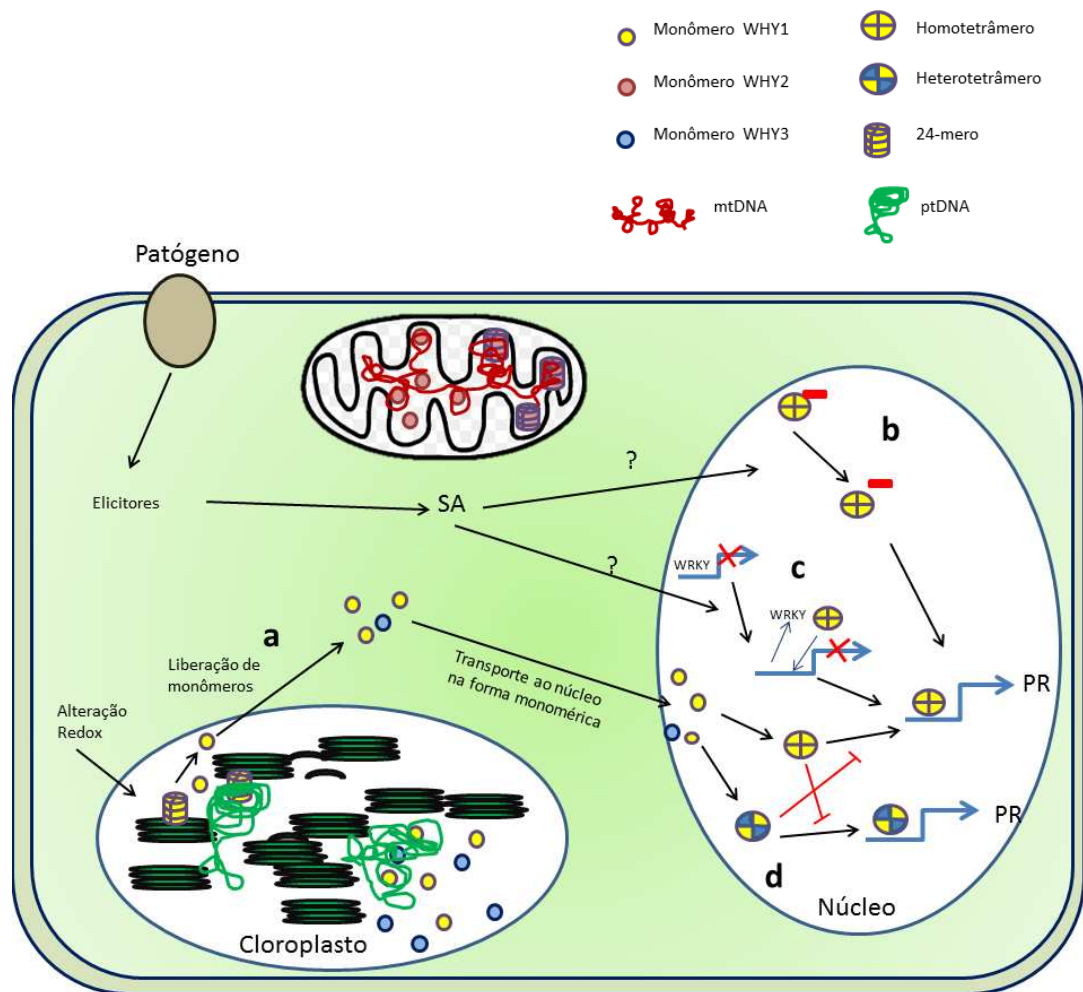


Figura 3: Possibilidades para a regulação de WHY1. Ao se considerar heterotetrameros, a ligação das proteínas WHIRLY ao mesmo cis-elemento reconhecido por fatores de transcrição WRKY, e a possibilidade de haver algum inibidor com carga negativa, várias opções para a regulação da transcrição através de proteínas WHIRLY podem ser discutidas. Além da regulação através da translocação de AtWHY1 do cloroplasto para o Núcleo(a), já discutido por Foyer et al. (2014), outras opções seriam: (b) a presença de uma carga negativa na proteína agiria como regulador negativo. Após a sinalização através de SA a carga seria retirada e AtWHY1 se ligaria ao DNA; (c) a ligação de proteínas WRKY seria impedida pela presença de AtWHY1, que agiria, dessa forma, como inibidor da transcrição de proteínas de defesa. O contrário também é possível, onde fatores WRKY agiriam inibindo a ativação de genes específicos ao interferir com a ligação de AtWHY1 ao promotor; (d) a presença de heterotetrameros pode interferir com a ligação ao DNA e até mesmo com a sequência do DNA que será reconhecida por fatores de transcrição WHIRLY.

subcelular, o padrão de expressão durante a resposta a estresses e o possível papel de Whirly como fator de transcrição para cada um dos transcritos observados para essa importante cultura agrônômica. Além disso, é imprescindível a realização de experimentos que visam elucidar as proteínas que interagem com WHY1 em soja e em outros organismos.

CONCLUSAO

Os dados apresentados indicam que a família Whirly esta envolvida em diversos processos celulares e aparentemente exerce diferentes funções, de acordo com a organela em que se encontra. Apesar de ainda não haver estudos conclusivos em relação a sua função, é notório que essas proteínas são importantes para a resposta os estresses bióticos, seja através da capacidade de ligação ao DNA, ou através da manutenção do cloroplasto e dos níveis de ROS na célula.

Uma vez revelada a função desta família e as vias as quais esta pertence, genes relacionados à família Whirly podem vir a ser alvo para a manipulação genética, visando a obtenção de cultivares resistentes a patógenos, e conseqüentemente diminuição nos custos de produção de diferentes culturas agrônômicas.

REFERÊNCIAS

Alves MS, Dadalto SP; Gonçalves AB, Souza GB, Barros VA, Fietto LG (2013) Plant bZIP Transcription Factors Responsive to Pathogens: A Review. *Int. J. Mol. Sci.* 14(4): 7815- 7828 .

Cabreira C, Cagliari A, Bucker-Neto L, Wiebke-Strohm B, Freitas LB, Marcelino- Guimaraes F, Nepomuceno AL, Margis-Pinheiro MMAN, Bodanese-Zanetini MH (2013) The Lesion Simulating Disease (LSD) gene Family as a variable in soybean response to *Phakopsora pachyrizi* infection and dehydration. *Funct. Integr. Genomics.* 13: 323-338.

Cappadocia L, Maréchal A, Parent JS, Lepage E, Sygusch J, Brisson N (2010) Crystal Structures of DNA-Whirly Complexes and Their Role in Arabidopsis Organelle Genome Repair. *The Plant Cell.* 22: 1849–1867.

Cappadocia L, Parent JS, Zampini E, Lepage E, Sygusch J, Brisson N (2012) A conserved lysine residue of plant Whirly proteins is necessary for higher order proteins assembly and protection against DNA damage. *Nucleic Acids Res.* 40:258-269.

Cappadocia L, Parent JB, Sygusch J, Brisson N (2013) A family portrait: structural comparison of the Whirly proteins from Arabidopsis thaliana and Solanum tuberosum. *Acta Crystal.* 69:12071211.

Derksen H, Rampitsch C, Daayf F (2013) Signaling cross-talk in plant disease resistance. *Plant Science.* 207: 79-87.

Desveaux D, Després C, Joyeux A, Subramaniam R, Brisson N (2000) PBF-2 is a novel single-stranded DNA binding factor implicated in PR10a gene activation in potato. *The Plant Cell, Vol. 12,* 1477–1489.

Desveaux D, Allard J, Brisson N, Sygusch J (2002) A new family of plant transcription factors displays a novel ssDNA- binding surface. *Nat. Struct. Biol.* 9, 512-517.

Desveaux D, Subramaniam R, Despres C, Mess J, Levesque C, Fobert PR, Dangi JL, Brisson N (2004) A “Whirly” transcription factor is required for salicylic acid-dependent disease resistance in Arabidopsis. *Cell Press.* 6:229-240.

Desveaux D, Maréchal A, Brisson N (2005) Whirly transcription factors: defense gene regulation and beyond. *TRENDS in Plant Sci.* 10: 95-102.

Estavillo GM, Chan KX, Phua SY, Pogson BJ (2013) Reconsidering the nature and mode of action of metabolite retrograde signals from the chloroplast. *Front. In Plant Sci.* 3: (300) 1-9.

Foyer CH, Karpinska K, Krupinska K (2014) The functions of WHIRLY1 and REDOX-RESPONSIVE TRANSCRIPTION FACTOR1 in cross tolerance responses in plants: a hypothesis. *Phil.Trans. R. Soc.* 369:20130226.

Grabowski E, Miao Y, Mulisch M, Krupinska K (2008) Single-Stranded DNA-Binding Protein Whirly1 in Barley Leaves Is Located in Plastids and the Nucleus of the Same Cell. *Plant Physiol.* 147: 1800-1804.

Green BR (2011) Chloroplast genomes of photosynthetic eukaryotes. *Plant J.* 66:34-44.

Holub EB, Beynon LJ, Crute IR (1994) Phenotypic and genotypic characterization of interactions between isolates of *Peronospora parasitica* and accessions of *Arabidopsis thaliana*. *Mol. Plant Microbe Interact.* 7:223-239.

Huner NP, Bode R, Dahal K, Hollis L, Rosso D, Krol M, Ivanov AG (2012) Chloroplast redox imbalance governs phenotypic plasticity: the “grand design of photosynthesis” revisited. *Front Plant Sci* 3:255.

Isemer R, Mulisch M, Schafer A, Kirchner S, Koop HU, Krupinska K (2011) Recombinant Whirly1 translocates from transplastomic chloroplasts to the nucleus. *FEBS Lett.* 586:85-88.

Isemer R, Krause K, Grabe N, Kitahata N, Asami T, Krupinska K (2012) Plastid located WHIRLY1 enhances the responsiveness of *Arabidopsis* seedlings toward abscisic acid. *Front. In Plant Sci.* 3:1-11.

Janack B, Sosoi P, Krupinska K, Humbeck K (2016) Knockdown of WHIRLY1 affects drought stress-induced leaf senescence and histone modifications of the senescence-associated gene HVS40. *Plants (Basel)* 5 (3): 37.

Janicka S, Kuhn K, Ret M, Bonnard G, Imbault P, Augustyniak H, Gualberto JM (2012) A RAD52like single-stranded DNA binding protein affects mitochondrial DNA repair by recombination. *The Plant J.* 72: 423-435.

Krause K, Kilbiński I, Mulisch M, Rodiger A, Schafer A, Krupinska K (2005) DNA-binding proteins of the Whirly family in *Arabidopsis thaliana* are targeted to the organelles. *FEBS Lett.* 579:3707-3712.

Krause K, Krupinska K (2009) Nuclear regulators with a second home in organelles. *Trends Plant Sci.* 14:194-199.

Krupinska K, Oetke S, Desel C, Mulisch M, Schafer A, Hollmann J, Kumlehn J, Hensel G (2014) WHIRLY1 is a major organizer of chloroplast nucleoids. *Frontiers in Plant Science.* 5: (432) 1-11.

Lepage E, Zampini E, Brisson N (2013) Plastid genome instability leads to reactive oxygen species production and plastid-to-nucleus retrograde signaling in *Arabidopsis*. *Plant Physiol.* 163:867-881.

Maleck K, Levine A, Eulgem T, Morgan A, Schmid J, Lawton KA, Dangl JL, Dietrich RA (2000) The transcriptome of *Arabidopsis thaliana* during systemic acquired resistance. *Nat Genet.* 26(4):403-10.

Maréchal A, Parent JS, Sabar M, Lafortune FV, C A Rached, Brisson N (2008) Overexpression of mtDNA-associated AtWHY2 compromises mitochondrial function. *BMC Plant Biology.* 8:42.

Maréchal A, Parent JS, Lafortune FV, Joyeux A, Lang BF, Brisson N (2009) Whirly proteins maintain plastid genome stability in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* 106:14693-14698.

Matton DP, Brisson N. (1989) Cloning, expression and sequence conservation of pathogenesis-related gene transcripts of potato. *Mol. Plant Microbe Interact.* 2, 325-331.

Matton DP, Prescott G, Bertrand C, Camirand A, Brisson N (1993) Identification of cis-acting elements involved in the regulation of the 17kDa pathogenesis-related gene STH-2 in potato. *Plant Mol. Biol.* 22: 279-291.

McCallum CM, Comai L, Greene EA, Henikoff S (2000) Targeted screening for induced mutations. *Nat. Biotechnol.* 18: 455-457.

Melonek J, Mulisch M, Schmitz-Linneweber C, Grabowski E, Hensel G, Krupinska K (2010) associates with intron containing RNAs and rarely co-localizes with nucleoids. *Planta.* 232:471481.

Miao Y, Jiang J, Ren Y, Zhao Z (2013) The single-stranded DNA-Binding protein WHIRLY1 represses WRKY53 expression and delays leaf senescence in a developmental stage-dependent manner in *Arabidopsis*. *Plant Physiol.* 163:746-756.

Moore JW, Loake J, Spoel SH (2011) Transcription Dynamics in plant immunity. *The Plant Cell*. 23: 2809–2820.

Parent JS, Lepage E, Brisson N (2011) Divergent roles for the two Poll-like organelle DNA polymerases of *Arabidopsis*. *Plant Physiol*. 156:254-262.

Powikrowska M, Khrouchtchova A, Martens HJ, Zygadlo-Nielsen A, Melonek J, Shulz A *et al.* (2014) SVR4 (suppressor of variegation 4) and SRV4-like: two proteins with a role in proper organization of the chloroplast genetic machinery. *Physiol. Plant* 150:477-492.

Prikryl J, Watkins KP, Friso G, Van Wijk KJ, Barkan A(2008) A member of the Whirly family is a multifunctional RNA- and DNA-Binding proteins that is essential for chloroplast biogenesis. *Nucleic Acids Res*. 36: 5152-5165.

Seo PJ, Hong SY, Kim SG, Park CM (2011) Competitive inhibition of transcription factors by small interfering peptides. *Trends Plant Sci*. 16: 541-549

Seo PJ, Park MJ, Park CM (2013) Alternative splicing of transcription factors in plant responses to low temperature stress: mechanisms and functions. *Planta*. 237:1415-1424.

Subramanian R, Després C, Brisson N (1997) A functional homolog of mammalian protein kinase C participates in the elicitor-induced defense response in potato. *The Plant Cell*. 9:653664.

Torres MA (2010) ROS in biotic interactions. *Physiol. Plant*. 138:414-429

Triantaphylidés C, Krischke M, Hoeberichts FA, Ksas B, Gresser G, Havaux M, Van Breusegem F, Mueller MJ (2008) Singlet oxygen is the major reactive oxygen species involved in photooxidative damage to plants. *Plant Physiol*. 148:960-968.

Xiao Y, Wang J, Dehesh K(2013) Review of stress specif organelles-to-nucleus metabolic signal molecules in plants. *Plant Science* 212: 102-107.

Xiong JY, Lai CX, Qu Z, Yang XY, Qin XH, Liu GQ (2009) Recruitment of AtWHY1 and AtWHY3 by a distal element upstream of the kinesin gene AtKP1 to mediate transcriptional repression. *Plant Mol. Biol*. 71:437-449.

Yoo HH, Kwon C, Lee MM, Chung IK (2007) Single-stranded DNA binding factor AtWHY1 modulates telomere length homeostasis in *Arabidopsis*. *Plant J*. 49:442-451.

**Chapter II: Characterisation of the WHIRLY family in soybean:
differential regulation of *GmWHY1A* mRNAs expression by
environmental stresses and phytohormones**

(Este artigo foi submetido à revista Annals of Botany)

ABSTRACT

Whirly proteins are important single stranded DNA binding proteins that regulate developmental and stress responses in chloroplasts and nuclei. However, this family remains poorly characterised in crop plants such as soybean. Analysis of the soybean Whirly genes revealed that this family is much larger and more complex than in *Arabidopsis thaliana*, being composed by 18 members, coded by seven *loci*. We designated each of the proteins based on their similarities to *Arabidopsis thaliana* Whirly proteins and *in silico* analysis. Each gene had a specific expression pattern in roots, stems, cotyledons, primary leaves and trifoliolate leaves. This characterisation of the Whirly protein family in soybean identified *GmWHY1A* as the most highly expressed form in terms of transcript abundance. However, the levels of *GmWHY1A* transcripts were decreased following exposure to abiotic (dark chilling) and biotic (*Phakopsora pachyrhizi*) stresses. While *GmWHY1A* transcripts were decreased by treatment with ethylene (ET), they were higher after spraying with either salicylic acid (SA) or jasmonic acid (JA). The differential regulation of *GmWHY1A* transcript abundance in response stress hormones suggests dynamic regulation of *WHY1A* expression in response to environmental stress triggers.

Keywords

Whirly, Soybean, Stress, Chilling, *Phakopsora pachyrhizi*.

INTRODUCTION

Plants are constantly challenged by environmental stresses that can impair growth and productivity. Environmental stresses can be caused by herbivores and pathogens, such as fungus and bacterial infections (biotic stresses) or by extremes of climate, such as drought, high light and cold (abiotic stress). Plant stress responses involve signalling by a plethora of hormones, such as salicylic acid (SA), jasmonic acid (JA) and ethylene (ET) (Dresken *et al.*, 2013; Eremina *et al.*, 2016). Stress signalling pathways are complex and often poorly resolved, requiring the simultaneous participation of different organelles and cellular compartments.

In addition to housing the photosynthetic processes, chloroplasts are directly involved in plant signalling pathways and in phytohormone synthesis. Reactive oxygen species (ROS), which are produced during photosynthesis can act as second messengers in stress signal transduction leading to genetic reprogramming in the nucleus (Huner *et al.*, 2012). The interaction of ROS signalling with phytohormone-dependent pathways that mediate plant stress responses involves many factors such as receptor-like kinases, MAP kinases, protein phosphatases and a wide range of transcription factors.

Whirly proteins are single stranded DNA binding proteins that can act as transcription factors (TF) during biotic and abiotic stress responses. They are also involved in plastid maintenance and the stability of chloroplasts genomes (Marechal *et al.*, 2009). WHY1 regulates the compaction of chloroplasts nucleoids (Krupinska *et al.*, 2014). There are three Whirly proteins in *Arabidopsis thaliana*: WHY1, WHY2 and WHY3. WHY1 and WHY3 are found in both plastids and nuclei while WHY2 is localized to the mitochondria (Grabowsky *et al.*, 2008). *A. thaliana* mutants lacking the WHY1 and WHY3 showed damage to plastidial DNA and greater ROS accumulation after treatment with ciprofloxacin than the wild type (Lepage *et al.*, 2013).

WHY1 was shown to be transported from chloroplasts to the nucleus in transplastomic tobacco plants (Isemer *et al.*, 2012). Moreover, in the nucleus WHY1 triggered the expression of defence proteins such as pathogenesis-related (PR) proteins such as PR1 and PR2 (Isemer *et al.*, 2012). Hence, Whirly proteins play a role in biotic stress responses (Desveaux *et al.*, 2000).

WHY1 was shown to bind Ethylene Response Elements (ERE) in potato, following contact with pathogens or after elicitor treatment (Desveaux *et al.*, 2000). Moreover, the overexpression of WHY1 in *A. thaliana* activates the PR1 expression in a NPR independent manner, while *why1* mutants showed lower resistance to *Peronospora parasitica* infection (Desveaux *et al.*, 2004).

The crystallographic structure of WHY1 from potato revealed that this protein occurs as a tetramer with C4 symmetry (Desveaux *et al.*, 2002). Each protomer is formed by two surfaces containing four beta sheets each and three alpha helix. The c-terminus has a LHL motif that is in contact to other protomers. The alpha helices converge to the core, and create a hydrophilic pore that deviate the beta sheets (Desveaux *et al.*, 2002). The structure as a whole has a quaternary form that is similar to a whirligig in appearance and the protein family was therefore named accordingly.

The KGAAL sequence found in all Whirly forms is important in the interaction between Whirly proteins especially at Lys102 and DNA, (Desveaux *et al.*, 2002). The N-terminal sequence, which is a largely disordered region that has 11 consecutive glutamines, is able to bind to the PR10a promotor activating expression in vitro (Desveaux *et al.*, 2002). DNA binding does not greatly change the conformation of the Whirly proteins. The symmetry of WHY1 orientates the DNA by 90° to the next monomer, stimulating dissociation of the DNA. The structures of the arabidopsis proteins and of Stwhy2 are similar to STwhy1, the key amino acids required for DNA binding being conserved between different plant families and species (Capadocia *et al.*, 2010).

Barley WHY1 has been shown to bind to the promoter regions of senescence associated genes such as HvS40 (Janack *et al.*, 2016). Knockdown of HvWHY1 contributes to drought tolerance (Janack *et al.*, 2016) and to tolerance of nitrogen deficiency (Comadira *et al.*, 2015). It has been suggested that WHY1 supresses the expression of senescence genes in the nucleus by binding to the GNNNAAATT sequence in the promoter of the WRKY53 transcription factor, negatively regulating its expression.

There are no reports in the literature of characterisation of the Whirly protein family in soybean. We therefore undertook this analysis and present evidence showing that the Whirly protein family is more complex in soybean than in other species such as barley or *A. thaliana*. Organ-specific patterns of

gene expression are presented together with data, showing that WHY1A transcripts are increased following SA and JA treatments, but decreased by exposure to Ethephon, chilling stress or *Phakopsora pachyrhizi* infection.

Material and Methods

In silico analysis:

Arabidopsis thaliana WHY1 sequence was used as query to select Whirly family in *Glycine max* (Phytozome 11.0). It was considered as Whirly proteins that had the WHIRLY domain. These proteins had its subcellular localization predicted using TargetP, iPSORT and Predotar programs. Phylogenetic analyses were performed with proteins sequences from soybean and *A. thaliana* using ClustalW (MEGA 6.0), Maximum Likelihood statistical method and bootstrap test of phylogeny.

Plant material and stress treatments:

G. max (L.) Merril were plated on water and transplanted after six days to vermiculite supplemented using Hoaligans solution. Plants were cultivated in walk-in chamber at 25°C/30°C temperature and 12hours day. After 15 days roots, stem, cotyledon, leaves and trifoliolate leaves were harvested and kept at -80 degrees until use. 100mg of each tissue was used for RNA extraction using Spectrum™ Plant Total RNA KIT (sigma). After extraction samples were measured using Nanodrop. 1µg of RNA was used to synthesize cDNA using the Quantitect Reverse Transcription KIT Qiagen.

Glycine max cultivar 'CD206' was cultivated in greenhouse 12hour day until reach V1 stage. The plants were submitted to dark chilling stress at 4°C. Trifoliolate leaves were collected after 1h, 2h and 12h of stress. As control, plants not submitted to stress were collected at the same times. All samples were stored at -80degrees.

For *Phakopsora pachyrhizi* infection analysis all RNA samples from EMBRAPA48 (susceptible) and PI561365 (resistant) infected and non-infected with PP were provided from EMBRAPA and extracted according previous report (Alves *et al.*, 2015).

Hormone treatments:

G. max CD206 plants were grown until v3 stage, and then submitted to hormones treatment. Jasmonic Acid (1,2uLMeJa/100mL), SA (1mM) and Ethephon (10mM) were diluted in EtOH 0,01% (SA and JA) or TBS Tween 0,01% (Ethephon) and sprayed towards all leaves and trifoliolate leaves accordingly to Lee *et al.* (2001) and Alves *et al.* (2015). After 1hour, 2 hours, 6hours and 12 hours of treatment trifoliolate leaves were collected and stored at -80 degrees until RNA extraction. All soybean CD206 samples had their RNA extracted using TRIZOL reagent. RNA quality was verified by electrophoresis (1% agarose). Qubit (ThermoFish) was used to quantify RNA samples. 1µg of RNA previously treated with DNase was used to synthesize cDNA (High Capacity cDNA Reverse Transcriptase kit – Applied Biosystems).

Quantitative PCR:

The qPCR experiments were performed in Applied Biosystems equipment Step one, using 5ng of cDNA synthesized previously and SYBER Green as fluorophore. Specific primers to each gene of soybean Whirly family as well for reference genes that were used in qPCR experiments are shown in supplementary Table1 [**supplementary Information**]. Helicase was used as reference gene in chilling stress experiments, Actin was used for infection and EIF1β was used to analyse the expression level of Whirly family in roots, stem, cotyledon, leaves and trifoliolate leaves (Jian *et al.*, 2008; Le *et al.*, 2012; Alves *et al.*, 2015). The statistical analyses were performed using REST 2009 software.

Results and discussion

Searching with AtWHY1 in Phytozome revealed 18 Whirly protein sequences in soybean genome that are located on seven loci (Table1), with an alignment of the Whirly proteins from soybean and *A. thaliana*, as shown in Figure 1. Sequences for chloroplast and mitochondria localized forms are highlighted by the shadow boxes, as well as the Nuclear Localization Signal (NLS) and the amino acids that are crucial for single strand DNA binding (indicated by a *).

The putative transactivation domain (PTD) in soybean sequences was located in a poly P region that is conserved between Glyma01G043000 and Glyma02G019900 and between Glyma08G297200 and Glyma18G124800. These sequences differ from AtWHY1 and AtWHY3, which have a poly S region, and from StWhy1, which has a poly Q sequence (Desveaux *et al.*, 2005).

The phylogenetic tree of Whirly protein Family shows that Glyma08G297200 and Glyma18G124800 are more similar to AtWHY1 and AtWHY3 (both located in chloroplasts). Glyma03G252100 and Glyma19G2499500 are more similar to AtWHY2 (mitochondrial), while Glyma01G043000, Glyma02G019900 and Glyma02G020400 have less similarity to Arabidopsis Thaliana proteins, and formed a separated group (Figure 2). Considering this similarity, the soybean proteins were named as showed in table 1: those proteins there were more similar to AtWHY1 and AtWHY3 were classed as GmWHY1. Glyma08G297200 and Glyma18g124800 were called GmWHY1A and GmWHY1B, respectively. Glyma03G252100 and Glyma19G2499500 were similar to AtWHY2, so they were called GmWHY2A and GmWHY2B, respectively. Glyma01G043000, Glyma02G019900 and Glyma02G020400 were called GmWHY4A, GmWHY4B and GmWHY4C in order to not be confused to AtWHY3 (table1). Higher complexity may have consequences to Whirly regulation in soybean plants. Whirly proteins from group 4 lack the KGKAAL conserved sequence (Figure 1) that is described as required for optimal DNA binding (Desveaux *et al.*, 2005). A heterotetramer formed with one of these proteins would probably

Table 1: Whirly family members in soybean. Soybean Whirly family is coded by seven loci. Proteins were named due its cellular localization and similarity to *Arabidopsis thaliana* Whirly proteins. Whirly proteins from soybean show high complexity and several alternatives transcripts.

Gene ID	Named	Group	Putative Location	Transcripts
Glyma.01G043000	GmWHY4A	WHY4	Undetermined	2
Glyma.02G019900	GmWHY4B	WHY4	Undetermined	1
Glyma.02G020400	GmWHY4C	WHY4	Undetermined	5
Glyma.03G252100	GmWHY2A	WHY2	Mithochondria	1
Glyma.08G297200	GmWHY1A	WHY1	Chloroplast	4
Glyma.18G124800	GmWHY1B	WHY1	Chloroplast	3
Glyma.19G249500	GmWHY2B	WHY2	Mithochondria	2

abolish the DNA binding activity of Whirly. The presence of alternative transcripts also implicates in more complex regulation, once the lack of some domains in TFs can alter the function of this TF upon a signal. Seo *et al.* (2013) discusses the importance of alternative splicing to low temperature response.

Roots, stem, cotyledon, leaves and trifoliolate leaves had their Whirly family expression analysed by quantitative real time using *ELF1 α* as reference gene (Figure 3). GmWHY1A is the most expressed Whirly protein while GmWHY4A presented lower expression in almost all analysed parts (except leaves). All Whirly members are more expressed in green tissues (cotyledon, leaves and trifoliolate leaves).

Whirly protein sequences were analysed in order to predict their subcellular location and by presence of a NLS (Nuclear Location Signal). Proteins that belong to group 1 (GmWHY1A and GmWHY1B) are predicted to located at chloroplast and also have a NLS in their sequences. Together with similarity to ATWHY1, it makes these two proteins good candidates to act during the retrograde signalling. Members of the group 2 (GmWHY2A and GmWHY2B) have their predicted location in mitochondria. For members of group 4 the prediction revealed two different locations (GmWHY4A and GmWHY4B) or none (GmWHY4C), so the predicted location was considered undetermined for these proteins. Considering the high similarity, the predicted location, the presence of a NLS and the expression patterns, the protein GmWHY1A was chosen for further analysis in order to verify its expression during stress response in soybean.

WHY1A had its expression suppressed after 1h, 2h and 12 hours of dark chilling stress (Figure 4). It remains unclear whether its suppression is a consequence of the lower global expression of cell or whether it is part of cold response. Similar results were obtained to WHY1B (data not shown). Ethephon treatment showed similar results to chilling stress, which is compatible with previous studies about ethylene during cold response (Eremina *et al.*, 2016). It has been demonstrated that Whirly proteins can bind to similar cis-elements of WRKY family trans-factors (Janack *et al.*, 2016). AtKP1 is a kinesin protein (proteins that are involved in transport along

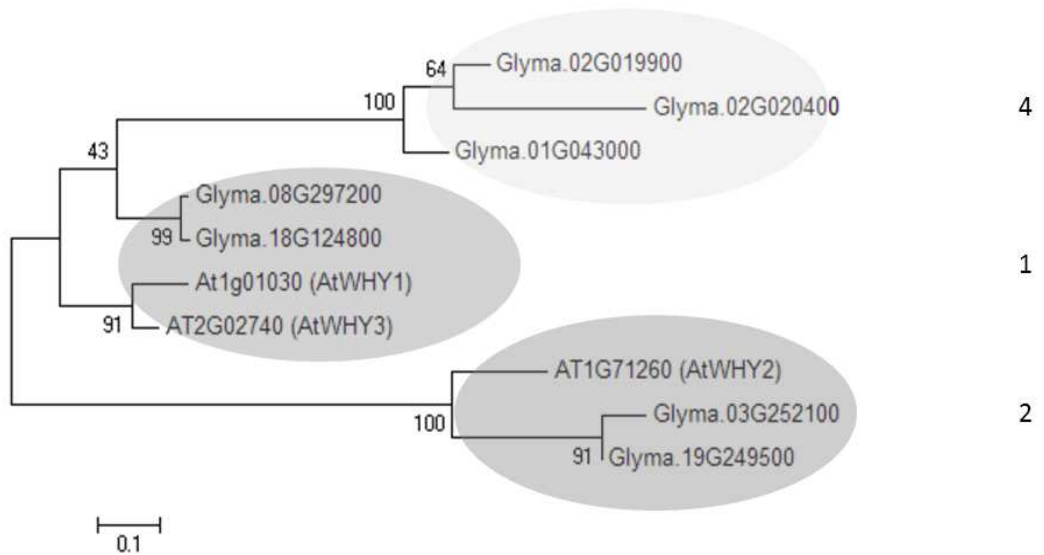


Figure 2: **Phylogenetic tree showing the similarity between members of Whirly family in soybean and *Arabidopsis thaliana*.** The tree was constructed using maximum likelihood algorithm. It is possible to see three distinct groups formed by plastidial proteins (1), mitochondrial proteins (2) and undetermined location proteins (4).

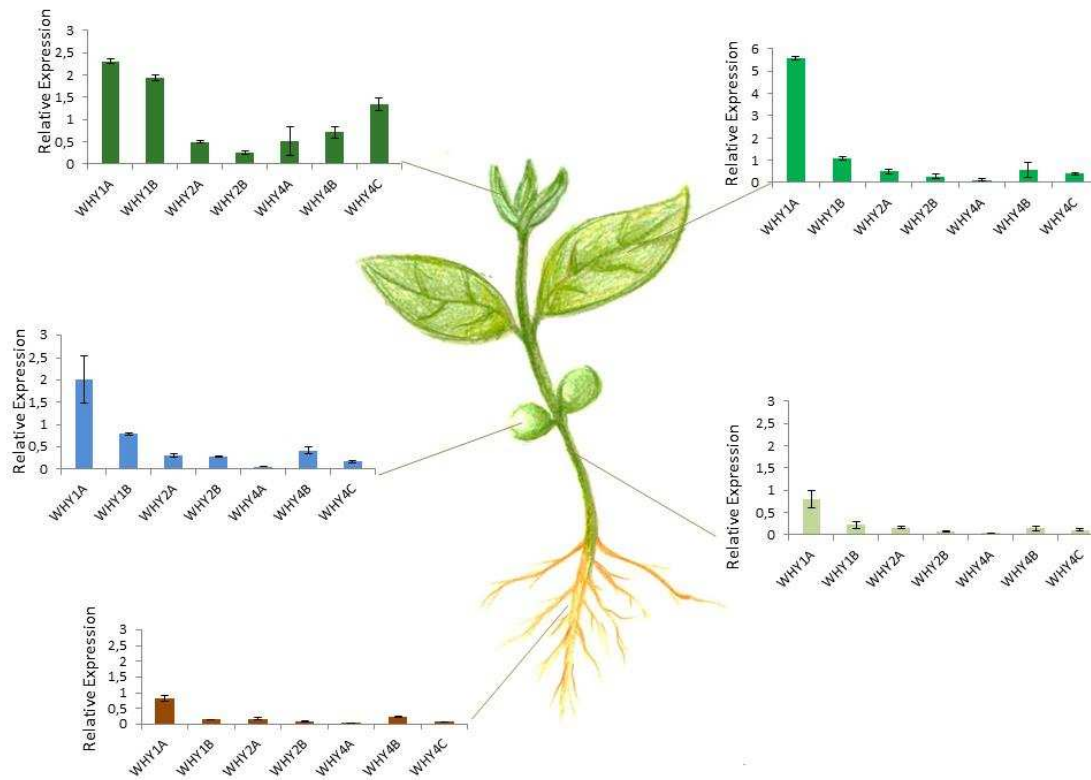


Figure 3: Expression patterns of Whirly genes in roots, stem, cotyledon, leaves, and trifoliolate leaves of soybean plant. ELF1 β was used as reference gene for quantitative real time experiments. WHY1A is the most expressed Whirly in all parts. All Whirly genes were more expressed in green tissues, as cotyledons, leaves and trifoliolate leaves.

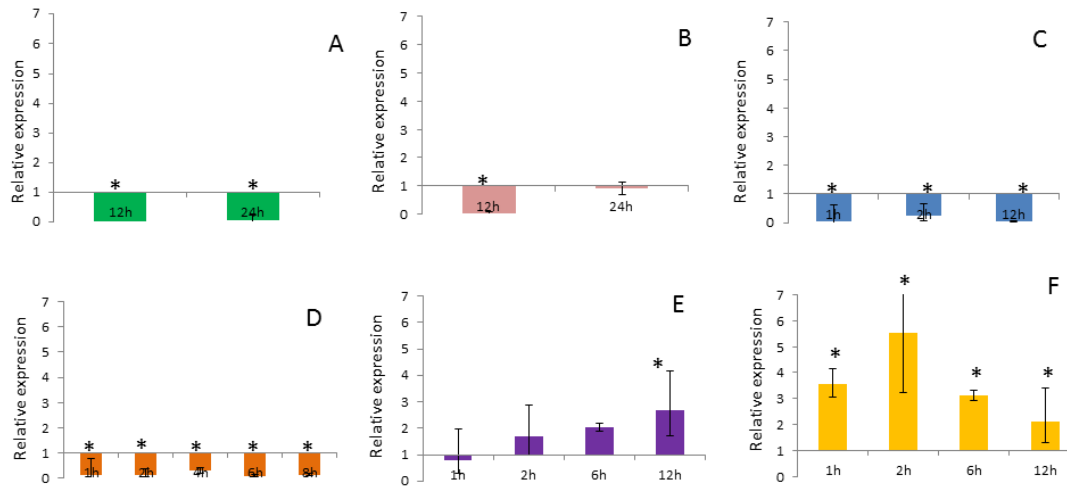


Figure 4: GmWHY1A expression under different conditions: Infection of *Phakopsora pachyrhizi* in susceptible (A) and resistant (B) plants; dark shilling stress (C); after treatment with ethephon (D), jasmonic acid (E) and salicylic acid (F). Both biotic and abiotic stress suppressed WHY1A expression, while SA and JA up regulated WHY1A expression. Ethephon Treatment strongly suppressed WHY1A expression. *indicates statistical significance at 5%.

microtubules). The promoter region of AtKP1 is similar to PB core – the sequence recognized by Stwhy1 – and shares sequence to w-box (recognized by WRKY TFs). After purification steps it was revealed that the repressor of AtKP1 was formed by WHY1 and WHY3 that together bind to its promoter region and repress AtKP1 expression (Xiong *et al.*, 2009). Recently, the WRKY protein CsWRKY46 was related to cold tolerance in cucumber. This protein binds to W-box in ABI5 promoter region and its overexpression resulted in chilling tolerance. Cold related genes COR47, RD29a and KIN1 were upregulated in WORKY46 OE plants (Zhang *et al.*, 2016). Thus, one possibility is that Whirly proteins act as repressors of proteins that are up-regulated during cold adaption by binding to cis-elements recognized by WRKY TFs and are suppressed during chilling response, when WRKY factors can bind and regulate COR genes.

GmWhy1A also is affected by infection. Susceptible (EMPRAPA48) and resistant (PI561356) soybean infected with *P. pachyrhizi* showed lower WHY1A expression when compared to same plants non-infected with fungus. Why1A is modulated in early stages of infection (Figure 4). Surprisingly, SA and JA treatment enhanced WHY1A expression after 12 hours (Figure4), similarly to levels of expression obtained for AtWHY1 and AtWHY3 proteins after SA treatment (Xiong *et al.*, 2009). AtWHY1 is important to PR1 expression during SA signalling in a NPR –independent manner (Desveaux *et al.*, 2004). The expression data obtained to GmWHY1A during fungal infection doesn't follow the SA pattern. This suggests that during contact with *P. pachyrhizi* the fungus is able to down regulate Whirly expression, even for resistant plants.

Another possibility of Whirly function upon stress comes from its plastidial activity. In *Arabidopsis thaliana*, Whirly is described as a chloroplast organizer and its suppression results in higher ROS levels and activation of abiotic stress response due to plastidial DNA damage (Lepage *et al.*, 2013). ROS are known to act as signalling molecules during stress and the lack of Whirly in chloroplasts would increase signal transduction during stress. Whirly proteins WHY1 and WHY3 were founded both in chloroplasts and nucleus (Grabowski *et al.*, 2008). Upon stress WHY1 can migrate from chloroplast to nucleus and both WHY1 and WHY3 can be act as transcriptional factors (Isemer *et al.*,

2011). This is another way that Whirly proteins can contribute for retrograde signalling upon stress.

CONCLUSION

The complex WHIRLY gene family in soybean has been characterised for the first time. We have designated three groups of WHIRLY genes, accordingly to protein sequences, predicted location, and similarity to *A. thaliana* WHIRLY sequences. The expression of GmWHY1A is differentially regulated in response to plant hormones, and to environmental stresses.

ACKNOWLEDGMENTS

We thank to Biotechnology and Biological Sciences Research Council (BBSRC, UK) [BB/M009130/1], Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for financial support. This work was part of Programa de Doutorado Sanduiche no Exterior (PDSE).

REFERENCES

- Alves MS, Soares ZG, Vidigal PMP, Barros EG, Poddanosqui AMP, Aoyagi LN, Abdelnoor RV, Marcelino-Guimarães FC, Fietto LG (2015) Differential expression of four soybean bZIP genes during *Phakopsora pachyrhizi* infection. *Func. & Integr. Genomics*. 15(6): 685–696
- Cappadocia L, Maréchal A, Parent JS, Lepage E, Sygusch J, Brisson N (2010) Crystal Structures of DNA-Whirly Complexes and Their Role in Arabidopsis Organelle Genome Repair. *The Plant Cell*. 22: 1849–1867
- Cappadocia L, Parent JB, Sygusch J, Brisson N (2013) A family portrait: structural comparison of the Whirly proteins from Arabidopsis thaliana and Solanum tuberosum. *Acta Crystal* 69:1207-1211.
- Comadira G, Rasool B, Karpinska B, Márquez García B, Morris J, Verrall SR, Bayer M, Hedley PE, Hancock RD, Foyer CH (2015) WHIRLY1 functions in the control of responses to N-deficiency but not aphid infestation in barley (*Hordeum vulgare*). *Plant Physiol*. 168: 1140-1151.
- Derksen H, Rampitsch C, Daayf F (2013) Signaling cross-talk in plant disease resistance. *Plant Sci* 207: 79-87
- Desveaux, D, Després C, Joyeux A, Subramanian R, Brisson N (2000) PBF-2 is a novel single-stranded DNA binding factor implicated in PR10a gene activation in potato. *The Plant Cell* 12: 1477–1489
- Desveaux D, Allard J, Brisson N, Sygusch J (2002) A new family of plant transcription factors displays a novel ssDNA- binding surface. *Nat Struct Biol* 9, 512-517
- Desveaux D, Subramaniam R, Despres C, Mess J, Levesque C, Fobert PR, Dangi JL, Brisson N (2004) A “Whirly” transcription factor is required for salicylic acid-dependent disease resistance in Arabidopsis. *Cell Press* 6: 229-240
- Desveaux, D, Maréchal A, Brisson N (2005) Whirly transcription factors: defense gene regulation and beyond. *TRENDS in Plant Sci* 10:2
- Eremina M, Rozhom W, Poppenberger B (2016) Hormonal control of cold stress responses in plants. *Cell Mol Life Sci* 73:797-810

Grabowski E, Miao Y, Mulisch M, Krupinska K (2008) Single-Stranded DNA-Binding Protein Whirly1 in Barley Leaves Is Located in Plastids and the Nucleus of the Same Cell. *Plant Physiol* 147: 1800-1804

Huner NP, Bode R, Dahal K, Hollis L, Rosso D, Krol M, Ivanov AG (2012) Chloroplast redox imbalance governs phenotypic plasticity: the “grand design of photosynthesis” revisited. *Front Plant Sci* 3:255

Isemer R, Mulisch M, Schafer A, Kirchner S, Koop HU, Krupinska K (2011) Recombinant Whirly1 translocates from transplastomic chloroplasts to the nucleus. *FEBS Lett* 586:85-88

Isemer R, Krause K, Grabe N, Kitahata N, Asami T, Krupinska K (2012) Plastid located WHIRLY1 enhances the responsiveness of Arabidopsis seedlings toward abscisic acid. *Front. In Plant Sci* 3:1-11

Janack B, Sosoi P, Krupinska K, Humbeck K (2016) Knockdown of WHIRLY1 affects drought stress-induced leaf senescence and histone modifications of the senescence-associated gene HVS40. *Plants (Basel)* 5 (3): 37.

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

Krupinska K, Oetke S, Desel C, Mulisch M, Schafer A, Hollmann J, Kumlehn J, Hensel G (2014) WHIRLY1 is a major organizer of chloroplast nucleoids. *Front in Plant Sci* 5: (432) 1-11

Le DT, Aldrich DL, Valliyodan B, Watanabe Y, Van Ha C, Nishiyama R, Guttikonda SK, Quach TN, Gutierrez-Gonzalez JJ, Tran LSP, Nguyen HT (2012) Evaluation of candidate reference genes for normalization of quantitative RT-PCR in soybean tissues under various abiotic stress conditions. *Plos one* 7(9): e46487

Lee SC, Kim YJ, Hwang K (2001) A pathogen-induced chitin-binding protein gene from pepper: Its isolation and differential expression in pepper tissues treated with pathogens, ethephon, methyl jasmonate or wounding. *42(12): 1321-1330.*

Lepage E, Zampini E, Brisson N (2013) Plastid genome instability leads to reactive oxygen species production and plastid-to-nucleus retrograde signaling in Arabidopsis. *Plant Physiol* 163:867-881

Maréchal A, Parent JS, Lafortune FV, Joyeux A, Lang BF, Brisson N (2009) Whirly proteins maintain plastid genome stability in Arabidopsis. *Proc Natl Acad Sci USA* 106:14693-14698

Seo PJ, Park MJ, Park CM (2013) Alternative splicing of transcription factors in plant responses to low temperature stress: mechanisms and functions. *Planta* 237:1415-1424

Shi Y, Ding Y, Yang S (2014) Cold signal transduction and its interplay with phytohormones during cold acclimation. Oxford University Press 1-28.

Xiong JY, Lai CX, Qu Z, Yang XY, Qin XH, Liu GQ (2009) Recruitment of AtWHY1 and AtWHY3 by a distal element upstream of the kinesin gene AtKP1 to mediate transcriptional repression. *Plant Mol Biol* 71:437-449

Zhang Y, Yu H, Yang X, Li Q, Ling J, Wang H, Gu X, Huang S, Jiang W (2016) CsWRKY46, a WRKY transcriptional factor from cucumber, confers cold resistance in transgenic-plant by regulating a set of cold-stress responsive genes in an ABA-dependent manner. *Plant Physio And Biochem* 108: 478 – 487.

CONCLUSÕES GERAIS

- Existe pouco ou nenhum material publicado sobre proteínas Whirly em soja;

- As proteínas Whirly em soja são compostas de 18 membros, codificados por 7 loci. Cada membro foi analisado quanto a predição de localização celular e similaridade com proteínas de *A. thaliana* e nomeado com base nessas informações.

- A família foi separada em três grupos 1, 2 e 4. O grupo 1 é mais similar às proteínas AtWHY1 e AtWHY2, possui localização predita para o cloroplasto, e sinal de localização nuclear. A família 1 é composta por GMWHY1A e WHY1B. O grupo 2 é similar a AtWHY2, predita como mitocondrial, e não apresentam PTD (putative transactivation domain). O grupo 4 é composto por três membros e é um grupo mais heterogêneo, composto por proteínas que não possuem peptídeo sinal. Um dos membros é uma forma truncada, que não possui domínio de ligação ao ssDNA.

- A expressão da família Whirly foi analisada em raízes, caule, cotilédones, folhas e trifólios. WHY1 é mais expresso em todas as partes estudadas. GMWHY1A foi escolhido para experimentos de expressão em situações de estresses.

- WHY1 é positivamente regulado após tratamento com jasmonato e com ácido salicílico. Após tratamento com Etefon, WHY1 foi fortemente reprimido, já a partir da primeira hora.

- WHY1 é reprimido após resfriamento a 4°C durante a noite, e após contato com o patógeno *Phakopsora pachyrhizi* em trifólios.

Anexo I

Tabela suplementar do capítulo 2.

Gene ID	Name	Fw	Rv
Glyma.01g043000	GmWHY4A	AACATTCCTTTGAACCGCAA	GCTTGGGTTGGACAAGGTTA
Glyma.02g019900	GmWHY4B	CATGCTCACCGTGATTCCTA	CATGCTCACCGTGATTCCTA
Glyma.02g020400	GmWHY4C	CATGGAGGGAATTCAGGAAA	TCCAAAGGCATTCCAACCTA
Glyma.03g252100	GmWHY2A	GTTCCCTCTCACTTCCCCTT	TGAATAGGTAAAATCGCGCC
Glyma.08g297200	GmWHY1A	CAGCTTCACTCTCCTCCACC	AAGGGTGGTTTGGGAATAGC
Glyma.18g124800	GmWHY1B	CTACACCACGACCAACGAGA	AGTCAATGTAAGCGCAGCCT
Glyma.19g249500	GmWHY2B	TGCTCCCTCTCACTTCCACT	ATGCCAGCTGAATGGCTAAG
Glyma.02g44460	ELF1 β	GTTGAAAAGCCAGGGGACA	TCTTACCCCTTGAGCGTGG
Glyma.19g00900	Actin	GAGACATCCGAGACCAGCTC	AATGCCTGATGCTTCCATTC
Glyma.12g170000	Helicase	TAACCCTAGCCCCTTCGCCT	GCCTTGTCGTCTTCCTCCTCG

Supplementary Table 1: Primers used for qPCR reactions. ELF1 β , actin and Helicase were used as reference genes in different experiments.