UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Caracterização de genes WRKY de soja [*Glycine max* (L.) Merrill] responsivos ao déficit hídrico e estudo dos seus promotores.

Leticia Pereira Dias

Porto Alegre, Agosto de 2018

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Caracterização de genes WRKY de soja [*Glycine max* (L.) Merrill] responsivos ao déficit hídrico e estudo dos seus promotores.

Leticia Pereira Dias

Tese submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Federal do Rio Grande do Sul como requisito parcial para obtenção do título de Doutor em Genética e Biologia Molecular.

Orientação: Prof. Dra. Maria Helena Bodanese Zanettini

Co-orientação: Dra. Luisa Abruzzi de Oliveira

Porto Alegre, Agosto de 2018

AGRADECIMENTOS

A Deus pelo amor, misericórdia e graça infinitos.

Aos meus pais, pelo amor incondicional e a torcida incansável.

A Universidade Federal do Rio Grande do Sul pelos 15 anos de educação pública de qualidade que me foram proporcionados. Minha caminhada como aluna desta instituição acaba aqui, mas por onde eu for levarei o nome da UFRGS com muito orgulho. Muito obrigada mesmo!

A todos do PPGBM, todos sem exceção: o pessoal da limpeza (pelo “Bom dia!” tímido, mas sincero que nos motiva a iniciar mais uma jornada diária de trabalho), a equipe da segurança (que muitas vezes me acompanharam até o carro nas noites de sábado quando eu estava sozinha no prédio) , o pessoal da administração (Valeu por tudo Elmo!!), aos professores e colegas.

À CAPES, pelo suporte financeiro e as oportunidades concedidas.

À Iowa State University, especialmente a todos do laboratório do Dr. Steve Whitham.

Obrigada.

INSTITUIÇÃO E FONTES FINANCIADORAS

Este trabalho foi desenvolvido nos seguintes laboratórios:

(i)Laboratório de Genética e Biologia Molecular Vegetal - Departamento de Genética - Universidade Federal do Rio Grande do Sul

(ii) Laboratório de Cultura de Tecidos e Transformação Genética de

Plantas - Departamento de Genética - Universidade Federal do Rio Grande do Sul

(iii) Whitham Lab –Department of Plant Pathology and Microbiology – Iowa State University – United States of America.

E recebeu apoio financeiro do Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), assim como do Consorcio Nacional do Genoma da Soja (GENOSOJA-CNPq) e BIOTECSUR (EU/MCT).

RESUMO

A seca é hoje, e talvez seja ainda mais no futuro, um grande desafio das culturas de grãos devido à redução drástica que causa no rendimento da produção. Sendo assim, é de grande interesse a obtenção de genótipos de soja tolerantes ao déficit hídrico. Por causa de sua base poligênica, a tolerância à seca é uma característica difícil de ser alcançada através do melhoramento clássico. Nesse contexto, a engenharia genética apresenta-se como um facilitador para a construção desse atributo. A fim de projetar estratégias bem sucedidas de manipulação genética, um conhecimento mais profundo sobre os genes e promotores gênicos da soja responsivos ao déficit hídrico é necessário. Os promotores induzíveis podem ser ferramentas poderosas para garantir o controle mais fino da expressão transgênica. A capacidade de modular a expressão de muitos genes faz com que os fatores de transcrição sejam promissores alvos biotecnológicos para o desenvolvimento de cultivares tolerantes ao estresse. As proteínas WRKY formam uma grande família de fatores de transcrição que estão envolvidos em importantes processos fisiológicos e bioquímicos em plantas, incluindo a resposta ao déficit hídrico. Neste estudo, o padrão de expressão determinado por qPCR mostrou que, os genes *GmWRKY6*, *GmWRKY46*, *GmWRKY56*, *GmWRKY106* e *GmWRKY149* são diferencialmente expressos entre um genótipo tolerante à seca e um suscetível em condições de estresse hídrico. As análises *in silico* do promotor e de coexpressão indicam que estes genes atuam em situações de estresse. Esses dados foram publicados no *Plant Physiology and Biochemistry Journal* no ano de 2016. A caracterização do promotor do gene *GmWRKY46* foi iniciada. Os resultados obtidos até o momento mostram que a atividade da sequência de 500pb a montante do sítio de início da transcrição é induzida pelo estresse iônico e abriga cis-elementos relacionados à luz e hormônios vegetais. Visando a continuidade desta pesquisa, vetores para a análise dos genes e promotores WRKY e também linhagens transgênicas de *Arabidopsis Thaliana* carregando as construções “promotor WRKY: gene repórter” foram obtidos. Certamente, essas ferramentas serão muito úteis para esse e futuros estudos.

ABSTRACT

Drought is today, and perhaps even more in the future, the main challenge for grain crops, resulting in a drastic yield reduction. Thus, it is of great interest to obtain soybean genotypes tolerant to water deficit. The drought tolerance trait is difficult to obtain through classical breeding due to its polygenic basis. In this context, genetic engineering is presented as a way to achieve this attribute. In order to project successful strategies of genetic manipulation, a deeper knowledge about soybean genes and their promoters, which are responsive to water deficit is needed. The inducible promoters can be powerful tools in genetic engineering to ensure finer control of transgene expression. The ability to modulate the expression of many genes placed the transcription factors as promising biotechnological targets to develop stress tolerant cultivars. The WRKY proteins form a large family of transcription factors that are involved in important physiological and biochemical processes in plants, including the response to water deficit. In this study, the expression pattern determined by qPCR showed that, *GmWRKY6*, *GmWRKY46*, *GmWRKY56*, *GmWRKY106* and *GmWRKY149* genes are differentially expressed between a drought tolerant and a susceptible soybean genotype in water stress conditions. The *in silico* promoter and coexpression analysis indicate that these genes act in a stress physiological background. These data were published in *Plant Physiology and Biochemistry Journal* in 2016. The *GmWRKY46* gene promoter characterization has also been initiated. In the early results, cis-elements related to light and plant hormones were identified in the 500 bp sequence upstream the transcription start site and this promoter fragment activity was induced by ionic stress. Aiming the continuity of this research, vectors to WRKY genes and promoters analysis were constructed and also *Arabidopsis thaliana* transgenic lines carrying WRKY promoter:reporter gene constructs were obtained. Surely, these tools will be very useful to this and future studies.

LISTA DE ABREVIATURAS E SIGLAS

ABA: A*bscisic acid*, Ácido abscísico

ABRE: ABA *responsive element*, Elemento responsivo ao ABA

AmpR : Resistência à ampicilina

bp: *Base pair*, Pares de bases

CaMV: *Cauliflower mosaic virus*, Vírus do mosaico da couve-flor

cDNA: *Complementary DNA*, DNA complementar

CO_2: Gás carbônico

Col-0 : Columbia (*Arabidopsis thaliana ecotype*), Columbia (Ecótipo de *Arabidopsis thaliana*)

CONAB: Companhia Nacional de Abastecimento

DNA: Ácido desoxirribonucleico

dNTP: Desoxirribonucleotideos trifosfatados

EMBRAPA: Empresa Brasileira de Pesquisa Agropecuária

FAO: *Food and Agriculture Organization of the United Nations,* Organização das Nações Unidas para Agricultura e Alimentação

Gm: *Glycine max*

Glyma: *Glycine max*

GUS: β-glicuronidase

gusA: Gene repórter que codifica a β-glucuronidase

kbp: Quilopares de bases

LEA: *Late embryogenesis abundant protein family,* Família de proteínas

abundantes na embriogênese tardia

MAPA: Ministério da Agricultura, Pecuária e Abastecimento

min: Minuto(s)

mM: Milimolar(es)

MU: Methylumbelliferyl

MYB: *Myeloblastosis Oncogene protein family*, Família de proteínas oncogênicas isoladas do vírus da mieloblastose

MYBR: *MYB recognition sequence, S*equência de reconhecimento das

proteínas MYB

MYC: *Myelocytomatosis Oncogene proteins family*, Família de proteínas oncogênicas isoladas do vírus da mielocitomatose

MYCR: *MYC recognition sequence,* Sequência de reconhecimento das

proteínas MYC

PCR: *Polymerase chain reaction*, Reação em cadeia da polimerase

PEG: Polietilenoglicol

PlantPAN: *Plant Promoter Analysis Navigator*

PlantCARE: Plant Cis Acting Regulatory Elements Database

qPCR: *Quantitative PCR.* PCR quantitativo

RNA: Ácido ribonucleico

RT-qPCR: *Reverse Transcription quantitative PCR*

s: segundo(s)

SA: *Salicylic acid*. Ácido salicílico

snp: S*ingle nucleotide polymorphisms*, Polimorfismo de nucleotídeo único

SAGE: *Serial analysis of gene expression,* Análise Serial da Expressão Gênica

T-DNA: Transfer DNA, DNA de transferência

TFs: *Transcription factors*, Fatores de transcrição

T0, T25 ,T50 ,T75 ,T100 , T125 ,T150 : Tempo de exposição ao déficit hídrico, em minutos

T6 : Tempo decorrido após tratamento com ABA e SA , em horas

U: u*nit(s)*, unidade(s)

VIGS : *Virus-induced gene silencing*, Silenciamento induzido por vírus

WT: *Wild-type,* Selvagem

µl: microlitro**(s)**

SUMÁRIO

INTRODUÇÃO

A SOJA E O DÉFICIT HÍDRICO

Na agricultura, estresse é entendido como o fator que limita a produtividade da cultura abaixo do seu potencial genético. Com base nesse conceito, é possível afirmar indiscutivelmente que a seca é um dos principais fatores estressores das plantas em todo o mundo. O caráter multifuncional da água no organismo vegetal faz com que sua baixa disponibilidade afete negativamente uma série de processos essenciais como como regulação térmica , transporte de nutrientes e células, preservação das vias fisiológicas e bioquímicas e até mesmo a manutenção da estrutura da planta, estabelecendo assim um estado de estresse (Taiz e Zeiger, 2013).

O relatório FAO (2018) informa que em comparação com outros desastres como inundações, queimadas, pragas e temperaturas extremas, a seca foi responsável pelo maior impacto negativo na produção agrícola no período de 2005 a 2015 nos países em desenvolvimento. E o cenário previsto para o futuro não é agradável. A frequência e a intensidade das secas tende a crescer sob efeito do aquecimento global. Somado a isso, o aumento da pressão demográfica demandará uma produção de alimentos 70% maior até 2050 a fim de que a segurança alimentar seja garantida (FAO, 2011). Para mitigar esses problemas, mudanças na forma como a atividade agrícola é exercida serão inevitáveis e deverão seguir por dois caminhos: 1. Redefinição do mapa agrícola mundial evitando o cultivo de culturas intensivas em água em áreas onde a carência hídrica já é uma realidade. 2. Aprimoramento dos sistemas de cultivos a fim de torná-los mais econômicos no uso de água sem causar prejuízos à produtividade.

A obtenção de cultivares tolerantes a seca é uma estratégia crucial para o estabelecimento de cultivos eficientes. E para espécies como a soja que agregam as características de grande demanda hídrica e ampla relevância econômica, essa estratégia deve ser seguida ainda com mais afinco. O requerimento hídrico da soja para o alcance de uma produtividade considerada elevada é aproximadamente 800 mm dia^{-1}, relativamente alto quando comparada a outras importantes culturas (Isoda,

2005; EMBRAPA, 2008). Somado a isso, segundo Berlato (1987), o consumo relativo de água durante o período crítico da cultura (do início da floração até o início do enchimento de grãos) é o principal fator determinante do rendimento de grãos de soja, explicando assim o enorme potencial limitante que a escassez hídrica possui sobre a produtividade dessa cultura.

Entretanto, para que desenvolvimento que cultivares efetivamente tolerantes seja possível, o mecanismo que atribui à soja a capacidade de suportar o estresse sem sofrer sérios danos no desenvolvimento e produção deve ser entendido em todos os níveis, inclusive o molecular. A insuficiência hídrica, condição primária da seca, quase sempre vem acompanhada por outros fatores como altas temperaturas e/ou ventos secos, o que explica a alta complexidade da reação da planta a esse tipo de estresse. A resposta molecular do vegetal a seca é composta por uma rede intrincada que demanda uma regulação fina da expressão gênica. Os fatores de transcrição AREB/ABF, NAM, ATAF, CUC, MYB/MYC e WRKY são alguns dos principais reguladores dessa resposta, atuando como orquestradores na composição do perfil tolerante (Singh e Laxmi, 2015).

FATORES DE TRANSCRIÇÃO WRKY

O controle da transcrição é um dos principais pontos de regulação gênica. Esse mecanismo é mediado pelos fatores de transcrição. Os fatores de transcrição são proteínas capazes de reconhecer e interagir com sequências específicas das regiões promotoras dos genes e, desta forma, definir quais deles serão expressos e em que níveis (Meshi e col., 1995).

As proteínas WRKY formam uma das maiores famílias de fatores de transcrição dos vegetais. Receberam este nome devido a sequência conservada de aminoácidos WRKYGQK (ou WRKYGKK) que identifica todos os componentes da família. Esta sequência faz parte do sítio de ligação da proteína ao DNA que integra 60 aminoácidos seguidos por uma estrutura do tipo dedo-de-zinco (Rushton e col. 2010). Baseando-se no número de motivos WRKY e nas características da estrutura tipo dedo-de-zinco, Eulgem e col., (2000) e Yin e col., (2013) propuseram que a família fosse organizada em três grupos. Proteínas com dois motivos WRKY pertencem ao grupo I. No grupo II estão as proteínas com apenas um motivo. Já o grupo III agrega as proteínas que apresentam uma estrutura dedo-de-zinco C_2-HC, diferente do padrão C_2-H_2 seguido pelos demais grupos.

Estudados a mais de duas décadas, os WRKY podem ser considerados uma das classes de fatores de transcrição mais bem caracterizadas nas plantas. Os primeiros trabalhos publicados ainda na década de 1990 os relacionavam somente a processos metabólicos (Ishiguro e col., 1994), fisiológicos (Rushton e col., 1995) e na defesa contra patógenos (Raventós e col., 1995). Até que no principio dos anos 2000 foram divulgados dados apontando o envolvimento destes fatores de transcrição na resposta das plantas às perturbações ambientais, entre elas a seca (Chen e col., 2011; Marè e col., 2004; Zhou e col. 2008; Ramamoorthy e col., 2008). Porém, devido ao menor tempo de pesquisa, o conhecimento sobre o desempenho dos WRKYs nas rotas de reação da planta aos estresses abióticos ainda é superficial se comparado ao que já se sabe sobre a relação destas proteínas com os estresses bióticos (Rushton e col., 2010).

Ao longo do tempo, novos resultados foram revelando não apenas a forte aptidão dos fatores de transcrição WRKY em regular a resposta da planta a diversas formas de estresses (Chen e col., 2011; Tripathi e col., 2013; Bakshi e Oelmüller, 2014; Banerjee e Roychoudhury, 2015) como também a interconexão entre essas redes de regulação, evidenciada pela intensa auto-regulação gênica e a redundância funcional que ocorre dentro da família. Paradoxalmente, essa complexidade pode ser um ponto vantajoso para o desenvolvimento de genótipos tolerantes à seca através da engenharia genética.

A seca é um estresse com resposta de base poligênica que desencadeia múltiplos efeitos na planta (Grillo e col., 2006). Portanto, considerando a pluralidade da família WRKY é possível prever que a manipulação de apenas um ou poucos genes serão suficientes para alterar rotas regulatórias inteiras e conferir a planta um perfil tolerante. Obviamente que o sucesso desse processo depende de uma escolha acurada do gene que será alvo da engenharia genética. E a acurácia dessa escolha, por sua vez, é completamente dependente do quanto a atuação dos fatores de transcrição WRKY nas resposta da planta aos estresses abióticos é conhecida e compreendida.

PROMOTORES

A compreensão plena do comportamento de um gene só é alcançada com o estudo de seu promotor. Em geral, considera-se como região promotora a fração de aproximadamente 1 a 4 kbp à montante do sítio de inicio da transcrição (Rombauts e col., 2003; Shahmuradov e col., 2005). No entanto, as formas variadas nas quais um promotor pode se apresentar, inclusive à jusante do sítio de inicio da transcrição como um dos íntrons, dificulta a definição dos seus limites (Rombauts e col., 2003). É nesta porção não codificante do gene que estão localizados os motivos onde se ligam os fatores de transcrição. Esses sítios de ligação são conhecidos como cis-elementos. São eles que carregam a informação sobre a potencial ativação ou repressão gênica, condição que só será definida pela interação com as proteínas reguladoras (Azad e col., 2011; Biłas e col., 2016)

Os promotores gênicos podem ser classificados em três diferentes grupos. Os constitutivos são aqueles que mantêm a expressão constante e ubíqua. Os espaço-temporais direcionam uma expressão específica para cada período e tecido. Já os promotores induzíveis dirigem a atividade do gene sob influência dos sinais químicos e físicos externos (Peremarti e col., 2010; Hernandez-Garcia e Finer, 2014).

Com o aumento da intensidade e frequência dos estresses ambientais, produto das mudanças climáticas mundiais (Bellard, 2012), uma atenção especial vem sendo dedicada à investigação dos promotores induzíveis. O objetivo é elucidar como a planta percebe os fatores ambientais estressores e como estes sinais são transmitidos até o acionamento dos mecanismos de tolerância ao estresse. É sobre esse entendimento que os programas de melhoramento se inclinam para elaborar estratégias que levem ao desenvolvimento de cultivares mais resilientes aos estresses abióticos, entre eles a seca.

A caracterização de novas sequências que possam ser aplicadas em engenharia genética é também um propósito no estudo dos promotores. A insuficiente variabilidade de sequências reguladoras disponíveis impede um ajuste fino da expressão dos transgenes . Por ser imprecisa e invariável, a condução da expressão gênica estabelecida por um promotor não induzível pode fazer com que o transgene

apresente um padrão de atividade completamente diferente do observado em seu organismo nativo, produzindo assim , um fenótipo não desejado (Hernandez-Garcia e Finer, 2014). Fenômeno esse já relatado por Liu e col. (1998), Yamaguchi-Shinozaki e Shinozaki (1999) e Chen e col. (2009) que observaram um severo retardo no crescimento e redução da produção de sementes em plantas modificadas geneticamente para expressar fatores de transcrição de forma constitutiva.

Sendo assim, mesmo muito relevantes para o progresso já alcançado na pesquisa em genética de plantas , os promotores dos genes da nopalina-sintetase e da octopina-sintetase , ambos oriundos de Agrobacterium tumefaciens, assim como o promotor 35S do vírus do mosaico da couve flor (CaMV 35S), não são, devido a natureza constitutiva, as melhores opções quando a intenção for um perfil de expressão consonante com os estímulos ambientais (Sassaki, 2008). A expressão constitutiva de trangenes que nativamente não seguem este padrão transcricional pode resultar em plantas de crescimento e desenvolvimento alterados, morfologia anormal e inesperada ativação de vias de defesa mesmo na ausência de estresses (Hernandez-Garcia e Finer, 2014).

Portanto o estudo de promotores é antes de tudo um esforço para preencher duas demandas ascendentes no contexto agrícola mundial. A primeira é a segurança dos alimentos geneticamente modificados. O uso de promotores induzíveis garantirá um controle mais efetivo da expressão dos transgenes, reduzindo, deste modo, os riscos de expressão indesejada, principalmente daqueles que conferem resistência a antibióticos e são usados para a seleção de transformantes (Huang e col., 2001). A segunda é a segurança alimentar da população . A disponibilidade de novas sequências reguladoras facilitará o desenvolvimento de genótipos que agregam características de alta produtividade e tolerância a estresses, assegurando assim qualidade e quantidade de alimentos suficientes para a nutrição da população.

REFERÊNCIAS

Azad, A. K. M., Shahid, S., Noman, N., Lee, H. (2011) Prediction of plant promoters based on hexamers and random triplet pair analysis. Algorithms for Molecular Biology *6*(1), 19

Bakshi, M., Oelmüller, R. (2014) WRKY transcription factors jack of many trades in plants. Plant Signaling and Behavior. 9, 1–18

Banerjee, A. e Roychoudhury, A. (2015) WRKY Proteins: Signaling and Regulation of Expression during Abiotic Stress Responses. The Scientific World Journal. ID 807560

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., and Courchamp, F. (2012) Impacts of climate change on the future of biodiversity. Ecology Letters. 15, 365–377

Berlato, M.A. Modelo de relação entre o rendimento de grãos de soja e o déficit hídrico para o Estado do Rio Grande do Sul. 1987. 93f. Tese (Doutorado em Meteorologia) Curso de Pós-graduação em Meteorologia, Instituto Nacional de Pesquisas Espaciais, 1987.

Biłas, R., Szafran, K., Hnatuszko-konka, K. (2016) *Cis*-regulatory elements used to control gene expression in plants. *Plant Cell, Tissue and Organ Culture. 127*(2), 269–287

Chen, M., Xu, Z., Xia, L., Li, L., Cheng, X., Dong, J., Wang, Q., Ma, Y. (2009) Cold-induced modulation and functional analyses of the DRE-binding transcription factor gene, GmDREB3, in soybean (*Glycine max L*.). Journal of Experimental Botany. 60(1), 121–35

Chen, L., Song, Y., Li, S., Zhang, L., Zou, C., Yu, D. (2011) The role of WRKY transcription factors in plant abiotic stresses. Biochimica et Biophysica Acta. 1819(2), 120–8

Chen, W., Provart, N.J., Glazebrook, J., Katagiri, F., Chang, H.S., Eulgem, T., Mauch, F., Luan, S., Zou, G., Whitham, S.A., Budworth, P.R., Tao, Y., Xie, Z., Chen, X., Lam, S., Kreps, J.A., Harper, J.F., Si-Ammour, A., Mauch-Mani, B., Heinlein, M., Kobayashi, K., Hohn, T., Dangl, J.L., Wang, X., and Zhu, T. (2002) Expression profile matrix of *Arabidopsis* transcription factor genes suggests their putative functions in response to environmental stresses. Plant Cell. 14, 559-574

EMBRAPA. Tecnologias de Produções de Soja – Região Central do Brasil 2008. Londrina: Embrapa Soja: Embrapa Cerrados: Embrapa Agropecuária Oeste, 2008. 208p. (Embrapa Soja, Sistema de Produções, 12)

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

Food and Agriculture Organization of the United Nations - FAO. (2018) *The impact of disasters and crises on agriculture and food security*.

Food and Agriculture Organization of the United Nations - FAO. (2011) *The State of the World's land and water resources for Food and Agriculture. Managing systems at risk*. *Food and Agriculture Organization*

Grillo, S., Blanco, A., Cattivelli, L., Coraggio, I., Leone, A., Salvi, S. (2006) Risposte genetico-molecolari delle piante a carenza idrica. *Stress: The International Journal on the Biology of Stress*. 617–638

Hernandez-Garcia, C. M., Finer, J. J. (2014) Identification and validation of promoters and *cis*-acting regulatory elements. *Plant Science*. *217–218*, 109–119

Huang N *et al.* (2001) The tissue-specific activity of a rice beta-glucanase promoter (*Gns9*) is used to select rice transformants. Plant Science. 161, 589-595

Ishiguro, S., Nakamura, K. (1994) Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions

of genes coding for sporamin and ß-amylase from sweet potato. Molecular Genetics and Genomics. 244, 563-571

Isoda, A. Adaptive responses of soybean and cotton to water stress I. Transpiration changes in relation to stomatal area and stomatal conductance. Plant Production Science, Tokyo, v.8, p.16-26, 2005.

Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura S., Yamaguchi-Shinozaki, K., Shinozaki, K (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought and low temperature responsive gene expression, respectively, in *Arabidopsis*. The Plant Cell. 10, 1391–1406

Marè, C., Mazzucotelli, E., Crosatti, C., Francia, E., Stanca, M.A., Cattivelli, L. (2004) Hv-WRKY38: a new transcription factor involved in cold- and drought-response in barley. Plant Molecular Biology. 55, 399-416

Meshi, T., Iwabuchi, M. (1995). Plant transcription factors. Plant Cell Physiology. *36*(8), 1405–1420

Peremarti, A., Twyman, R. M., Gómez-Galera, S., Naqvi, S., Farré, G., Sabalza, Miralpeix, B.,Dashevskaya, S.,Yuan, D.,Ramessar, K.,Christou, P., Zhu, C., Bassie, L., Capell, T. (2010) Promoter diversity in multigene transformation. Plant Molecular Biology. *73*, 363–378

Ramamoorthy, R., Jiang, S-Y., Kumar, N., Venkatesh, P.N., Ramachandran, S. (2008) A comprehensive transcriptional profiling of the WRKY gene family in rice under various abiotic and phytohormone treatments. Plant Cell Physiology. 49(6), 865-879.

Raventós, D., Jensen, A.B., Rask, M. B., Casacuberta, J.M., Mundy, J., San Segundo, B. (1995). A 20 bp cis-acting element is both necessary and sufficient to mediate elicitor response of a maize PRms gene. Plant Journal. 7, 147-156

Rombauts, S., Florquin, K., Lescot, M., Marchal, K., Rouze, P., Peer, Y. V. (2003) Computational approaches to Identify Promoters and *cis*- Regulatory Elements in Plant Genomes. Plant Physiology. 132, 1162–1176

Rushton, P. J., Somssich, I. E., Ringler, P., Shen, Q. J. (2010) WRKY transcription factors. Trends in Plant Science. 15(5), 247–258

Rushton, P. J., Macdonald H., Huttly, A. K., Lazarus, C.M., Hooley, R. (1995) Members of a new family of DNA-binding proteins bind to a conserved *cis*-element in the promoters of α-Amy2 genes. Plant Molecular Biology. 29, 691-702

Sassaki, F. T. (2008) Isolamento e caracterização de promotores órgão específicos a partir de informações do Banco FORESTs (*Eucalyptus* Genome Sequencing Project Consortium). Tese (Doutorado em Ciências Biológicas – Genética) - Instituto de Biociências, Universidade Estadual Paulista, Botucatu, 2008.

Shahmuradov, I. A., Solovyev, V. V., Gammerman, A. J. (2005) Plant promoter prediction with confidence estimation. Nucleic Acids Research. 33(3), 1069–1076

Singh, D., Laxmi, A. (2015) Transcriptional regulation of drought response: a tortuous network of transcriptional factors. Frontiers in Plant Science. *6*, 1–11

Taiz, L., Zeiger, E. Fisiologia vegetal. 5.ed. Porto Alegre- RS. Artmed, 2013. 918 p.

Tripathi, P., Rabara, R. C., Rushton, P. J. (2013) A systems biology perspective on the role of WRKY transcription factors in drought responses in plants. *Planta*. 239, 1-12

Yamaguchi-Shinozaki, K., Shinozaki, K. (1999) Improving plant drought, salt and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nature Biotechnology. 17, 287-291

Yin, G., Xu, H., Xiao, S., Qin, Y., Li, Y., Yan, Y., Hu, Y. (2013) The large soybean (*Glycine max*) WRKY TF family expanded by segmental duplication events and subsequent divergent selection among subgroups. BMC Plant Biology. 13, 148

Zhou, Q. Y., Tian, A. G., Zou, H. F., Xie, Z. M., Lei, G., Huang, J., Wang, C. M., Wang, H. W., Zhang, J. S., Chen, S. Y. (2008) Soybean WRKY-type transcription factor genes, *GmWRKY13*, *GmWRKY21*, and *GmWRKY54*, confer differential tolerance to abiotic stresses in transgenic *Arabidopsis* plants. Plant Biotechnology Journal. 6, 486-503

OBJETIVOS

OBJETIVO GERAL

O objetivo geral deste trabalho é a caracterização de genes WRKY de soja e seus promotores em resposta ao déficit hídrico e estresses relacionados .

OBJETIVOS ESPECÍFICOS

1) Determinação do perfil de transcrição de genes WRKY de soja em resposta a desidratação.

2) Sequenciamento e comparação das regiões promotoras de genes WRKY isoladas das cultivares de soja tolerante (EMBRAPA48) e suscetível (BR16) à seca;

3) Avaliação da atividade dos promotores em condições de estresse;

4) Obtenção de plantas de Arabidopsis thaliana transgênicas que expressem um gene repórter direcionado pelos promotores WRKY em estudo.

5) Desenvolvimento de ferramentas biotecnológicas para estudo dos promotores de genes WRKY de soja.

CAPÍTULO 1

The differential expression of soybean [Glycine max (L.) Merrill] WRKY genes in response to water deficit

---

Artigo publicado no *Plant Physiology and Biochemistry Journal* em junho de 2016

Plant Physiology and Biochemistry 107 (2016) 288–300

Contents lists available at ScienceDirect

Plant Physiology and Biochemistry

journal homepage: www.elsevier.com/locate/plaphy

Research article

The differential expression of soybean [*Glycine max* (L.) Merrill] WRKY genes in response to water deficit

Letícia Pereira Dias, Luisa Abruzzi de Oliveira-Busatto, Maria Helena Bodanese-Zanettini*

Programa de Pós-Graduação em Genética e Biologia Molecular, Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul (UFRGS), 91501-970, Porto Alegre, RS, Brazil

ARTICLE INFO

Article history:
Received 11 March 2016
Received in revised form
13 June 2016
Accepted 14 June 2016
Available online 16 June 2016

Keywords:
WRKY genes
Drought stress
Expression profile
Cis-elements
Coexpression analysis

ABSTRACT

Drought is today, and perhaps even more in the future, the main challenge for grain crops, resulting in a drastic yield reduction. Thus, it is of great interest to obtain soybean genotypes tolerant to water deficit. The drought tolerance trait is difficult to obtain through classical breeding due to its polygenic basis. In this context, genetic engineering is presented as a way to achieve this attribute. The ability to modulate the expression of many genes placed the transcription factors as promising biotechnological targets to develop stress tolerant cultivars. The WRKY proteins form a large family of transcription factors that are involved in important physiological and biochemical processes in plants, including the response to water deficit. In this study, the expression pattern determined by qPCR showed that, GmWRKY6, GmWRKY46, GmWRKY56, GmWRKY106 and GmWRKY149 genes are differentially expressed between a drought tolerant and a susceptible soybean genotype in water stress conditions. The *in silico* promoter and coexpression analysis indicate that these genes act in a stress physiological background.

1. Introduction

Soybean crop (*Glycine max* (L.) Merrill) has a great importance for the world economy. Its grains have a high versatility, giving rise to products and sub products widely used by agribusiness (production of oilseed and animal feed), chemical and food industry. More recently, soybean use is also increasing as an alternative source of biofuel (Masuda and Goldsmith, 2009).

Drought stress is a major constraint to the soybean production. Thus, the goal of plant breeding is the development of new cultivars by accumulating genes that contribute to dehydration stress tolerance (Varshney et al., 2011). However, considering that the response to stress involves a large genetic network and complex molecular mechanisms, a critical point is to define which and how many genes are necessary to combine in one plant in order to make it tolerant. Given the difficulty in deciphering this puzzle prompts genetic engineering using regulatory genes as pieces. This biotech strategy requires the introgression of only one or few genes for development of tolerant genotypes.

Among the various classes of regulatory genes, those encoding transcription factors are a wise choice to implement that strategy. This is due to the action of transcription factors in the recruitment/blocking of RNA polymerases to DNA, thereby activating/repressing gene transcription. It makes them true orchestrators of plant response to environmental variations (Udvardi et al., 2007). The signal transduction pathways can ultimately regulate the expression of drought responsive genes through transcription factors (Ku et al., 2010). In the soybean genome, 4342 loci encoding 5035 transcription factors protein models were *in silico* annotated (Mochida et al., 2009), among them 176 belong to the WRKY family (Song et al., 2016).

The WRKY transcription factors comprise one of the largest families of regulatory proteins in plants (Eulgem and Somssich, 2007). The WRKY proteins, named because of the invariant WRKY amino acid sequence at the N-terminus, are a group of zinc-fingers proteins with a Cx4–5Cx22–23HxH or Cx7Cx23HxC structure (Eulgem et al., 2000). Besides many developmental processes such as seed germination, flower development, secondary metabolism, morphogenesis, dormancy and senescence, WRKY genes are involved in regulating the plant response to stress (Journot-Catalino et al., 2006; Li et al., 2006; Wang et al., 2006; Zheng et al., 2006; Feng et al., 2012). Studies in several plant species

* Corresponding author. Universidade Federal do Rio Grande do Sul, Instituto de Biociências, Departamento de Genética, Av. Bento Gonçalves 9500, Bloco 3, Prédio 43312 Sala 212D, Campus do Vale, CEP:91501-970, Porto Alegre, RS, Brazil.
E-mail addresses: dias.pereira.leticia@gmail.com (L.P. Dias), luisaabruzzi@gmail.com (L.A. de Oliveira-Busatto), maria.zanettini@ufrgs.br (M.H. Bodanese-Zanettini).

http://dx.doi.org/10.1016/j.plaphy.2016.06.018
0981-9428/

have reported the WRKY genes differential expression under drought stress conditions (Ramamoorthy et al., 2008; Zhou et al., 2008; Berri et al., 2009; Wu et al., 2009; Qiu and Yu, 2009; Wang et al., 2009; Ren et al., 2010; Xiong et al., 2010; Shekhawat et al., 2011; Jiang et al., 2012; Niu et al., 2012; Luo et al., 2013; Wang et al., 2013), but few studies have been performed in soybean. Recently, based on RNA-seq data a total number of 31 soybean WRKY genes were found to be differentially expressed under dehydration treatment (Song et al., 2016).

The wide involvement of WRKY proteins in plant physiological processes makes the elucidation of their coexpression network very operative in unveil molecular subsystems underlying complex traits (Rinerson et al., 2015; Choura et al., 2015; Wanke et al., 2010; Berri et al., 2009). Boosted by extensive generation of microarray data, the coexpression analysis has been very useful for the study of plant biology (Wei et al., 2006; Aoki et al., 2007; Saito et al., 2008; Wanke et al., 2010; Ficklin and Feltus, 2011; Deihimi et al., 2012). Although coexpressed genes do not always establish regulatory relationship (Stuart et al., 2003), the "Guilt-by-Association" principle allows us to infer similarity of function among genes which present a coordinated pattern of expression. This heuristic principle is based upon the fact that the behavior of an organism in a condition/stage is dependent on both the action of many genes as well as their interdependence. Therefore, genes expressed jointly are most likely under the control of a same regulatory system (Saito et al., 2008). This becomes even more reasoned when the guide-gene of coexpression analysis encodes a transcription factor as WRKY.

In this paper, experimental and *in silico* approaches were combined to better understand the expression demeanor of ten WRKY transcription factors in soybean under water deficit conditions.

2. Material and methods

2.1. Selection of WRKY genes

Ten WRKY genes induced in drought conditions were selected based on the serial analysis of gene expression (SuperSAGE) from the Soybean Genome Project database – UNICAMP (http://bioinfo03.ibi.unicamp.br/soja/). According to *Glycine max* genome v1.0 from Phytozome (Goodstein et al., 2012; Schmutz et al., 2010), the access number of the selected genes are: GmWRKY5 (Glyma01g31921), GmWRKY6 (Glyma08g15050), GmWRKY11 (Glyma05g20710), GmWRKY46 (Glyma05g36970), GmWRKY56 (Glyma08g23380), GmWRKY90 (Glyma05g29310), GmWRKY106 (Glyma07g02630), GmWRKY120 (Glyma09g37470), GmWRKY125 (Glyma09g41050) and GmWRKY149 (Glyma15g11680) (http://www.phytozome.net/). To validate the gene expression patterns, transcriptional analysis by qPCR were performed.

2.2. Plant material

In an experiment conducted at Embrapa Soja, Londrina, Paraná, Brazil, the BR16 and EMBRAPA48 soybean cultivar plants, highly and slightly sensitive to dehydration stress, respectively (Oya et al., 2004), were cultivated hydroponically in a greenhouse at 25 ± 2 °C and 60± 5% relative humidity, as described by Kulcheski et al. (2011). Briefly, seedlings in V2 stage were removed from hydroponic solution and kept in a tray in darkness and without nutrients. Leaves and roots were collected at 0 (control), 25, 50, 75, 100, 125, and 150 min after the initiation of dehydration stress, frozen in liquid nitrogen and stored at −80 °C until RNA extraction. Three biological replicates (three plants/replicate) were sampled for each organ/genotype/treatment point. The RNA extraction from each sample and the cDNA synthesis reaction were performed according to the procedure described by Cabreira et al. (2013). To evaluate the relative gene expression, the first-strand cDNA reaction product was diluted 1:100.

2.3. Gene expression analysis

The gene expression was analyzed using real-time quantitative polymerase chain reaction (qPCR). Primers were designed using Primer3 software (http://frodo.wi.mit.edu/) (Koressaar and Remm, 2007; Untergasser et al., 2012) according to the genes sequences (Table 1). The reactions were performed in 25 μl volume containing 10 μM of each primer, 12.5 μl of diluted cDNA sample, 1 × PCR buffer, 50 mM MgCl2, 10 mM of each dNTP, 2.5 μl SYBR-Green solution (1:100,000, Molecular Probes Inc., Eugene, OR) and 0.06 U Platinum Taq DNA Polymerase (Life Technologies). qPCR was carried out on StepOne Applied Biosystem Real-Time Cycler in a 96-well plate reaction. The cycling conditions were implemented as follows: 5 min at 94 °C for an initial denaturation, 40 cycles of a 10 s denaturation step at 94 °C, a 15 s annealing step at 60 °C, and a 15 s extension step at 72 °C ending with 2 min at 72 °C for a final extension. A melting curve analysis was performed at the end of the PCR run over a range of 55–99 °C, increasing the temperature stepwise by 0.1 °C every 1 s. Technical quadruplicate reactions were done for each sample. The f-box (Glyma12g05510), ACTII (Glyma18g52780), and ELF (Glyma02g44460) housekeeping genes (Jian et al., 2008) were used for data normalization. The relative expression fold changes were determined using the $2^{-\Delta\Delta Ct}$ method described by Livak and Schmittgen (2001).

2.4. Statistical analysis

To allow appropriate comparisons between genotypes, measured expression levels were calibrated by the control sample (T0) transcript level of BR16 leaves or roots. When comparing organs within genotypes, the expression data were calibrated by the transcript level of BR16 or EMBRAPA48 leaves at T0. The comparisons were performed by analysis of variance, using Ordinary Least Square method to improve the variance homogeneity. To correct for inflation in type I error, means were compared using the Bonferroni multiple comparison test. The SAS 9.2 and the SPSS/PASWSTAT 18

Table 1
Primers used in the RT-qPCR to determine the expression profile of WRKY soybean genes.

Genes	Prímers (5′→3′)	
GmWRKY5	Forward	GGTGGAGGAAATATGGACAGA
	Reverse	TGTGATTACGACTTATGAAGGGAA
GmWRKY6	Forward	GTCCAAGGATCGCGTTTAGA
	Reverse	GAAGAAAAGGGTGGAAAGGG
GmWRKY11	Forward	GATGGCTGATATTCCACCAGA
	Reverse	AACCTACGAGGGAGAGCACA
GmWRKY46	Forward	CCCAAATGGATGGATCATGTA
	Reverse	GCAACAAAGCAAGTGCAGAG
GmWRKY56	Forward	TGGACAAAAGGTGACCAGAG
	Reverse	AGCACAATCATCCTCAGGCT
GmWRKY90	Forward	ATGGGCATGGAGAAAGTACG
	Reverse	TTACTCCTCCGACCACAACC
GmWRKY106	Forward	AGCTGCCCTGTCAAAAAGAA
	Reverse	TGTACCGCATCTCTCAGCAC
GmWRKY120	Forward	GCCAGTGGAGGAAATATGGA
	Reverse	CCTCTTCCAGTTTCAGCCAC
GmWRKY125	Forward	CCTGAATGCCAAATTCCCTA
	Reverse	CATACTTGACCCCATGTCCC
GmWRKY149	Forward	ATGGATGCCAATGGAGAAAA
	Reverse	ACAACATATGAAGGCACGCA

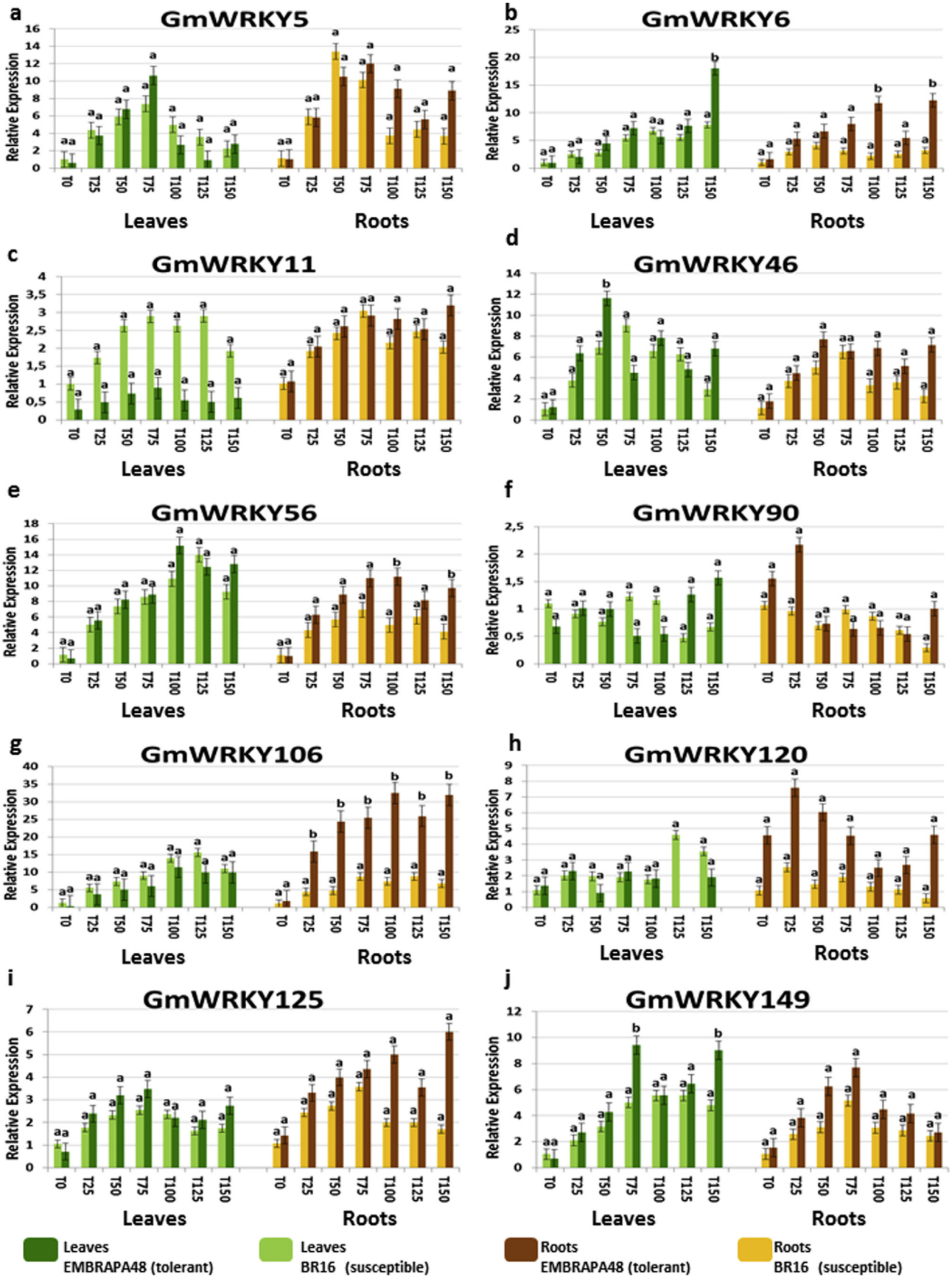

Fig. 1. Comparison of WRKY genes expression levels between cultivars (BR16 X EM48) within each organ in response to dehydration stress. The relative expression levels of genes were measured by qPCR at T0 (control), T25, T50, T75, T100, T125, and T150 min of dehydration stress. The f-box, ACTII, and ELF reference genes were used as internal controls to normalize the amount of mRNA present in each sample. All transcript levels were calibrated in relation to BR16 leaves/roots at T0. Data represent means of three biological replicates with four technical replicates each. Means labeled with the same letter do not differ significantly (Bonferroni multiple comparison test, $p < 0.05$) between genotypes. Error bars represent standard errors.

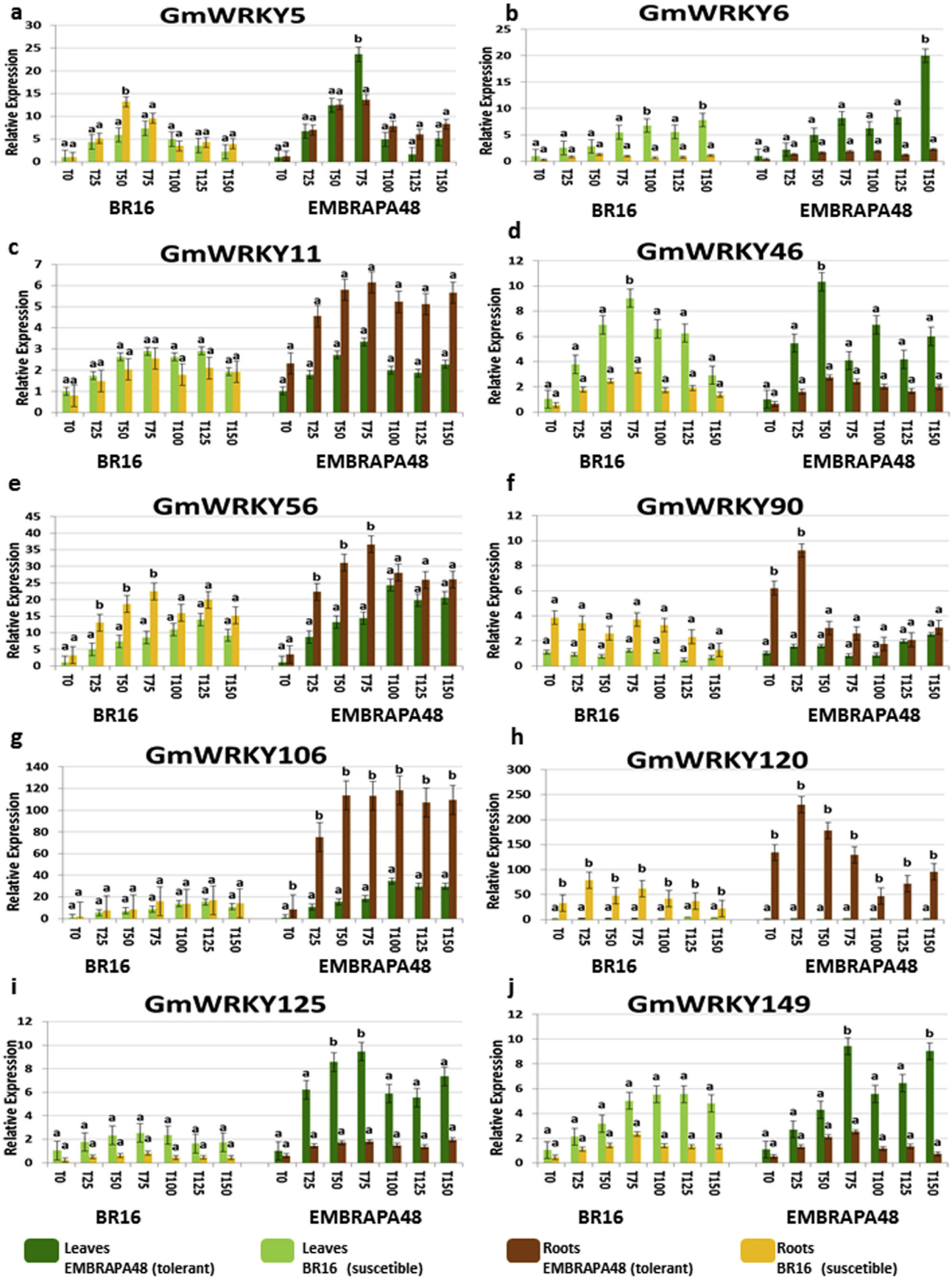

Fig. 2. Comparison of WRKY genes expression levels between organs (leaves × roots) within each cultivar in response to dehydration stress. The relative expression levels of the genes were measured by qPCR at T0 (control), T25, T50, T75, T100, T125, and T150 min of dehydration stress. The f-box, ACTII, and ELF reference genes were used as internal controls to normalize the amount of mRNA present in each sample. All transcript levels were calibrated in relation to BR16/EM48 leaves at T0. Data represent means of three biological replicates with four technical replicates each. Means labeled with the same letter do not differ significantly (Bonferroni multiple comparison test, $p < 0.05$) between genotypes. Error bars represent standard errors.

Table 2
Number of copies and biological function of the putative *cis*-elements related to water stress that were identified in the WRKY genes promoters.

Cis-element	Consensus sequence	Transcription factor	Function	GmWRKY5	GmWRKY6	GmWRKY11	GmWRKY46	GmWRKY56	GmWRKY90	GmWRKY106	GmWRKY120	GmWRKY125	GmWRKY149	References
MYBR	YAACKG	MYB2CONSENSUSAT	MYB recognition site found in the promoters of the dehydration-responsive gene rd22	2	4	1	1	1			1	2		Abe et al. (2002)
MYBR	WAACCA	MYB1AT	MYB recognition site found in the promoters of the dehydration-responsive gene rd22	4	3	4	2	2	4	2	2	2	1	Abe et al. (2002)
MYBR	CTAACCA	AtMYB2	MYB recognition site found in the promoters of the dehydration-responsive gene rd22	1										Abe et al. (1997)
MYBR	TAACTG	MYB2AT	Binding site for MYB plant protein (ATMYB2)		2			1	2		2		1	Urao et al. (1993)
MYBR	CNGTTR	MYBCORE	involved in regulation of water stress responsive genes	5	7	2	3	4	3	3	7	3	2	Abe et al. (1997)
MYCR	CANNTG	MYCCONSENSUSAT	MYC recognition site found in the promoters of the dehydration-responsive gene rd22	3	8	2	6	5	7	5	8	6	6	Abe et al. (2002)
MYCR	CACATG	AtMYC2	MYC recognition site found in the promoters of the dehydration-responsive gene rd22	1				3		1		2		Abe et al. (1997)
MYCR	CATGTG	MYCATERD1	MYC recognition sequence; necessary for expression of erd1 in dehydrated *Arabidopsis*			1					1	3		Simpson et al. (2003)
ABRE	ACGTGKC	ABRE	Cis-elements found in the promoters of the ABA-regulated genes and recognized by bZIP proteins									1		Shinozaki and Yamaguchi-Shinozaki (1997)
ABRE	BACGTGKM	ABRE-like			2							1	1	Shinozaki and Yamaguchi-Shinozaki (2000)
WBOX	TTGAC	WBOXATNPR1	W-box found in the promoter of *Arabidopsis* NPR1 gene; recognized specifically by WRKY DNA binding proteins	5	6	4	8	3	3	4	5	5	7	Yu et al. (2001)

Table 3
Coexpression analysis.

Guide-gene	Coexpression strenght (MR)	MR < 10 genes	*Arabidopsis* orthologs	Orthology -based annotation	(Possible) involvement in stress response	Also belonging to the expression network[a] of				
						GmWRKY46	GmWRKY106	GmWRKY6	GmWRKY56	GmWRKY149
GmWRKY6										
	1.0	Glyma06g42950	AT4G34370	ARI1 (Ubiquitin protein ligase ARIADNE 1)	Signaling			X		
	2.8	Glyma20g26600	AT3G05110	Putative Heat Shock Protein	Heat shock protein			X		
	2.8	Glyma16g28970	AT5G24090	CHIA (Chitinase A - Class III)	Stress protein			X		
	3.0	Glyma10g40720	AT2G15220	BSP (Plant basic secretory protein)	Stress protein			X		
	4.9	Glyma16g27250	AT1G34420	Protein kinase	Signaling			X		
	5.3	Glyma15g06790	AT2G14580	PRB1/ATCAPE7 (Basic Pathogenesis Related Protein 1)	Stress protein/ Signaling			X		
	7.3	Glyma01g03910	AT4G06546	Gypsy-like retrotransposon family	Gene expression control			X		
	8.0	Glyma18g49540	AT5G39670	Calcium-binding EF-hand family protein	Signaling			X		
	8.1	Glyma08g04920	AT1G73805	SARD1 (SAR Deficient 1)	Growth regulator			X		
GmWRKY46										
	1.4	Glyma08g02580	AT4G26640	Transcription factor WRKY20	Gene expression control	X	X		X	
	3.5	Glyma01g31921	AT5G56270	Transcription factor WRKY2	Gene expression control	X				X
	3.9	Glyma14g01990	AT2G40140	Transcription factor ATSZF2 (Salt-Inducible Zinc Finger 2)	Gene expression control	X				
	4.0	Glyma10g09720	AT2G28250	Protein kinase	Signaling	X	X			X
	4.0	Glyma06g14090	AT5G44650	AtCEST (Chloroplast Protein-Enhancing Stress Tolerance)	Reactive oxygen scavenging	X				
	4.6	Glyma20g25990	AT3G03280	Unknown protein	……	X				
	4.9	Glyma07g37790	AT3G29300	Unknown protein	……	X				X
	5.9	Glyma12g12260	AT5G54710	Ankyrin protein - Membrane adaptor	Signaling	X	X		X	X
	6.9	Glyma06g44430	AT3G15210	Transcription factor ERF4 (Ethylene Responsive Element Binding Factor 4)	Gene expression control	X		X		
	7.3	Glyma12g09830	AT5G22250	Deadenylase complex subunit CAF1B (CCR4-Associated Factor 1B)	Gene expression control	X				
	8.8	Glyma05g28280	AT5G64280	Chloroplast envelope transporter DIT2.2 (Dicarboxylate Transporter 2.2)	Ion/proton transporter	X				
	9.6	Glyma05g36980	AT1G64355	Unknown protein	……	X	X		X	
GmWRKY56										
	1.4	Glyma07g02630	AT1G80840	Transcription factor WRKY40	Gene expression control	X			X	
	1.7	Glyma10g08540	AT2G05580	Glycine-rich protein family	Stress protein	X	X		X	
	3.2	Glyma20g10920	AT1G14370	Protein kinase	Signaling		X		X	
	3.9	Glyma15g00570	AT1G80840	Transcription factor WRKY40	Gene expression control	X	X		X	X
	4.0	Glyma08g12530	AT5G62480	ATGSTU9 (Glutathione S-Transferase)	Reactive oxygen scavenging		X	X	X	X
	4.5	Glyma04g04500	AT3G17420	Protein kinase	Signaling		X		X	X
	4.9	Glyma07g27370	AT5G24080	Protein kinase	Signaling		X		X	
	5.1	Glyma17g36530	AT2G23760	Transcription factor SAW2	Gene expression control	X	X		X	
	6.3	Glyma15g01990	AT1G52200	PLAC8 family protein	Reactive oxygen scavenging		X		X	X
	6.3	Glyma08g26830	AT2G12980	CACTA-like transposase family	Gene expression control	X	X	X	X	X
	6.6	Glyma15g40290	AT1G78380	Glutathione S-Transferase	Reactive oxygen scavenging		X	X	X	
	7.0	Glyma13g38630	AT1G62300	Transcription factor WRKY6	Gene expression control		X		X	X
	7.2	Glyma02g41490	AT1G07570	Protein kinase	Signaling				X	X
	7.7	Glyma19g21690	AT1G10740	a/b Hydrolase fold protein	Stress protein		X		X	
	8.1	Glyma20g27470	AT1G2751	EX2 (Executer 2)	Reactive oxygen scavenging		X		X	
	8.8	Glyma06g01240	AT4G34588	CPUORF2 (Conserved Peptide Upstream Open Reading Frame)	Gene expression control		X		X	X

(*continued on next page*)

Table 3 (*continued*)

Guide-gene	Coexpression strenght (MR)	MR < 10 genes	*Arabidopsis* orthologs	Orthology -based annotation	(Possible) involvement in stress response	Also belonging to the expression network[a] of				
						GmWRKY46	GmWRKY106	GmWRKY6	GmWRKY56	GmWRKY149
	8.8	Glyma13g34100	AT1G53420	Protein kinase	Signaling		X		X	
	9.9	Glyma02g45970	AT5G36930	Disease resistance protein (TIR-NBS-LRR class)	Signaling		X		X	
GmWRKY106										
	1.4	Glyma08g23380	AT1G80840	Transcription factor WRKY40	Gene expression control	X	X			X
	5.3	Glyma18g03190	AT2G40095	a/b Hydrolase fold protein	Stress protein	X	X		X	X
	5.7	Glyma04g04500	AT3G17420	Protein kinase	Signaling		X		X	X
	8.5	Glyma11g29720	AT2G38470	Transcription factor WRKY33	Gene expression control		X	X	X	X
	9.2	Glyma15g00570	AT1G80840	Transcription factor WRKY40	Gene expression control	X	X		X	X
GmWRKY149										
	1.0	Glyma09g00820	AT4G04450	Transcription factor WRKY42	Gene expression control	X	X		X	X
	3.2	Glyma14g00380	AT2G17220	Protein kinase	Signaling		X			X
	5.3	Glyma14g27290	AT1G21450	Transcription factor SCL1 (SCARECROW-like 1)	Gene expression control	X				X
	6.5	Glyma06g16440	AT5G63790	Transcription factor NAC102	Gene expression control		X	X		X
	7.9	Glyma19g24160	AT4G09430	Disease resistance protein (TIR-NBS-LRR class)	Signaling				X	X
	8.5	Glyma02g39870	AT2G38470	Transcription factor ATWRKY33	Gene expression control	X			X	X
	9.2	Glyma13g38630	AT1G62300	Transcription factor ATWRKY6	Gene expression control		X		X	X
	9.4	Glyma13g28090	AT3G23570	a/b Hydrolase fold protein	Stress protein	X				X
	9.9	Glyma17g13750	AT1G67580	Protein kinase	Signaling	X	X			X

Each network is composed for 300 coexpressed genes and was generated based on SFGD – Soybean Functional Genomics Database (http://bioinformatics.cau.edu.cn/SFGD/).
[a] Coexpressed genes including those presenting MR > 10.

software packages were used for the analysis.

2.5. Promoters in silico analysis

The putative promoter region from the 2000-base pairs (bp) upstream of the transcription start site (TSS) of each WRKY gene was used to search for putative *cis*-elements. The analysis was performed using the Plant Pan database (Chang et al., 2008) (http://plantpan2.itps.ncku.edu.tw/).

2.6. Coexpression analysis using public transcriptome datasets

The microarray data (255 samples from 14 microarray expression experiments) available in SFGD – Soybean Functional Genomics Database (Yu et al., 2014) (http://bioinformatics.cau.edu.cn/SFGD/) was used to identify the genes coexpressed with GmWRKY6, GmWRKY46, GmWRKY56, GmWRKY106 and GmWRKY149. A coexpression network composed by 300 genes was set for each WRKY gene and assessed based on mutual ranking (MR), a coexpression coefficient. The genes presenting MR < 10 were selected for a closer analysis that includes an annotation based on *Arabidopsis* orthologs and a survey of SuperSAGE data (Soybean Genome Project database – http://bioinfo03.ibi.unicamp.br/soja/).

3. Results and discussion

3.1. Five of the ten analyzed genes exhibit differential expression between genotypes and organs under plant dehydrating conditions

Consistently with SuperSAGE data, the qPCR results revealed that GmWRKY5, 6, 11, 46, 56, 90, 106, 120, 125 and 149 expression was increased under dehydration stress. In agreement, Song et al. (2016) showed, based on ESTs libraries and RNA-Seq data, that GmWRKY5, 6, 46, 56, 106 and 120 genes expression is responsive to water deficit stress.

The transcriptional profiles of GmWRKY6, GmWRKY46, GmWRKY56, GmWRKY106 and GmWRKY149 showed higher expression levels in EMBRAPA48, the tolerant cultivar, when compared to the BR16, the susceptible one, at least at one of the six time points under dehydration (Fig. 1). In leaves, statistically significant ($p < 0.05$) higher transcription levels were observed at, T150 for GmWRKY6 (Fig. 1b), T50 for GmWRKY46 (Fig. 1d), and at T75 and T150 for GmWRKY149 (Fig. 1j). In roots, this pattern was observed at T100 and T150 for GmWRKY6 and GmWRKY56 (Fig. 1b and e) and at all periods of plant exposure to drought stress for GmWRKY106 (Fig. 1g). For all other genes analyzed, no significant differences were detected between the susceptible and the tolerant cultivar (Fig. 1a, c, f, h and i). The statistically significant increase in the expression of GmWRKY6, GmWRKY46, GmWRKY56, GmWRKY106 and GmWRKY149 genes in cultivar EMBRAPA48 compared to BR16 indicates a robust performance of these genes in building a profile tolerant to drought.

A comparison of gene expression levels between organs (leaves × roots) within each cultivar was performed (Fig. 2). GmWRKY6, GmWRKY46 and GmWRKY149 genes presented some points with higher expression in leaves of both genotypes (EMBRAPA48 and BR16). In this organ, GmWRKY6 showed significantly higher expression at T150 in the tolerant genotype and at T100 and T150 in the susceptible (Fig. 2b). GmWRKY46 expression was higher at T75 in BR16 and at T50 in EMBRAPA48 (Fig. 2d). For GmWRKY149, the expression was significantly higher at T75 and T150 in EMBRAPA48 (Fig. 2j). On the other hand, higher expression level was observed in roots for GmWRKY56 and GmWRKY106 genes. For GmWRKY56,

the expression was higher in roots of both cultivars from T25 to T75 (Fig. 2e). GmWRKY106 gene expression in roots of EMBRAPA48 plants surpassed the expression in leaves at all evaluated points (Fig. 2g). For the remaining genes, significant differences of transcriptional levels between organs were also detected (Fig. 2a, f, h and i). However, all of them were statistically deemed as equally expressed between susceptible and tolerant genotypes (Fig. 1a, f, h and i).

Genes expressed differentially in roots under drought conditions are objects of great interest to crop breeding. Chemical signals are transmitted from the root tissues that inform the plant on the soil water status (Schachtman and Goodger, 2008). In dry conditions, these signals regulate stomatal behavior, leaf initiation, leaf expansion and other developmental processes (Davies and Zhang, 1991). The GmWRKY106 gene stood out due to the higher expression in roots of the tolerant cultivar. This gene emerges as candidate for genetic engineering strategies aiming the development of soybean cultivars tolerant to dehydration. Its ortholog in *Arabidopsis*, WRKY40 (Yin et al., 2013), was described as a negative regulator of ABA response (Chen et al., 2010). This was supported by the interaction of WRKY40 with the promoters of ABA responsive genes, ABI4 and ABI5, repressing their expressions (Liu et al., 2012). In addition, an increase in root length was observed in *Arabidopsis* lines overexpressing WRKY40 (Chen et al., 2010).

The higher expression of GmWRKY6, GmWRKY46 and GmWRKY149 genes in leaves of the tolerant cultivar may indicate their involvement in regulatory routes that lead to minimizing the water deficit damaging effects like reductions in CO_2 assimilation and leaf area (Stolf et al., 2009). The GmWRKY46 ortholog gene in *Arabidopsis*, WRKY41, has been identified as a flagellin-inducible gene by Higashi et al. (2008), conferring increased resistance against the bacteria *Pseudomonas syringae* when overexpressed. Concerning WRKY6, GmWRKY149 ortholog gene in *Arabidopsis*, its role as a component of senescence-inducing signaling complex and defense responses was demonstrated by Robatzek and Somssich (2001).

Besides being the main mechanism used by the plant to avoid water loss, the stomatal closure is part of the innate immune response in restricting bacterial invasion (Melotto et al., 2006). In the same way, the decline in photosynthetic activity, besides being one of the effects of drought, is among the initiator factors of the leaf senescence (Zentgraf et al., 2010). Facing these evidences, it is not only common but also expected that the plant senescence, bacterial attack and drought responses pathways are interconnected by the performances of transcription factors as convergence points.

3.2. WRKY genes promoters harbor a notable number of W box binding sites and cis elements related to water deficit stress

The *cis*-acting elements are conserved regions of gene promoters that act as transcription factors' binding sites; as such, crucial pieces of this complicated gears system that constitute the gene regulation (Yamaguchi-Shinozaki and Shinozaki, 2005). The definition of the correspondence between the conserved promoter motifs and the behavior of gene expression has provided advances in stress comprehension at the molecular level (Chen et al., 2012).

The *in silico* analysis showed that the putative promoters of studied genes contain *cis* elements related to the regulation of several physiological processes such as the plant response to phytohormones (ABA, gibberellic acid, salicylic acid, auxin), infection by pathogens, deficiency of nutrients (phosphorus, sulfur), light stress, high temperature stress and salinity. *Cis*-acting elements associated with roots nodulation and flowering were also found. As expected, *cis*-elements involved in plant response to water stress as well as the W-box, the DNA binding site of WRKY transcription factors, are present in all investigated putative promoters (Table 2).

The presence of W-box reveals the complex regulatory network established by the components of this transcription factors family that involves not just the interaction with other proteins but also the self-regulation previously reported (Eulgem et al., 1999; Cormack et al., 2002; Robatzek and Somssich, 2002; Dong et al., 2003; Song et al., 2016).

Taking into account the well-known involvement of the WRKY family in plant response to stress, efforts have been made to understand the relationship between these transcription factors and ABA. This phytohormone is referred to as the "stress hormone", due to its massive accumulation in the plant under environmental stress conditions (Davies and Zhang, 1991; Oliveira et al., 2010). The participation of specific WRKY genes in ABA signaling pathways has been demonstrated (Jiang and Yu, 2009; Ren et al., 2010; Shang et al., 2010). Under drought conditions, the ABA operates mainly promoting stomatal closure, thus reducing water loss by transpiration (Campalans et al., 1999). There are at least four pathways of gene expression induction in response to drought, two of them are ABA-dependent and contain ABRE (ABA responsive element), MYBR (MYB recognition sequence) and MYCR (MYC recognition sequence) as key *cis*-acting elements (Shinozaki and Yamaguchi-Shinozaki, 1997).

The presence of MYBR and MYCR *cis*-acting elements are remarkable in the studied promoters (Table 2), suggesting a major involvement of MYC and MYB transcription factors in the WRKY genes regulation. The direct interaction between MYBR and MYB2-box-like elements has been reported in the promoters of osmotic, drought, and ABA-induced genes (Urao et al., 1993; Jung et al., 2008). In addition, it has been observed that a significant over-representation of MYBR in the promoter regions of pathogen/salicylic acid regulated and unregulated *Arabidopsis* WRKY genes, indicating that MYB proteins may play a role in regulating the basal and constitutive expression of WRKY genes (Dong et al., 2003). The MYB2-box-like elements were the second most common transcription factor binding site motif found in the *in silico* analysis of 3922 plant promoters (monocotyledons and dicotyledons) carried out by Liu et al. (2013).

In the present study, the investigation of putative WRKY promoters did not reveal the presence of *cis*-elements related to ABA independent pathways of response to water deficit, corroborating the role of WRKY proteins as regulatory factors of ABA's action in plants, as revised by Rushton et al. (2012).

3.3. Coexpression networks

The term "coexpression" means the likeness of gene expression patterns across a variety of transcriptome modulations (Aoki et al., 2007). So, the range of experiments from which a coexpression network is constructed has a great influence on how much stronger the Guilt-by-association inferences are. The capture of stable coexpression relationships becomes more reliable when based in a diverse collection of experimental perturbations. For this purpose, the use of publicly available DNA microarray datasets is very helpful (Saito et al., 2008; Ficklin and Feltus, 2011).

Through the knowledge of the co-induced genes, it is possible to decipher the physiological background in which the guide-gene is inserted, expanding the comprehension of the transcriptional profile. The GmWRKY6, GmWRKY46, GmWRKY56, GmWRKY106 and GmWRKY149 genes were used as guide-genes to construct the coexpression networks (Table 3) in order to better understand why they presented a differential expression under stress.

Used as one of coexpression coefficients, the Mutual Ranks (MR) stands out for not only measuring expression similarity, but also

Table 4
SuperSAGE data of coexpressed genes.

(Possible) Involvement in stress response	MR<10 genes	SuperSAGE Data		
		Drought		Asian Rust
		EMBRAPA48	BR16	
Signaling				
	Glyma10g09720	■	■	
	Glyma04g04500	■		
	Glyma12g12260	■	■	
	Glyma06g42950			■
	Glyma20g10920		■	
	Glyma04g04500	■		■
	Glyma07g27370			■
	Glyma13g34100	■		■
	Glyma14g00380			■
	Glyma19g24160	■		
	Glyma17g13750	■	■	■
Stress protein				
	Glyma15g06790			■
	Glyma16g28970		■	
Gene expression control				
	Glyma08g02580	■		
	Glyma01g31921	■	■	
	Glyma14g01990	■	■	
	Glyma12g09830	■	■	
	Glyma08g23380	■		
	Glyma15g00570	■	■	■
	Glyma07g02630	■	■	
	Glyma15g00570	■	■	■
	Glyma17g36530	■	■	
	Glyma09g00820	■	■	■
	Glyma14g27290	■	■	■
	Glyma06g16440	■	■	
Ion/proton transporter				
	Glyma05g28280	■	■	
Reactive oxygen scavenger				
	Glyma08g12530	■		
	Glyma15g01990	■		
Unknown protein				
	Glyma20g25990	■	■	
	Glyma07g37790	■	■	

■ Down-regulated ■ Up-regulated

considering a biological significance between gene pairs of interest. The "biological significance" is an indication of how much two genes are functionally related. The lower numerical amount of the MR, the greater is the strength of a gene coexpression relationship (Obayashi and Kinoshita, 2009).

The annotation based on *Arabidopsis* orthologs indicated that the soybean genes presenting MR < 10 (the most related with the WRKY guide-genes) fit in the stress responsive group (Table 3) and the SuperSAGE data validate it (Table 4). These results corroborate the involvement of GmWRKY6, GmWRKY46, GmWRKY56, GmWRKY106 and GmWRKY149 genes in plant defense mechanisms against stress. The activation of stress responsive genes triggers a range of molecular processes to overcome the plant limitations brought by biotic and abiotic perturbations. Signaling components, heat shock proteins, grown regulators, reactive oxygen scavengers, stress proteins, gene expression regulators and ion/proton transporters are codified and lead the plant towards stress tolerance.

The plant inability to escape from adverse environmental conditions due to its immobility demands a greater dynamic balance between the regulator elements with each other and with the entire regulatory machinery to achieve an effective response to stress. Considering that the WRKY genes codify transcription factors, it is no coincidence that among the genes coexpressed, those possibly involved in control of gene expression and stress signaling are more numerous. Coexpression analysis revealed that many putative products of the MR < 10 coexpressed genes like Ankyrin protein, Calcium-binding EF-hand protein, Ubiquitin ligase and Kinase proteins can be part of the stress signal transduction route.

A signal transduction starts by the binding of extracellular signal molecules to the membrane proteins like Ankyrin. The Ankyrin proteins are membrane adaptors which mediate protein–protein interactions and are related to defense signaling (Becerra et al., 2004; Bennett and Chen, 2001; Sakamoto et al., 2013; Seong et al., 2007). The signals are then transmitted into the cells through the intracellular signaling molecules, the secondary messengers. Calcium ion (Ca^{2+}) is one of them. The binding of Ca^{2+} is performed by the calcium-binding EF-hand protein, a calcium sensor, which operates in Ca buffering in the cytosol (Tuteja and Mahajan 2007; Trewavas and Malho, 1997; Lewit-Bentley and Réty, 2000). The secondary messengers, in turn, can activate protein modifier enzymes like ubiquitin. Ariadne proteins are ubiquitin protein ligases E3, which have the role of transferring the small protein ubiquitin to the targeted substrate. This post-translational modification regulates a vast array of cellular processes, including stress responses (Berndsen and Wolberger, 2014; Zhang et al., 2014).

The stress signal relayed by the secondary messengers also alters the kinase proteins (Li and Qian, 2003). Unsurprisingly, many genes encoding these proteins are part of the coexpression networks of the WRKY genes analyzed. The relation between kinases and regulatory proteins is well established and reflects the kinases ability in direct leading processes like cell cycle control, plant hormones action and stress responses. The kinases add one or more high-energy phosphate groups to a protein, thereby switching a system from one state to another (Neill and Burnett, 1999; Laurie and Halford, 2001).

Regarding the gene expression control, the MR < 10 coexpression networks bear a noteworthy number of genes known for integrating the gene expression management in stress events (Table 3). Their presumably encoding products are the deadenylase complex subunit CAF1, transposable elements and the transcription factors ATSZF2, ERF4, SAW2, SCL1, NAC102 and WRKYs. The gene expression regulation can occur at different levels. At the RNA stability level, the CAF1 acts in mRNA degradation (Liang et al., 2009; Walley et al., 2010). The RNA decay is a very important mechanism because it provides a rapid gene expression reprogramming, especially in response to environmental signals (Liang et al., 2009). Ma and Bohnert (2007) identified CAF1 as a common stress response gene. At RNA processing level, the activation of transposable elements by stress makes them not only the major source of genetic variability, but also a very fine tuning tool of the genes expression (Wessler, 1996; Grandbastien, 1998; Capy et al., 2000).

At transcriptional level of gene expression control, studies in *Arabidopsis* showed that ERF4, ATSZF2 and SAW2 proteins act as negative regulators of ethylene and ABA sensitivity, salt stress responses and growth, respectively (Yang et al., 2005; Sun et al., 2007; Kumar et al., 2007). Also in *Arabidopsis* studies, Christianson et al. (2009) concluded that ANAC102 primary role is very likely to respond to stress. The function of SCL proteins as a regulatory link between stress and developmental signaling was reviewed by Golldack et al. (2013).

The WRKY transcription factors family is overrepresented in the MR < 10 group (Table 3), supporting the evidence of interaction between WRKY genes as indicated by the W-box abundance in their promoters (Table 2). WRKY33 and WRKY40, both defined as "hubs" by Choura et al. (2015), are among the *Arabidopsis* orthologs of the coexpressed genes. The hubs are considered the maintainers of WRKY web robustness due to many connections that are established from them. Their multifunctional character enables them to mediate key biological pathways such as signal transduction. Additionally, according to Ma and Bohnert (2007), WRKY33 and WRKY40 are generally upregulated by stress.

The WRKY genes interconnection and their intricate transcriptional circuits are also asserted by some overlap observed among the analyzed coexpression networks, mainly between GmWRKY56 and GmWRKY106 (Table 3). However, this does not occur to the MR < 10 genes coexpressed with GmWRKY6. Interestingly, unlike the other guide-genes, the GmWRKY6 codify a protein that does not bear the WRKY family conserved heptameric peptide WRKYGQK. Instead, it contains a variant motif, the WRKYGKK. The requirement of the WRKY canonical motif for a proper DNA-binding was reported by Maeo et al. (2001). In soybean, Zhou et al. (2008) showed that WRKY proteins with the family signature altered, including GmWRKY6, lost the ability to bind the W-box. This observation suggests that they possibly have different binding sites, target genes and consequently distinct coexpression networks (Ciolkowski et al., 2008; Van Verk et al., 2008).

The broad overlap of the GmWRKY56 and GmWRKY106 coexpression network may be explained by the high sequence similarity between them. The close relationship among these genes has been previously confirmed by Bencke-Malato et al. (2014) through phylogenetic analysis. In the same study, it was shown that GmWRKY56 and GmWRKY106 present similar expression patterns in soybean response to *Phakopsora pachyrhizi* infection. Besides, the transcriptional profiles determined in our study show that these genes also have a similar expression behavior in soybean under water deficit. Both genes were induced in roots. Together, these data point that GmWRKY56 and GmWRKY106 might share functions in plant stress defense.

4. Conclusion

In this study, the transcriptional profile of ten soybean WRKY genes in response to water deficit was outlined. The higher expression in tolerant genotype presented by five of them may be understood as they have a more determinative role in this process. The involvement in defense routes against other stresses, the regulation by ABA and the overlap among the coexpression

networks have been revealed through the *cis*-elements and coexpression analyses. Altogether, the obtained results are in agreement with the complex regulatory and functional network of WRKY transcription factors. We believe that our data carry the starting point information to justify the application of these genes in plant genetic engineering to promote drought tolerance or even as phenotyping tools in breeding programs.

Contributions

LPD: Carried out experimental work.

LPD, LAOB and MHBZ: Designed experiments, data discussion and wrote the manuscript.

Acknowledgements

This work was supported by grants from CNPq, CAPES and FAPERGS.

References

Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., Yamaguchi-Shinozaki, K., 2002. Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. Plant Cell Online 15 (1), 63–78.

Abe, H., Yamaguchi-Shinozaki, K., Urao, T., Iwasaki, T., Hosokawa, D., Shinozaki, K., 1997. Role of Arabidopsis MYC and MYB homologs in drought and abscisic acid-regulated gene expression. Plant Cell 9 (10), 1859.

Aoki, K., Ogata, Y., Shibata, D., 2007. Approaches for extracting practical information from gene coexpression networks in plant biology. Plant Cell Physiol. 48 (3), 381–390.

Becerra, C., Jahrmann, T., Puigdomènech, P., Vicient, C.M., 2004. Ankyrin repeat-containing proteins in *Arabidopsis*: characterization of a novel and abundant group of genes coding ankyrin-transmembrane proteins. Gene 340, 111–121.

Bencke-Malato, M., Cabreira, C., Wiebke-Strohm, B., Bücker-Neto, L., Mancini, E., Osorio, M.B., Homrich, Milena, S., Turchetto-Zolet, A., De Carvalho, M., Stolf, R., Weber, R., Westergaard, G., Castagnaro, A.P., Abdelnoor, R.V., Marcelino-Guimarães, F.C., Margis-Pinheiro, M., Bodanese-Zanettini, M., 2014. Genome-wide annotation of the soybean WRKY family and functional characterization of genes involved in response to *Phakopsora pachyrhizi* infection. BMC Plant Biol. 14 (1), 236.

Bennett, V., Chen, L., 2001. Ankyrins and cellular targeting of diverse membrane proteins to physiological sites. Curr. Opin. Cell Biol. 13, 61–67.

Berndsen, C.E., Wolberger, C., 2014. New insights into ubiquitin E3 ligase mechanism. Nat. Struct. Mol. Biol. 21 (4), 301–307.

Berri, S., Abbruscato, P., Faivre-Rampant, O., Brasileiro, A.C., Fumasoni, I., Satoh, K., Kikuchi, S., Mizzi, L., Morandini, P., Pè Mario, Piffanelli, P., 2009. Characterization of WRKY co-regulatory networks in rice and *Arabidopsis*. BMC Plant Biol. 9, 120.

Cabreira, C., Cagliari, A., Bücker-Neto, L., Strohm, B.W., Freitas, L.B., Francismar, C., Guimarães, F.M., Nepomuceno, A.L., Margis-Pinheiro, M.M.A.N., Zanettini, M.H.B., 2013. The lesion simulating disease (LSD) gene family as a variable in soybean response to *Phakopsora pachyrhizi* infection and dehydration. Funct. Integr. Genom. 13, 323–338.

Campalans, A., Messeguer, R., Goday, A., Pagès, M., 1999. Plant responses to drought, from ABA signal transduction events to the action of the induced proteins. Plant Physiol. Biochem. 37 (5), 327–340.

Capy, P., Gasperi, G., Biémont, C., Bazin, C., 2000. Stress and transposable elements: co-evolution or useful parasites? Heredity 85, 101–106.

Chang, W.C., Lee, T.Y., Huang, H.D., Huang, H.Y., Pan, R.L., 2008. PlantPAN: plant promoter analysis navigator, for identifying combinatorial cisregulatory elements with distance constraint in plant gene group. BMC Genom. 9, 561.

Chen, H., Lai, Z., Shi, J., Xiao, Y., Chen, Z., Xu, X., 2010. Roles of *Arabidopsis* WRKY18, WRKY40 and WRKY60 transcription factors in plant responses to abscisic acid and abiotic stress. BMC Plant Biol. 10 (1), 281.

Chen, L., Song, Y., Li, S., Zhang, L., Zou, C., Yu, D., 2012. The role of WRKY transcription factors in plant abiotic stresses. Biochim. Biophys. Acta 1819 (2), 120–128.

Choura, M., Rebaï, A., Masmoudi, K., 2015. Unraveling the WRKY transcription factors network in *Arabidopsis thaliana* by integrative approach. Netw. Biol. 5 (2), 55–61.

Christianson, J.A., Wilson, I.W., Llewellyn, D.J., Dennis, E.S., 2009. The low-oxygen-induced NAC domain transcription factor ANAC102 affects viability of *Arabidopsis* seeds following low-oxygen treatment. Plant Physiol. 149, 1724–1738.

Ciolkowski, I., Wanke, D., Birkenbihl, R.P., Somssich, I.E., 2008. Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. Plant Mol. Biol. 68 (1–2), 81–92.

Cormack, R.S., Eulgem, T., Rushton, P.J., Köchner, P., Hahlbrock, K., Somssich, I.E., 2002. Leucine zipper-containing WRKY proteins widen the spectrum of immediate early elicitor-induced WRKY transcription factors in parsley. Biochim. Biophys. Acta 1576 (1–2), 92–100.

Davies, W.J., Zhang, J., 1991. Roots signals and the regulation of growth and development of plants in drying soil. Annu. Rev. Plant Physiol. 42, 55–76.

Deihimi, T., Niazi, A., Ebrahimi, M., Kajbaf, K., Fanaee, S., Bakhtiarizadeh, M.R., Ebrahimie, E., 2012. Finding the undiscovered roles of genes: an approach using mutual ranking of coexpressed genes and promoter architecture-case study: dual roles of thaumatin like proteins in biotic and abiotic stresses. SpringerPlus 1, 30.

Dong, J., Chen, C., Chen, Z., 2003. Expression profiles of the *Arabidopsis* WRKY gene superfamily during plant defense response. Plant Mol. Biol. 51 (1), 21–37.

Eulgem, T., Rushton, P.J., Robatzek, S., Somssich, I.E., 2000. The WRKY superfamily of plant transcription factors. Trends Plant Sci. 5 (5), 199–206.

Eulgem, T., Rushton, P.J., Schmelzer, E., Hahlbrock, K., Somssich, I.E., 1999. Early nuclear events in plant defence signalling: rapid gene activation by WRKY transcription factors. EMBO J. 18 (17), 4689–4699.

Eulgem, T., Somssich, I.E., 2007. Networks of WRKY transcription factors in defense signaling. Curr. Opin. Plant Biol. 10 (4), 366–371.

Feng, B., Yang, S., Du, H., Hou, X., Zhang, J., Liu, H., Huang, Y., 2012. Molecular characterization and functional analysis of plant WRKY genes. Afr. J. Biotechnol. 11 (72), 13606–13613.

Ficklin, S.P., Feltus, F.A., 2011. Gene Coexpression Network alignment and conservation of gene modules between two grass species: maize and rice. Plant Physiol. 156, 1244–1256.

Golldack, D., Li, C., Mohan, H., Probst, N., 2013. Gibberellins and abscisic acid signal crosstalk: living and developing under unfavorable conditions. Plant Cell Rep. 32, 1007–1016.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., Rokhsar, D.S., 2012. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 40, 1178–1186.

Grandbastien, M.A., 1998. Activation of plant retrotransposons under stress conditions. Trends Plant Sci. 3 (5), 181–187.

Higashi, K., Ishiga, Y., Inagaki, Y., Toyoda, K., Shiraishi, T., Ichinose, Y., 2008. Modulation of defense signal transduction by flagellin-induced WRKY41 transcription factor in *Arabidopsis thaliana*. Mol. Genet. Genom. 279, 303–312.

Jiang, W., Yu, D., 2009. *Arabidopsis* WRKY2 transcription factor mediates seed germination and postgermination arrest of development by abscisic acid. BMC Plant Biol. 9, 96.

Jiang, Y., Liang, G., Yu, D., 2012. Activated expression of WRKY57 confers drought tolerance in *Arabidopsis*. Mol. Plant 5, 1375–1388.

Jian, B., Liu, B., Bi, Y., Hou, 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.

Journot-Catalino, N., Somssich, I.E., Roby, D., Kroj, T., 2006. The transcription factors WRKY11 and WRKY17 act as negative regulators of basal resistance in *Arabidopsis* thaliana. Plant Cell 18, 3289–3302.

Jung, C., Seo, J.S., Han, S.W., Koo, Y.J., Kim, C.H., Song, S.I., Nahm, B.H., Choi, Y.D., Cheong, J.J., 2008. Overexpression of AtMYB44 enhances stomatal closure to confer abiotic stress tolerance in transgenic *Arabidopsis*. Plant Physiol. 146 (2), 623–635.

Koressaar, T., Remm, M., 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23 (10), 1289–1291.

Ku, Y., Yung, Y., Li, M., Wen, C., Liu, X., 2010. Drought stress and tolerance in soybean. In: Board J E (Org.) (Ed.), A Comprehensive Survey of International Soybean Research – Genetics, Physiology, Agronomy and Nitrogen Relationships. InTech, pp. 209–237.

Kulcheski, F.R., de Oliveira, L.F., Molina, L.G., Almerão, M.P., Rodrigues, F.A., Marcolino, J., Barbosa, J.F., Stolf-Moreira, R., Nepomuceno, A.L., Marcelino-Guimarães, F.C., Abdelnoor, R.V., Nascimento, L.C., Carazzolle, M.F., Pereira, G.A., Margis, R., 2011. Identification of novel soybean microRNAs involved in abiotic and biotic stresses. BMC Genom. 12, 307–324.

Kumar, R., Kushalappa, K., Godt, D., Pidkowich, M.S., Pastorelli, S., Hepworth, S.R., Haughn, G.W., 2007. The *Arabidopsis* BEL1-LIKE HOMEODOMAIN proteins SAW1 and SAW2 act redundantly to regulate KNOX expression spatially in leaf margins. Plant Cell 19, 2719–2735.

Laurie, S., Halford, N.G., 2001. The role of protein kinases in the regulation of plant growth and development. Plant Growth Regul. 34, 253–265.

Lewit-Bentley, A., Réty, S., 2000. EF-hand calcium-binding proteins. Curr. Opin. Struct. Biol. 10, 637–643.

Li, G., Qian, H., 2003. Sensitivity and specificity amplification in signal transduction. Cell Biochem. Biophys. 39, 45–59.

Li, J., Brader, G., Kariola, T., Tapio Palva, E., 2006. WRKY70 modulates the selection of signaling pathways in plant defense. Plant J. 46, 477–491.

Liang, W., Li, C., Liu, F., Jiang, H., Li, S., Sun, J., Wu, X., Li, C., 2009. The *Arabidopsis* homologs of CCR4-associated factor 1 show mRNA deadenylation activity and play a role in plant defence responses. Cell Res. 19, 307–316.

Liu, Y., Yin, J., Xiao, M., Mason, A.S., Gao, C., Liu, H., Li, J., Fu, D., 2013. Characterization of structure, divergence and regulation patterns of plant promoters. J. Mol. Biol. Res. 3 (1), 23–36.

Liu, Z., Yan, L., Wu, Z., Mei, C., Lu, K., Yu, Y., Liang, S., Zhang, X., Wang, X., Zhang, D., 2012. Cooperation of three WRKY-domain transcription factors WRKY18, WRKY40, and WRKY60 in repressing two ABA-responsive genes ABI4 and ABI5 in *Arabidopsis*. J. Exp. Bot. 63 (2), 695–709.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 25 (4), 402–408.

Luo, X., Bai, X., Sun, X., Zhu, D., Liu, B., Ji, W., Cai, H., Cao, L., Wu, J., Hu, M., Liu, X., Tang, L., Zhu, Y., 2013. Expression of wild soybean WRKY20 in *Arabidopsis* enhances drought tolerance and regulates ABA signalling. J. Exp. Bot. 64, 2155–2169.

Ma, S., Bohnert, H.J., 2007. Integration of *Arabidopsis thaliana* stress-related transcript profiles, promoter structures, and cell-specific expression. Genome Biol. 8 (4), R49.

Maeo, K., Hayashi, S., Kojima-Suzuki, H., Morikami, A., Nakamura, K., 2001. Role of conserved residues of the WRKY domain in the DNA-binding of tobacco WRKY family proteins. Biosci. Biotechnol. Biochem. 65, 2428–2436.

Masuda, T., Goldsmith, P.D., 2009. World Soybean production: area harvested, yield, and long-term projections. Int. Food Agribus. Manag. Rev. 12 (4), 143–162.

Melotto, M., Underwood, W., Koczan, J., Nomura, K., He, S.Y., 2006. Plant stomata function in innate immunity against bacterial invasion. Cell 126 (5), 969–980.

Mochida, K., Yoshida, T., Sakurai, T., Yamaguchi-Shinozaki, K., Shinozaki, K., Tran, L.S.P., 2009. *In silico* analysis of transcription factor repertoire and prediction of stress responsive transcription factors in soybean. DNA Res. 16 (6), 353–369.

Neill, S.J., Burnett, E.C., 1999. Regulation of gene expression during water deficit stress. Plant Growth Regul. 29, 23–33.

Niu, C.F., Wei, W., Zhou, Q.Y., Tian, A.G., Hao, Y.J., Zhang, W.K., Ma, B., Lin, Q., Zhang, Z.B., Zhang, J.S., Chen, S.Y., 2012. Wheat WRKY genes TaWRKY2 and TaWRKY19 regulate abiotic stress tolerance in transgenic *Arabidopsis* plants. Plant Cell Environ. 35, 1156–1170.

Obayashi, T., Kinoshita, K., 2009. Rank of correlation coefficient as a comparable measure for biological significance of gene coexpression. DNA Res. 16, 249–260.

Oliveira, A.C., Magalhães Jr., A.M., Silveira, C.F., Braga, E.J.B., 2010. Ácido abscísico e o estresse abiótico. In: EMBRAPA Clima Temperado. Documentos, vol. 307, p. 30. ISSN 1516–8840.

Oya, T., Nepomuceno, A.L., Neumaier, N., Farias, J.R.B., Tobita, S., Ito, O., 2004. Drought tolerance characteristics of Brazilian soybean cultivars—evaluation and characterization of drought tolerance of various Brazilian soybean cultivars in the field. Plant Prod. Sci. 7, 129–137.

Qiu, Y.P., Yu, D.Q., 2009. Over-expression of the stress-induced OsWRKY45 enhances disease resistance and drought tolerance in *Arabidopsis*. Environ. Exp. Bot. 65, 35–47.

Ramamoorthy, R., Jiang, S.Y., Kumar, N., Venkatesh, P.N., Ramachandran, S., 2008. A comprehensive transcriptional profiling of the WRKY gene family in rice under various abiotic and phytohormone treatments. Plant Cell Physiol. 49, 865–879.

Ren, X., Chen, Z., Liu, Y., Zhang, H., Zhang, M., Liu, Q., Hong, X., Zhu, J.K., Gong, Z., 2010. ABO3, a WRKY transcription factor, mediates plant responses to abscisic acid and drought tolerance in *Arabidopsis*. Plant J. 63, 417–429.

Rinerson, C.I., Scully, E.D., Palmer, N.A., Donze-Reiner, T., Rabara, R.C., Tripathi, P., Shen, Qingxi J., Sattler, S.E., Rohila, J.S., Sarath, G., Rushton, P.J., 2015. The WRKY transcription factor family and senescence in switchgrass. BMC Genomics 16, 912.

Robatzek, S., Somssich, I.E., 2001. A new member of the *Arabidopsis* WRKY transcription factor family, AtWRKY6, is associated with both senescence and defence related processes. Plant J. 28 (2), 123–133.

Robatzek, S., Somssich, I.E., 2002. Targets of AtWRKY6 regulation during plant senescence and pathogen defense. Genes Dev. 16 (9), 1139–1149.

Rushton, D.L., Tripathi, P., Rabara, R.C., Lin, J., Ringler, P., Boken, A.K., Langum, J., Smidt, L., Boomsma, D.D., Emme, N.J., Chen, X., Finer, J.J., Rushton, P.J., 2012. WRKY transcription factors: key components in abscisic acid signalling. Plant Biotechnol. J. 10 (1), 2–11.

Saito, K., Hirai, M.Y., Yonekura-Sakakibara, K., 2008. Decoding genes with coexpression networks and metabolomics – "majority report by precogs". Trends Plant Sci. 13, 36–43.

Sakamoto, H., Nakagawara, Y., Oguri, S., 2013. The expression of a novel gene encoding an Ankyrin-repeat protein, DRA1, is regulated by drought-responsive alternative splicing. Int. J. Biol. Biomol. Agric. Food Biotechnol. Eng. 7 (12), 791–794.

Schachtman, D.P., Goodger, J.Q.D., 2008. Chemical root to shoot signaling under drought. Trends Plant Sci. 13 (6), 281–287.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song Q,Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.-C., Shinozaki, K., Nguyen, H.T., Wing, R., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C., Jackson, S., 2010. Genome sequence of the palaeopolyploid soybean. Nature 463 (7278), 178–183.

Seong, E.S., Choi, D., Cho, H.S., Lim, C.K., Cho, H.J., Wang, M.H., 2007. Characterization of a stress-responsive ankyrin repeat-containing zinc finger protein of Capsicum annuum (CaKR1). J. Biochem. Mol. Biol. 40 (6), 952–958.

Shang, Y., Yan, L., Liu, Z.Q., Cao, Z., Mei, C., Xin, Q., Wu, F.Q., Wang, X.F., Du, S.Y., Jiang, T., Zhang, X.F., Zhao, R., Sun, H.L., Liu, R., Yu, Y.T., Zhang, D.P., 2010. The Mg-chelatase H subunit of *arabidopsis* antagonizes a group of WRKY transcription repressors to relieve ABA-responsive genes of inhibition. Plant Cell 22, 1909–1935.

Shekhawat, U.K., Ganapathi, T.R., Srinivas, L., 2011. Cloning and characterization of a novel stress-responsive WRKY transcription factor gene (MusaWRKY71) from Musa spp. cv. Karibale Monthan (ABB group) using transformed banana cells. Mol. Biol. Rep. 38, 4023–4035.

Shinozaki, K., Yamaguchi-Shinozaki, K., 2000. Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. Curr. Opin. Plant Biol. 3, 217.

Shinozaki, K., Yamaguchi-Shinozaki, K., 1997. Gene expression and signal transduction in water-stress response. Plant Physiol. 115, 327–334.

Simpson, S.D., Nakashima, K., Narusaka, Y., Seki, M., Shinozaki, K., Yamaguchi-Shinozaki, K., 2003. Two different novel cis-acting elements of erd1, a clpA homologous Arabidopsis gene function in induction by dehydration stress and dark-induced senescence. Plant J. Cell Mol. Biol. 33 (2), 59–70.

Song, H., Wang, P., Hou, L., Zhao, S., Zhao, C., Xia, H., Li, P., 2016. Global analysis of WRKY genes and their response to dehydration and salt stress in soybean. Front. Plant Sci. 7, 1–15.

Stolf, R., et al., 2009. Morpho-anatomical and micromorphometrical evaluations in soybean genotypes during water stress. Braz. Arch. Biol. Technol. 52 (6), 1321–1331.

Stuart, J.M., Segal, E., Koller, D., Kim, S.K., 2003. A gene-coexpression network for global discovery of conserved genetic modules. Science 302, 249–255.

Sun, J., Jiang, H., Xu, Y., Li, H., Wu, X., Xie, Q., Li, C., 2007. The CCCH-type zinc finger proteins AtSZF1 and AtSZF2 regulate salt stress responses in *Arabidopsis*. Plant Cell Physiol. 48 (8), 1148–1158.

Trewavas, A., Malho, R., 1997. Signal perception and transduction: the origin of the phenotype. Plant Cell 9, 1181–1195.

Tuteja, N., Mahajan, S., 2007. Calcium signaling network in plants: an overview. Plant Signal. Behav. 2, 79–85.

Udvardi, M.K., Kakar, K., Wandrey, M., Montanari, O., Murray, J., Andriankaja, A., Zhang, J.-Y., Benedito, V., Hofer, J.M.I., Chueng, F., Town, C.D., 2007. Legume transcription factors: global regulators of plant development and response to the environment. Plant Physiol. 144 (2), 538–549.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G., 2012. Primer3-new capabilities and interfaces. Nucleic Acids Res. 40 (15), e115.

Urao, T., Yamaguchi-Shinozaki, K., Urao, S., Shinozaki, K., 1993. An *Arabidopsis* myb homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence. Plant Cell 5 (11), 1529–1539.

Van Verk, M.C., Pappaioannou, D., Neeleman, L., Bol, J.F., Linthorst, H.J.M., 2008. A Novel WRKY transcription factor is required for induction of PR-1a gene expression by salicylic acid and bacterial elicitors. Plant Physiol. 146, 1983–1995.

Varshney, R.K., Bansal, K.C., Aggarwal, P.K., Datta, S.K., Craufurd, P.Q., 2011. Agricultural biotechnology for crop improvement in a variable climate: hope or hype? Trends Plant Sci. 16 (7), 363–371.

Walley, J.W., Kelley, D.R., Nestorova, G., Hirschberg, D.L., Dehesh, K., 2010. *Arabidopsis* deadenylases AtCAF1a and AtCAF1b play overlapping and distinct roles in mediating environmental stress responses. Plant Physiol. 152, 866–875.

Wang, C., Deng, P., Chen, L., Wang, X., Ma, H., Hu, W., Yao, N., Feng, Y., Chai, R., Yang, G., He, G., 2013. A wheat WRKY transcription factor TaWRKY10 confers tolerance to multiple abiotic stresses in transgenic tobacco. PLoS One 8, e65120.

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

Wang, Z., Zhu, Y., Wang, L.L., Liu, X., Liu, Y.X., Phillips, J., Deng, X., 2009. A WRKY transcription factor participates in dehydration tolerance in Boea hygrometrica by binding to the W-box elements of the galactinol synthase (BhGolS1) promoter. Planta 230, 1155–1166.

Wanke, D., Hahn, A., Kilian, J., Harter, K., Berendzen, K.W., 2010. Regulatory networks: inferring functional relationships through co-expression. QP-PQ: quantum probability and white noise analysis – Vol. 26: 405–424. In: Quantum Bio-informatics III: from Quantum Information to Bio-informatics. World Scientific, 2010; ISBN: 978-981-4304-05-4.

Wei, H., Persson, S., Mehta, T., Srinivasasainagendra, V., Chen, L., Page, G.P., Somerville, C., Loraine, A., 2006. Transcriptional coordination of the metabolic network in *Arabidopsis*. Plant Physiol. 142, 762–774.

Wessler, S.R., 1996. Plant retrotransposons: turned on by stress. Curr. Biol. 6, 959–961.

Wu, X., Shiroto, Y., Kishitani, S., Ito, Y., Toriyama, K., 2009. Enhanced heat and drought tolerance in transgenic rice seedlings overexpressing OsWRKY11 under the control of HSP101 promoter. Plant Cell Rep. 28, 21–30.

Xiong, X., James, V.A., Zhang, H., Altpeter, F., 2010. Constitutive expression of the barley HvWRKY38 transcription factor enhances drought tolerance in turf and forage grass (Paspalum notatum Flugge). Mol. Breed. 25, 419–432.

Yamaguchi-Shinozaki, K., Shinozaki, K., 2005. Organization of cis-acting regulatory elements in osmotic and cold stress responsive promoters. Trends Plant Sci. 10 (2), 88–94.

Yang, Z., Tian, L., Latoszek-Green, M., Brown, D., Wu, K., 2005. *Arabidopsis* ERF4 is a transcriptional repressor capable of modulating ethylene and abscisic acid responses. Plant Mol. Biol. 58, 585–596.

Yin, G., Xu, H., Xiao, S., Qin, Y., Li, Y., Yan, Y., Hu, Y., 2013. The large soybean (*Glycine max*) WRKY TF family expanded by segmental duplication events and subsequent divergent selection among subgroups. BMC Plant Biol. 13, 148.

Yu, J., Zhang, Z., Wei, J., Ling, Y., Xu, W., Su, Z., 2014. SFGD: a comprehensive platform for mining functional information from soybean transcriptome data and its use in identifying acyl-lipid metabolism pathways. BMC Genom. 15 (1), 271.

Yu, D., Chen, C., Chen, Z., 2001. Evidence for an important role of WRKY DNA binding proteins in the regulation of *NPR1* gene expression. Plant Cell 13, 1527.

Zentgraf, U., Laun, T., Miao, Y., 2010. The complex regulation of WRKY53 during leaf

senescence of *Arabidopsis* thaliana. Eur. J. Cell Biol. 89 (2–3), 133–137.
Zhang, X., Wang, N., Chen, P., Gao, M., Liu, J., Wang, Y., Zhao, T., Li, Y., Gai, J., 2014. Overexpression of a soybean ariadne-like ubiquitin ligase gene GmARI1 enhances aluminum tolerance in *Arabidopsis*. PLoS One 9 (11), e111120.
Zheng, Z., Qamar, S.A., Chen, Z., Mengiste, T., 2006. *Arabidopsis* WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. Plant J. 48, 592–605.
Zhou, Q.Y., Tian, A.G., Zou, H.F., Xie, Z.M., Lei, G., Huang, J., Wang, C.M., Wang, H.W., Zhang, J.S., Chen, S.Y., 2008. Soybean WRKY-type transcription factor genes, GmWRKY13, GmWRKY21, and GmWRKY54, confer differential tolerance to abiotic stresses in transgenic *Arabidopsis* plants. Plant Biotechnol. J. 6, 486–550.

CAPÍTULO 2

The characterization of GmWRKY46 gene (Glycine max L. Merril) under drought stress components

---

Manuscrito em redação

1. INTRODUCTION

In the last decades, the rapid progress of biotechnology has enabled hundreds of stress-induced genes to be identified. However, the function of most of them is still unknown. The pursuit of drought-associated genes has been an active area of plant research **(**Singh and Laxmi, 2015). The great potential for yield loss during drought stress has spurred the efforts of breeding programs to develop drought tolerant cultivars. With respect to soybean (*Glycine max* L. Merril), there is still an urgent need to improve drought tolerance. Due to its world-wide importance, extensive losses in soybean yield mean not only economical but also social harm, because many families depend directly or indirectly on the jobs generated by the soybean production, transportation and industrialization (FAO, 2018).

The WRKY proteins constitute a large family of transcriptions factors in plants. They participate in diverse biological functions and are well-known for their roles in plant defense against stresses (Eulgem *et al*., 2000; Eulgem, 2006; Eulgem and Somssich, 2007; Chen *et al*., 2011; Feng *et al*., 2012; Bakshi and Oelmüller, 2014; Jiang *et al.* 2017). In Dias *et al*. (2016), the *GmWRKY46* gene was differentially expressed in the drought tolerant cultivar “EMBRAPA48” versus the drought susceptible cultivar “BR16” during dehydration conditions. *GmWRKY46* showed a higher transcriptional level in the leaves of the tolerant cultivar. Furthermore, the action of *AtWRKY41*, the *GmWRKY46* ortholog in *Arabidopsis thaliana*, was strongly involved in the plant reaction to stress. In previous studies, the expression of *AtWRKY41* is tightly related to stress response pathways as a modulator of the crosstalk between the salicylic acid (SA) and jasmonic acid (JA) pathways (Higashi *et al*., 2008), as an abscisic acid partner in the regulation of seed dormancy (Ding *et al*., 2014), as a control factor of anthocyanin (a known stress metabolite) biosynthesis (Duan *et al*., 2018**;** Chalker-Scott, 1999), or even as a factor that interacts with and is targeted by pathogen effector proteins (Gupta *et al*., 2017; Sarris *et al*., 2015).

But the identification and manipulation of genes for improved drought tolerance requires more information than their expression levels under stress conditions. The characterization of the roles of candidate genes in the complex stress response

pathways is essential. Therefore, when it comes to water deficit stress, this can be challenging since the plant responses to salinity, extreme temperatures and oxidative stresses have significant similarities and cross talk with drought stress (Ingram and Bartels, 1996; Bray, 1997). Such observations illustrate how much the gene networks comprising these response pathways are interconnected.

The analysis of promoter sequences arises as a helpful strategy to accomplish the difficult task of elucidating a gene function. The promoter is the regulatory portion of a gene that contains many *cis*-regulatory elements and is typically located upstream of the transcription start site and thus the coding region. C*is*-elements are DNA motifs to which the transcriptions factors bind. There are many factors that influence this binding, such as the timing of when it happens, other regulatory factors that might participate and their clustering into a complex that will define when, where and how a gene is activated. The decoding of a promoter's cis-elements provides access to the set of instructions that describe the task to be performed by the gene.

Therefore, a deeper comprehension of the role of a gene in the plant perception and reaction to the environment can be reached through the comparison of its promoter sequence between two genotypes that have divergent responses to a stress. The possibility of discovering new *cis*-elements unveiled by the single nucleotide polymorphisms may be useful to clarify the molecular basis of the contrasting responses of cultivars when facing a stress. In addition, there is the advantage of identifying new promoters. The characterization of more promoters of plant genes is necessary to increase the options of regulatory sequences available for the use in genetic engineering in order to ensure finer control of transgene expression.

The study of regulatory sequences in soybean genes has revealed interesting options to add to the toolbox of promoters for plant biotechnology (Gunadi *et al*., 2016). A variety of promoters have been characterized in *Glycine max* (L.) Merrill, including inducible by auxin (Li *et al*., 1994), low temperatures (Chen *et al*., 2009) and cyst nematode (Liu *et al*., 2014), as well as seed-specific promoters (Cho *et al*., 1995; Buenrostro-Nava *et al*., 2006; Caiyin *et al*., 2007) and also constitutive promoters have led to gene expression comparable to or even higher than 35S CaMV promoter (Chiera

et al., 2007; Hernandez-Garcia *et al*., 2010, Zhang *et al*., 2015; Zhao *et al*., 2012). However, the study of promoters of soybean genes promoters responsive to water deficit stress *(*Tripathi *et al.,* 2013*;* Tripathi *et al.,* 2016*;* Conforte *et al.,* 2017*)* is lagging somewhat behind.

In this work, we conducted promoter analysis and gene silencing experiments to assess the *GmWRKY46* gene expression under some drought-related stimuli including abscisic acid (ABA) application and osmotic and ionic imbalances. Using this approach, we expanded the understanding of the role of *GmWRKY46* in the soybean response to water deficit and also its position in other analogous stress response pathways.

2. MATERIALS AND METHODS

2.1 Nicotiana benthamiana experiment

2.1.1 Construction of vectors

The 500bp sequence upstream of the transcription start site (TSS) of the *GmWRKY46* gene was amplified from the genomic DNA of EMBRAPA48 and BR16 soybean cultivars using specific primers (Table 2a, b). The promoter fragment was subcloned into the pENTR/D-TOPO vector (Invitrogen, Carlsbad, CA, USA) to generate entry clones. It was then recombined from pENTR/D-TOPO into pHGWFS7 (Karimi *et al*., 2002) using the Gateway cloning system (Invitrogen, Carlsbad, CA, USA) to make the pHGWFS7:Gm46.5BR and pHGWFS7:Gm46.5EM constructs (Table 2a, b). To make the positive control construct, the CaMV 35S promoter was subjected to the same cloning process (Table 2c). The empty pHGWFS7 vector was used as negative control.

2.1.2 Agroinfiltration-mediated transient transformation

Agroinfiltration experiments were performed on 4–6-week-old *Nicotiana benthamiana* plants cultivated in vermiculite. The constructs were introduced into *Agrobacterium tumefaciens* strain GV3101. The infiltration of *N. benthamiana* leaves by the Agrobacterium cultures was performed as follows. The overnight cultures of *A.*

tumefaciens strains carrying the respective constructs were spun down gently and the pellets were resuspended to a final OD600 of 0.2 in agroinfiltration buffer (10mM $MgCl_2$, 10mM MES – 2-(N-morpholino)ethanesulfonic acid and 200µM acetosyringone). All suspensions were incubated at room temperature for 3 hours prior to infiltration. Leaf infiltration was conducted by pressing a needleless 1-ml disposable syringe to the under-surface of fully expanded leaves and slowly depressing the plunger.

2.1.3 Stress treatments

Plants were transferred to the hydroponic system described by O'Rourke *et al*. (2009) five days before the agroinfiltration to acclimate. Three plants were assessed per treatment, and two leaves were infiltrated per plant. The suspensions corresponding to the constructs pHGWFS7:Gm46.5BR, pHGWFS7:Gm46.5EM, pHGWFS7:35S (Table 2a, b and c), and the pHGWFS7 empty vector were infiltrated into each leaf, totaling four infiltration spots per leaf. One day after the agroinfiltration, the plants were exposed to the following treatments: NaCl 75mM, NaCl 100mM, PEG 6000 5% or PEG 6000 10%. The stress treatments were applied by adding NaCl/PEG 6000 to the full hydroponic nutrient solution. For the control plants, no alterations were made in the nutrient solution. After 24 hours, leaf discs were sampled from the infiltrated areas adjacent to the puncture hole, transferred to *Eppendorf* tubes, frozen in liquid nitrogen, and stored at -80 °C.

2.1.4 GUS activity assay

GUS enzyme activity was measured by the fluorescence of 4-methylumbelliferone (4-MU) produced by GUS cleavage of 4-methylumbelliferyl-β-D-glucuronide (4-MUG). The sampled leaf discs were ground in 150µl GUS extraction buffer (50mM NaHPO4 pH 7.0, 10mM 2-mercaptoethanol, 10mM Na2EDTA, 0.1% sodium lauryl sarcosine, 0.1% Triton X-100). After centrifugation at 4°C, 50 µl of supernatant was mixed with 100 µl of GUS assay solution (2mM 4-MUG in extraction buffer) and 50 µl of GUS extraction buffer. The mixture was incubated at 37°C for 30

min. Then, a 50 µl aliquot was promptly removed and added into 950µl stop buffer (0.2M sodium carbonate). The fluorescence level of 4-MU was determined using a fluorescence spectrophotometer (BioTek Synergy HT, Winooski, VT, USA) at excitation/emission of 365/455 nm. The 4-MU concentration was calculated from a standard curve. Protein concentrations of tissue extracts were measured by the Bradford method (Bradford, 1976).

2.1.5 Statistical analysis

The means of GUS activity assays that were performed in triplicate from six independent leaf samples were statistically analyzed using variance analysis and compared using the Tukey multiple comparison test at the 5% probability level. Error bars shown in figures represent standard deviation (SD). Statistical analysis was performed, using R-project statistical software version 3.4.3 (http://www.r-project.org)

2.2 I*n silico* promoter analysis

Putative cis-elements were searched in the 500-base pairs (bp) sequence upstream of the transcription start site (TSS) of GmWRKY46 gene isolated from BR16 and EMBRAPA48 soybean genotypes. The search was performed using the Plant Cis Acting Regulatory Elements Database – PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

2.3 Soybean experiment

2.3.1 Generation of silenced plants

Virus-induced gene silencing (VIGS)-mediated suppression of GmWRKY46 was performed according to Basic protocols 1 and 2 from Whitham *et al*. (2016). A 284 bp fragment of the *GmWRKY46* open reading frame was amplified using specific forward and reverse primers (Table 2d) which contain, respectively, XhoI and BamHI restriction

sites at their 5´ ends. This fragment was then cloned into pGEM-Teasy vector (Promega, Madison, USA) and recombined into pBPMV-IA-V2 vector using the BamHI and XhoI sites (Zhang *et al*., 2010). The construct was biolistically inoculated into soybean seedlings. The leaves showing viral symptoms were harvested and stored at -20 °C to be used as an inoculum in the next experiments. The rub inoculation was performed according to Support protocol 3 (Whitham *et al*., 2016). Once the plants showed viral symptoms at four weeks after inoculation, the stress treatments were applied. Control plants were inoculated with the empty vector and followed the same stress treatments.

2.3.2 Stress treatments

Three different treatments composed the experiment. They are: 100µL ABA, 100µL SA and water (control). For each treatment 12 plants were sampled. They were divided in 4 pots, named A, B, C and D groups, each one containing 3 plants. In the A and C pots, the *GmWRKY46*-silenced individuals were grown, and in the B and D pots, the control plants were grown. The leaf tissue discs are collected at T0 (control), and T6 hours after the plants were sprayed with ABA and SA. Fluridone, an ABA inhibitor, was sprayed only on the plants in the C and D groups of all treatments three hours before the T6 sampling (Figure 4). Silwet L-77 (0.2 ml/L) was added to spray solutions as a surfactant. All pots were kept at 75% of water content for the duration of the experiment. The samples were frozen in liquid nitrogen and stored at -80°C until RNA extraction.

2.3.3 RNA extraction and real-time qRT-PCR analysis

Total RNA was extracted using Qiagen plant RNeasy kit, treated with Qiagen RNase free DNase (Qiagen, Valencia, CA, U.S.A.) and then reverse-transcribed into cDNA using SuperScript III Reverse Transcriptase according to the manufacturer's protocol (Invitrogen, Carlsbad, CA, USA). Each cDNA sample was diluted 1:10 in sterile ddH2O, and 5µL of this dilution was applied as template for qPCR using iTaq Universal SYBR Green Supermix (Bio-Rad, Hercules, CA, EUA) and specific primers (Table

2e).The qPCR was performed in 384-well plates using the CFX384 Touch Real-Time PCR Detection System (Bio-Rad, Hercules, CA, EUA). The reactions thermal profile was: 95°C for 2 min, 45 cycles of 95°C for 5 s and 60°C for 30 s. Moreover, to analyze the quality of the dissociation curves, the following program was added after the 45 PCR cycles: 95°C for 10 s, followed by a constant increase of the temperature (0.5°C/0.5s) from 65°C to 95°C. Three biological replicates (three plants) were sampled for each plant group/treatment point and each sample was run in technical triplicates. Relative expression levels of *GmWRKY46* were calculated by normalizing with Actin (Glyma.15G050200) and F-box (Glyma.12G051100) genes (Table 2f and g) according to the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). The data are being statistically analyzed.

3. RESULTS AND DISCUSSION

3.1 The 500bp promoter fragments of *GmWRKY46* isolated from the EMBRAPA48 and BR16 genotypes show a statistically significant higher activity in response to NaCl treatment.

Although reported as a dehydration responsive gene in Dias *et al*. (2016), the function of *GmWRKY46* in the drought response is not well-understood. Here, a portion of its regulatory region were identified, cloned, fused with the GUS reporter gene and transiently introduced in *N. benthamiana* in order to evaluate the promoter responsiveness in other stress pathways known to be overlapping with the drought response pathway.

A promoter is generally structured in three regions: core, proximal and distal. The core promoter encompasses the -35 to +35bp stretch from the transcription start site (TSS). This region contains motifs that bind to the machinery required for mRNA transcription, which makes it the minimal sequence needed for precise transcriptional initiation**.** Both proximal and distal promoter regions cover the regulatory elements. The proximal region, roughly set as 250bp upstream of the TSS, comprises many primary *cis*-elements while the distal portion, also comprises enhancer, silencer and insulator

sequences (Butler and Kadonaga, 2002; Liu and Stewart 2016; Porto *et al*., 2014). In our analysis, the activity of 500 bp promoter fragments of *GmWRKY46* isolated from EMBRAPA48 and BR16 were assessed under PEG and NaCl treatments, compounds broadly known as agents of osmotic and ionic stress, respectively (Michel and Kaufmann, 1973; Shavrukov, 2013).

Osmotic stress is the first effect imposed by the salt and drought stress, leading the plant to elicit a quite similar reaction to both adverse conditions. The plant attempts to overcome the differences in osmotic pressure between extracellular and intracellular spaces by the water transferring from the cytoplasm to the apoplast resulting in cellular osmotic dehydration (Bartels and Sunkar, 2005; Shavrukov, 2013). With respect to salinity, there is a second effect, the ionic stress. As the salt stress process follows, the ionic stress is caused by the imbalance in cellular ion homeostasis due to excessive uptake and consequent ion accumulation (Uddin *et al*., 2016; Gaspar *et al*., 2002).

The promoter activity comparison among treatments showed that the 500 bp *GmWRKY46* promoter fragment from both the drought tolerant (EMBRAPA48) and the drought susceptible (BR16) genotypes were significantly induced by the 100 mM NaCl treatments (Figure 1). Altered expression of GUS was not observed in either the 5 mM or 10 mM PEG treatment, indicating that neither promoter was induced by osmotic stress. These results indicated an association between this 500 bp promoter region and transcription factors that mediate plant responses to salt stress. The response to salt stress and lack of response to osmotic stress suggests that the 500 bp *GmWRKY46* promoter fragment from both cultivars is responsive to the second salt stress stage.

3.2 Cis-elements related to light and plant hormones responses compose the *GmWRKY46* gene promoter

In the promoter region, the transcription factors interact with short sequences called *cis*-elements. A *cis*-element is the DNA binding site of a specific transcription factor family. Beyond the presence of these motifs, their frequency and arrangement inside the regulatory sequence determine what proteins will be recruited to combinatorially direct the spatial and temporal patterns of gene expression (Priest *et al*., 2009). Therefore, the study of *cis*-elements is crucial to gain a detailed understanding of the mechanisms by which genes are induced or repressed in response to external stimuli (Lemon and Tjian, 2000).

A promoter *in silico* analysis was performed on the putative 500 bp sequence upstream the TSS of *GmWRKY46* in order to identify known cis-elements which are present. The analysis revealed three single nucleotide polymorphisms (SNPs) between the *GmWRKY46* promoter sequences from the BR16 and EMBRAPA48 genotypes. These SNPs occur at -130, -289 and -336 upstream of the TSS. In the promoter sequence of the BR16 genotype, the SNPs at the -130 and -289 positions disrupt the “Unnamed__4” and a “CAAT-box” elements, respectively (Figure 3).

Among the *cis*-elements identified, the extensive presence of ones related to responses to light are noteworthy. The motifs involved in the plant stress responses are also well represented. They act as binding sites to proteins responsive to hormones known for signaling during plant stress, such as methyl jasmonate (MeJA), ethylene, salicylic acid and also gibberellin.

Among the many roles that MeJA, ethylene, salicylic acid and gibberellin play in the plant stress signaling, reviewed respectively by Ahmad *et al*. (2016), Gamalero and Glick (2012), Yuan and Lin (2008) and Colebrook *et al*.(2014), is the increase of salinity tolerance in soybean. The results reported by Yoon *et al*. (2009), Ma *et al*. (2012), Jaiswal *et a*l. (2014) and Hamayun *et al.* (2010) showed that the stress signaling cascade triggered by these hormones is associated with the ability of soybean plants to overcome salinity conditions.

Others *cis*-elements including TATA-Box, a core promoter binding site for the transcription initiation machinery, MYB protein binding site and elements related to endosperm and meristem expression were also found in the *GmWRY46* promoter fragment assessed (Table 1).

REFERENCES

Ahmad, P., Rasool, S., Gul, A., Sheikh, S. A., Akram, N. A., Ashraf, M., Kazi, A. M., Gucel, S. (2016) Jasmonates: Multifunctional Roles in Stress Tolerance. Frontiers in Plant Science. 7, 813

Bakshi, M., Oelmüller, R. (2014) WRKY transcription factors jack of many trades in plants. Plant Signaling and Behavior. Plant Signaling and Behavior. 9, 1–18

Bartels, D., Sunkar, R. (2005) Drought and salt tolerance in plants. Critical Reviews in Plant Sciences. 24(1), 23–58

Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry. 72(1–2), 248–254

Bray, E. A. (1997) Plant Response to Water Deficit Stress. Trends in Plant Science, 2(2), 48–54

Buenrostro-Nava, M. T., Ling, P. P., Finer, J. J. (2006) Comparative analysis of 35S and lectin promoters in transgenic soybean tissue using an automated image acquisition system and image analysis. Plant Cell Reports. 25(9), 920–926

Butler, J. E. F., Kadonaga, J. T. (2002) The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes and Development. 16(20), 2583–2592

Caiyin, Q., Li, M., Wei, D., Cai, Y., Xing, L. (2007) Isolation and sequencing analysis on the seed-specific promoter from soybean. Frontiers of Agriculture in China. 1(1), 17–23

Chalker-Scott, L. (1999) Environmental significance of anthocyanins in plant stress responses. Photochemistry and Photobiology. 70(1), 1–9

Chen, M., Xu, Z., Xia, L., Li, L., Cheng, X., Dong, J., Wang, Q., Ma, Y. (2009) Cold-induced modulation and functional analyses of the DRE-binding transcription factor gene, GmDREB3, in soybean (*Glycine max L*.). Journal of Experimental Botany. 60(1), 121–35

Chen, L., Song, Y., Li, S., Zhang, L., Zou, C., Yu, D. (2011) The role of WRKY transcription factors in plant abiotic stresses. Biochimica et Biophysica Acta. 1819(2), 120–8

Chiera, J. M., Bouchard, R. A., Dorsey, S. L., Park, E., Buenrostro-Nava, M. T., Ling, P. P., Finer, J. J. (2007) Isolation of two highly active soybean (Glycine max (L.) Merr.) promoters and their characterization using a new automated image collection and analysis system. Plant Cell Reports. 26(9), 1501–1509

Cho, M., Widhohn, J. M., Vodkin, L. O. (1995) Cassettes for seed-specific expression tested in transformed embryogenic cultures of soybean. Plant Molecular Biology Reporter. 13(3), 255–269

Colebrook, E. H., Thomas, S. G., Phillips, A. L., Hedden, P. (2014) The role of gibberellin signalling in plant responses to abiotic stress. Journal of Experimental Biology. 217(1), 67–75

Conforte, A. J., Guimarães-Dias, F., Neves-Borges, A. C., Bencke-Malato, M., Felix-Whipps, D., Alves-Ferreira, M. (2017) Isolation and characterization of a promoter responsive to salt, osmotic and dehydration stresses in soybean. Genetics and Molecular Biology. 40(1), 226–237

Dias, L. P., Oliveira-Busatto, L. A., Bodanese-Zanettini, M. H. (2016) The differential expression of soybean [Glycine max (L.) Merrill] WRKY genes in response to water deficit. Plant Physiology and Biochemistry, 107, 288–300.

Ding, Z. J., Yan, J. Y., Li, G. X., Wu, Z. C., Zhang, S. Q., Zheng, S. J. (2014) WRKY41 controls *Arabidopsis* seed dormancy via direct regulation of ABI3 transcript levels not downstream of ABA. Plant Journal. 79, 810–823

Duan, S., Wang, J., Gao, C., Jin, C., Li, D., Peng, D., Du, G., Li, Y., Chen, M. (2018) Functional characterization of a heterologously expressed Brassica napus WRKY41-1 transcription factor in regulating anthocyanin biosynthesis in *Arabidopsis* thaliana. Plant Science. 268, 47-53

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

Eulgem, T. (2006) Dissecting the WRKY web of plant defense regulators. PLoS Pathogens. 2(11), 1028–1030

Eulgem, T., Somssich, I. E. (2007) Networks of WRKY transcription factors in defense signaling. Current Opinion in Plant Biology 10(4), 366–71

Feng, B. (2012) Molecular characterization and functional analysis of plant WRKY genes. African Journal of Biotechnology. 11(72), 13606–13613

Food and Agriculture Organization of the United Nations - FAO (2018) The impact of disasters and crises on agriculture and food security.

Gamalero, E., Glick, B., R. Ethylene and Abiotic Stress Tolerance in Plants. In: Ahmad, P., Prasad, M. N. V. Environmental adaptations and stress tolerance of plants in the era of climate change. Springer Science, 2012. Chapter 18, 1–515

Gaspar, T., Franck, T., Bisbis, B., Kevers, C. (2002) Concepts in plant stress physiology. Application to plant tissue cultures. Plant Growth Regulation. 37, 263–285

Gunadi, A., Rushton, P. J., McHale, L. K., Gutek, A. H., and Finer, J. J. (2016) Characterization of 40 soybean (Glycine max) promoters, isolated from across 5 thematic gene groups. Plant Cell, Tissue and Organ Culture. 127(1), 145–160

Gupta, S., Bhar, A., Chatterjee, M., Ghosh, A., Das, S., Gupta, V. (2017) Transcriptomic dissection reveals wide spread differential expression in chickpea during early time points of Fusarium oxysporum f. sp. ciceri. PLoS ONE. 12(5), e0178164

Hamayun, M., Khan, S. A., Khan, A. L., Shin, J. H., Ahmad, B., Shin, D. H., Lee, I. J. (2010) Exogenous gibberellic acid reprograms soybean to higher growth and salt stress tolerance. Journal of Agricultural and Food Chemistry. 58(12), 7226–7232

Hernandez-Garcia, C. M., Bouchard, R. A., Rushton, P. J., Jones, M. L., Chen, X., Timko, M. P., Finer, J. J. (2010) High level transgenic expression of soybean (*Glycine max*) GmERF and Gmubi gene promoters isolated by a novel promoter analysis pipeline. BMC Plant Biology. 10(1), 237

Higashi, K., Ishiga, Y., Inagaki, Y., Toyoda, K., Shiraishi, T., Ichinose, Y. (2008) Modulation of defense signal transduction by flagellin-induced WRKY41 transcription factor in *Arabidopsis* thaliana. Molecular Genetics and Genomics. 279, 303–312

Ingram, J., Bartels, D. (1996) The Molecular Basis of Dehydration Tolerance in Plants. Annual Review of Plant Physiology and Plant Molecular Biology. *47*, 377–403

Jaiswal, A., Pandurangam, V. and Sharma, S.K (2014) Effect of Salicylic Acid in Soybean (Glycine Max L. Meril).The Bioscan Journal 9(2), 671–676

Jiang, J., Ma, S., Ye, N., Jiang, M., Cao, J., Zhang, J. (2017) WRKY transcription factors in plant responses to stresses. Journal of Integrative Plant Biology. 59(2), 86–101

Karimi, M., Inze, Depicker, A. (2002) Gateway vectors for Agrobacterium-mediated plant transformation. Trends in Plant Science. 7, 193–195

Lemon, B. and Tjian, R. (2000) Orchestrated response: A symphony of transcription factors for gene control. Genes and Development. 14(20), 2551–2569.

Li, Y., Liu, Z. B., Shi, X., Hagen, G., Guilfoyle, T. J. (1994) An auxin-inducible element in soybean SAUR promoters. Plant Physiology. 106(1), 37–43

Liu, W., Stewart, C. N. (2016) Plant synthetic promoters and transcription factors. Current Opinion in Biotechnology. 37, 36–44

Liu, W., Mazarei, M., Peng, Y., Fethe, M. H., Rudis, M. R., Lin, J, Millwood, R.J., Arelli, P.R., Stewart, C. N. (2014) Computational discovery of soybean promoter *cis*-regulatory elements for the construction of soybean cyst nematode-inducible synthetic promoters. Plant Biotechnology Journal. 12(8), 1015–1026

Livak, K. J., Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 25(4), 402-408

Ma, H., Song, L., Shu, Y., Wang, S., Niu, J., Wang, Z., Yu, T., Gu, W., Ma, H. (2012) Comparative proteomic analysis of seedling leaves of different salt tolerant soybean genotypes. Journal of Proteomics. 75(5), 1529–1546

Michel, B. E., Kaufmann, M. R. (1973) The Osmotic Potential of Polyethylene Glycol 6000. Plant Physiology. *51*(5), 914–916

Porto, M. S., Pinheiro, M. P. N., Batista, V. G. L., Dos Santos, R. C., De Albuquerque Melo Filho, P., De Lima, L. M. (2014) Plant promoters: An approach of structure and function. Molecular Biotechnology. 56, 38–49.

Priest, H. D., Filichkin, S. A., and Mockler, T. C. (2009) *Cis*-Regulatory elements in plant cell signaling. Current Opinion in Plant Biology. 12(5), 643–649

O'Rourke, J. A., Nelson, R. T., Grant, D., Schmutz, J., Grimwood, J., Cannon, S., Vance, C. P, Graham, M. A, Shoemaker, R. C. (2009) Integrating microarray analysis and the soybean genome to understand the soybeans iron deficiency response. BMC Genomics. 10, 1–17

Sarris, P. F., Duxbury, Z., Huh, S. U., Ma, Y., Segonzac, C., Sklenar, J., Sklenar, Jan, Derbyshire, P., Cevik, V., Rallapalli, G., Saucet,, B., Wirthmueller, L., Menke, F. L.H., Sohn, K.H., Jones, J. D. G. (2015) A plant immune receptor detects pathogen effectors that target WRKY transcription factors. Cell. 161, 1089–1100

Shavrukov, Y. (2013) Salt stress or salt shock: which genes are we studying ?. Journal of Experimental Botany. 64: 1, 119–127

Singh, D., Laxmi, A. (2015). Transcriptional regulation of drought response: a tortuous network of transcriptional factors. Frontiers in Plant Science. 6, 1–11

The R Project for statistical computing [http://www.r-project.org]

Tripathi, P., Rabara, R. C., Lin, J., and Rushton, P. J. (2013) GmWRKY53, a water-and salt-inducible soybean gene for rapid dissection of regulatory elements in BY-2 cell culture. Plant Signaling and Behavior. 8(5), 1–6

Tripathi, P., Rabara, R. C., Reese, R. N., Miller, M. A., Rohila, J. S., Subramanian, S., Shen, Q., Morandi, D., Bücking, H., Shulaev, V., Rushton, P. J. (2016) A toolbox of genes, proteins, metabolites and promoters for improving drought tolerance in soybean includes the metabolite coumestrol and stomatal development genes. BMC Genomics. 17(1), 1–22

Uddin, M. N., Hossain, M. A., Burritt, D. J. Salinity and drought stress: Similarities and differences in oxidative responses and cellular redox regulation. In: Water Stress and Crop Plants: A Sustainable Approach. John Wiley & Sons Ltd, 2016. Volume 1. Chapter 7, 86–101

Whitham, S. A., Lincoln, L. M., Chowda-Reddy, R. V., Dittman, J. D., O'Rourke, J. A., Graham, M. A. (2016) Virus-Induced Gene Silencing and Transient Gene Expression in Soybean (Glycine max) Using Bean Pod Mottle Virus Infectious Clones. Current Protocols in Plant Biology. 1, 263–283

Yoon, J. Y., Hamayun, M., Lee, S. K. and Lee, I. J. (2009) Methyl jasmonate alleviated salinity stress in soybean. Journal of Crop Science and Biotechnology. 12(2), 63–68

Yuan, S., Lin, H. (2008) Role of Salicylic Acid in Plant Abiotic Stress. Zeitschrift für Naturforschung C. 63c, 313-320

Zhang, C., Bradshaw, J. D., Whitham, S. A., Hill, J. H. (2010) The Development of an Efficient Multipurpose Bean Pod Mottle Virus Viral Vector Set for Foreign Gene Expression and RNA Silencing. Plant Physiology 153(1), 52–65

Zhang, N., McHale, L. K., Finer, J. J. (2015) Isolation and characterization of "GmScream" promoters that regulate highly expressing soybean (Glycine max Merr.) genes. Plant Science. 241, 189–198

Zhao, Y., Shao, S., Li, X., Zhai, Y., Zhang, Q., Qian, D., Wang, Q. (2012) Isolation and Activity Analysis of a Seed-Abundant soyAP1 Gene Promoter from Soybean. Plant Molecular Biology Reporter. 30(6), 1400–1407

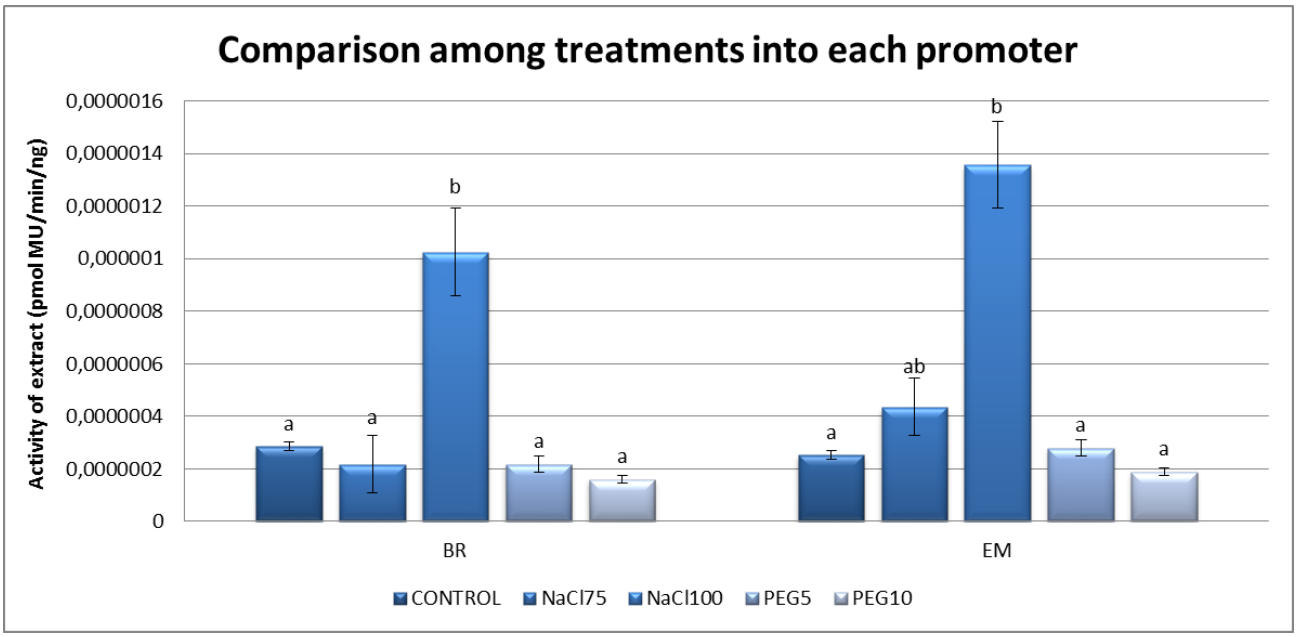

Fig 1. The GUS activity directed by a 500bp promoter fragment of *GmWRKY46* isolated from EMBRAPA48 and BR16 soybean genotypes were measured from leaves of *Nicotiana benthamiana* plants exposed to NaCl 75mM, NaCl 100mM, PEG 5% or PEG10% and water (control) treatments. EM: EMBRAPA48. BR: BR16. The values are means of three biological replicates. Means labeled with the same letter do not differ significantly (Tukey test, $p < 0.05$) between treatments. MU, Methylumbelliferyl. Error bars represent standard errors.

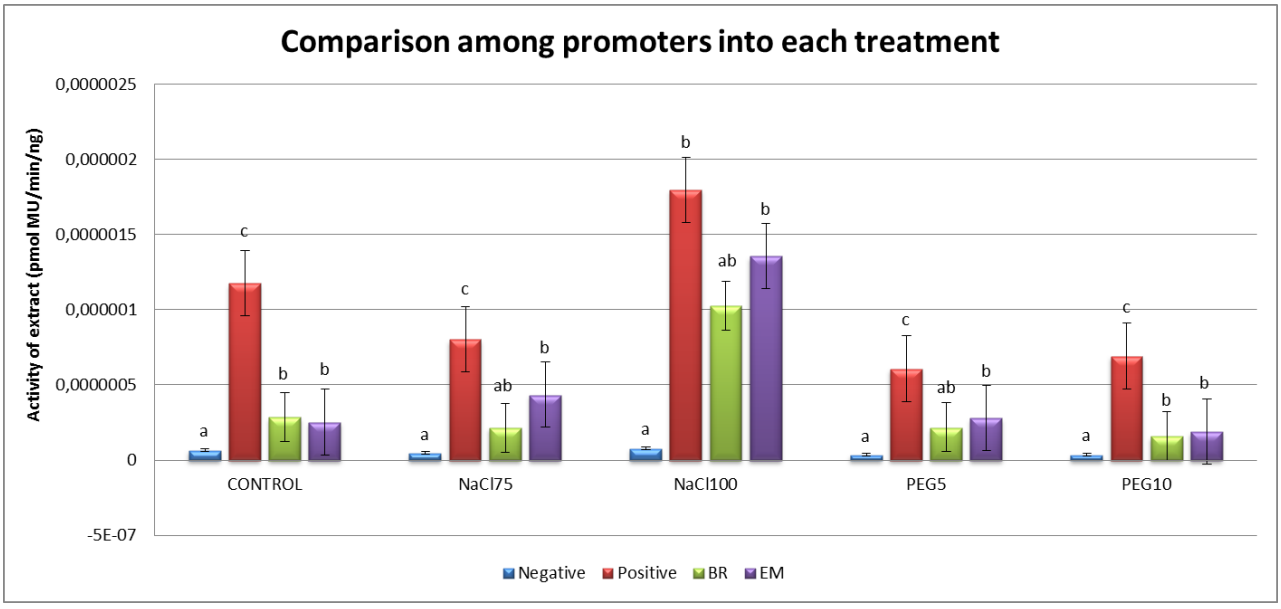

Fig 2. The GUS activity directed by the following sequences: 35S CaMV promoter (positive control) and the 500bp promoter fragments of *GmWRKY46* isolated from EMBRAPA48 and BR16 soybean genotypes were measured from leaves of *Nicotiana benthamiana* plants exposed to NaCl 75mM, NaCl 100mM, PEG 5% or PEG10% and water (control) treatments. The empty pHGWFS7 vector was used to agroinfiltrate the "negative control" leaves. . EM: EMBRAPA48. BR: BR16. The values are means of three biological replicates. Means labeled with the same letter do not differ significantly (Tukey test, $p < 0.05$) between treatments. MU, Methylumbelliferyl. Error bars represent standard errors.

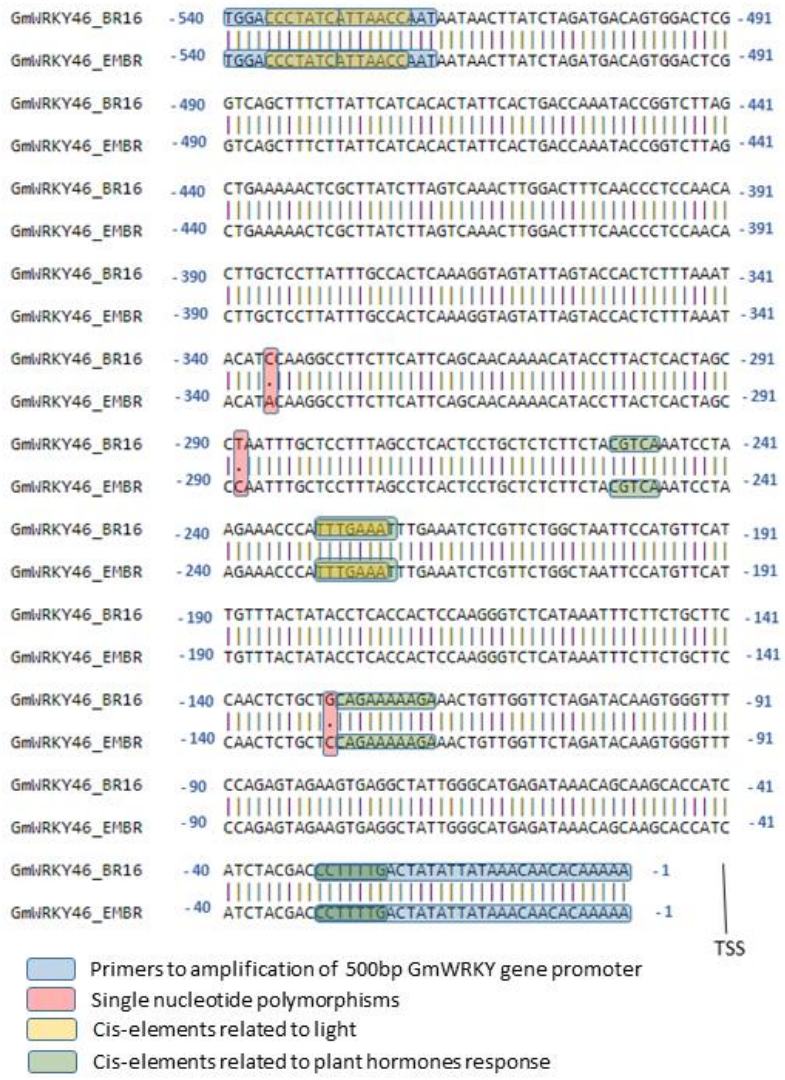

Fig 3. Cis- elements and single nucleotide polymorphisms between the 500bp promoter sequences isolated from the BR16 and EMBRAPA48 soybean genotypes

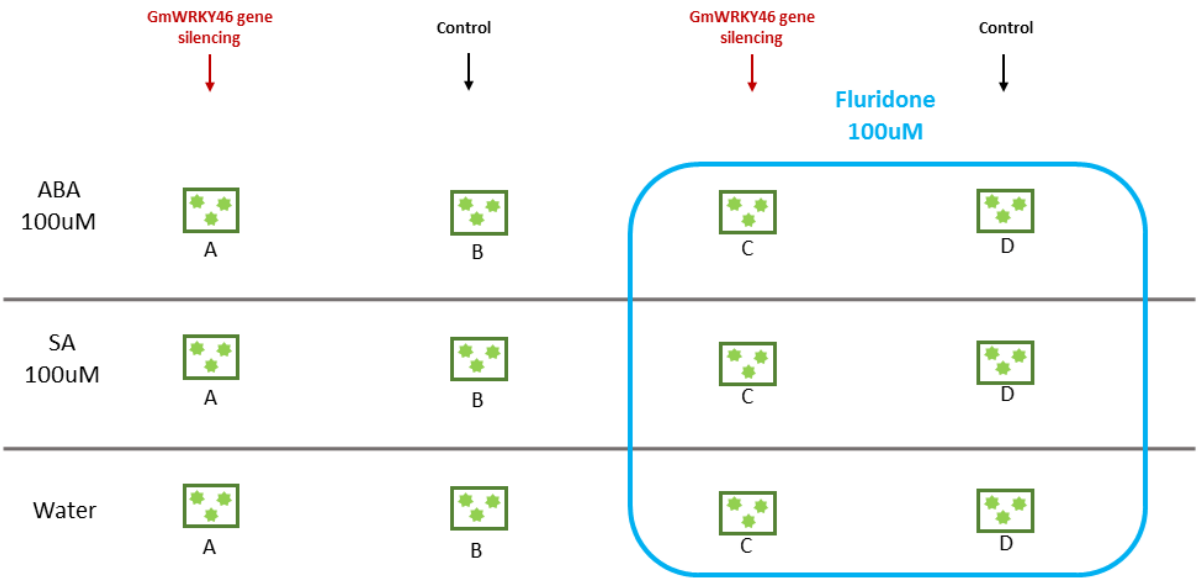

Fig 4 Design of the soybean BPMV system experiment

Table 1. Number of copies and biological function of the putative cis-elements identified in 500bp promoter fragment of GmWRKY46 gene from BR16 and EMBRAPA48 soybean genotypes

Cis-element	Consensus sequence	Function	BR16	EMBRAPA48
Box I	TTTCAAA	Light responsive element	1	1
GATA-motif	GATAGGG	Light responsive element	1	1
GT1-motif	GGTTAA	Light responsive element	1	1
I-box	GATAGGG	Light responsive element	1	1
CGTCA-motif	CGTCA	Methyl jasmonate responsive element	1	1
TGACG-motif	TGACG	Methyl jasmonate responsive element	1	1
P-box	CCTTTTG	Gibberellin responsive element	1	1
Skn-1_motif	GTCAT	Related to endosperm expression	1	1
ERE	ATTTCAAA	Ethylene responsive element	1	1
TCA-element	CAGAAAAGGA	Salicylic acid responsive element	1	1
MBS	CGGTCA	MYB protein binding site	1	1
CAT-box	GCCACT	Related to meristem expression	1	1
CAAT-box	CAAT/CAAAT/CCAAT	Related to enhancer promoter regions	9	10
Unnamed__4	CTCC	.....	5	6

Table 2. Primers sequences and constructs purposes

	Constructs / Primers Names	Primers sequence (5'→ 3')	Fragment amplified	Constructs Purpose
a	pHGWFS7: Gm46.5 BR	F: **CACC**TGGACCCTATCATTAACCAAT	500bp sequence upstream TSS of GmWRKY46 gene isolated from BR16 cultivar	Promoter activity analysis
		R: TTTTTGTGTTGTTTATAATATAGTCAAAAGG		Promoter fragment fused to gfp::gus
b	pHGWFS7: Gm46.5 EM	F: **CACC**TGGACCCTATCATTAACCAAT	500bp sequence upstream TSS of GmWRKY46 gene isolated from EMBRAPA46 cultivar	Promoter activity analysis
		R: TTTTTGTGTTGTTTATAATATAGTCAAAAGG		Promoter fragment fused to gfp::gus
c	pHGWFS7: 35S	F: **CACC**ACTAGAGCCAAGCTGAT	35S CaMV promoter	Promoter activity analysis
		R: ACATTACAATTTACTATTCTAGTCGA		Promoter fragment fused to gfp::gus
d	pBPMV:GmWRKY46	F: CCGCTCGAGAGGTAGCAAGGAAGTTGAAGGC	ORF fragment of GmWRKY46 gene	GmWRKY46 gene silencing
		R: CGCGGATCCGGGCGTTGCCTTTCTTTTCTT		
e	Gm46-qPCR	F: CCCAAATGGATGGATCATGTA	ORF fragment of GmWRKY46 gene	Detection of GmWRKY46 gene expression via qPCR
		R: CTCTGCACTTGCTTTGTTGC		
f	GmAct-qPCR	F: GAGCTATGAATTGCCTGATGG	ORF fagment of soybean actin gene (Glyma.15G050200)	Detection of actin gene expression via qPCR
		R: CGTTTCATGAATTCCAGTAGC		
g	GmFbox-qPCR	F: GAAAGCAGAAAGATGGGGTTGG	ORF fragment of soybean f-box gene (Glyma.12G051100)	Detection of f-box gene expression via qPCR
		R: CACACACGCCACTCTCGCAA		

CONCLUSÕES E PERSPECTIVAS

A importância da identificação de genes responsivos ao déficit hídrico assim como a caracterização de seus promotores foi destacada ao longo de toda tese. O perfil transcricional de dez genes WRKY foi determinado em duas cultivares de soja submetidas a desidratação. Para cinco deles, a expressão no genótipo EMBRAPA48, tolerante à seca, foi significativamente maior do que no genótipo BR16, suscetível à seca (Vide Capítulo 1). A publicação que descreve esses dados foi incluída na lista do "*Mark Plant Institute for Plant Breeding Research*" como uma das referências relevantes sobre WRKY genes do ano de 2016 (http://www.mpipz.mpg.de/20985/WRKY_References).

A caracterização da região promotora do gene *GmWRKY46* foi iniciada e os resultados iniciais apontam um perfil responsivo à salinidade (Vide Capítulo 2). Na análise *in silico* de *cis*-elementos, a marcante presença de motivos relacionados à luz e à resposta ao MeJa abrem o caminho para novas abordagens de investigação. A busca de um maior entendimento sobre a influência destes fatores/sinalizadores na atividade da região promotora do gene *GmWRKY46* surge como mais uma meta a ser alcançada nos próximos passos desta pesquisa.

Além disso, ferramentas indispensáveis para a continuidade desse trabalho, como vetores próprios para a avaliação da atividade dos promotores gênicos e linhagens de *Arabidopsis thaliana* carregando construções "Promotor: Gene repórter", foram obtidas , checadas e estão prontas para serem aplicadas em futuros experimentos (Vide Apêndice). É importante salientar que a utilidade desses vetores e linhagens não se limita apenas a essa pesquisa. Muitos alunos podem ser beneficiados utilizando-os em estudos que contemplem diferentes objetivos, como a análise da atividade dos promotores WRKY sob estresse biótico, metais pesados, nodulação, deficiência nutricional e muitos outros.

APÊNDICE

Ferramentas biotecnológicas desenvolvidas

1. Generation of transgenic *Arabidopsis* lines

The *GmWRKY46*, *GmWRKY56* and *GmWRKY106* putative promoter fragments sequences were amplified by PCR from EMBRAPA48 and BR16 genomic DNA with specific primers (Table 1a to h). The obtained sequences were introduced into pENTR/D-TOPO entry vector (Invitrogen, Carlsbad, CA, USA). The LR Clonase reaction combining the promoter fragment in the entry vector and destination vector pHGWFS7 (Karimi *et al*., 2002) (Figure 1 a) was used to produce the following constructs: pHGWFS7: Gm46.5BR, pHGWFS7: Gm46.5EM, pHGWFS7: Gm106.5BR, pHGWFS7: Gm106.5EM, pHGWFS7: Gm56.5EM, pHGWFS7: Gm106.1EM, pHGWFS7: Gm56.1EM (Table 1a to g). The 35S CaMV promoter was subjected to the same cloning process described in order to obtain the positive control vector pHGWFS7: 35S (Table 1h). All the constructs were checking by sequencing. Stable *Arabidopsis thaliana* Col-0 transgenic lines were generated by *Agrobacterium tumefaciens* using the floral dip method (Clough and Bent, 1998). The seeds from the T2 generation were harvested and the homozygous plants will be screened through hygromycin selection. The lines selected will be used in future experiments to analyze the activity of WRKY promoters under stress conditions.

2. Construction of vectors to promoter analysis

Fragments of the sequence upstream the TSS of *GmWRKY46*, *GmWRKY56* and *GmWRKY106* were amplified by PCR from EMBRAPA48 and BR16 genomic DNA using specific primers (Table 1i to p). The obtained sequences were subcloned into pENTR/D-TOPO entry vector (Invitrogen, Carlsbad, CA, USA) and then transferred into pGWB504 (Nakagawa *et al*., 2007) (Figure 1 b) by gateway system. These are the obtained constructs: pGWB504:Gm46.5BR, pGWB504:Gm46.5EM, pGWB504:Gm106.5BR, pGWB504:Gm106.5EM, pGWB504:Gm106.1BR, pGWB504: Gm106.1EM, pGWB504:Gm56.1EM, pGWB504:35S (Table 1i to p). The main purpose and expectation in the use of these constructs is the possibility of evaluate the activity of promoters through GFP fluorescence, making the enzymatic or chemical assays unnecessary to the analysis and thus simplifying analyses.

3. *Bean pod mottle virus* (BPMV)*:* a tool for functional genomics studies

BPMV was engineered as a plant virus vector to induce gene silencing or express foreign genes without losing its infection capacity (Whitham *et al*. 2016). Ironically, the virus-induced gene silencing (VIGS) happens by the own plant defense system against virus called post-transcriptional gene silencing (PTGS). In the PTGS process, sequence fragments of the virus genome are identify and degraded to prevent the infection or promote the plant recovery. So, by introducing a transgene into a virus vector and then use it to inoculate the vegetal, the VIGS technique causes the plant to identify its own transcripts as viral RNA targeting them for degradation and resulting in a silencing phenotype for the gene of interest (Vaucheret *et al*., 2001 ; Guo *et al*., 2000; Todd *et al*., 2010).

On account of being mediated by a fully functional virus, this plant transformation approach has some advantages over the stable system (Dawson, 2011; Gleba *et al*., 2013):

- The high amplification level of heterologous gene and the consequent rapid accumulation of its product.

- The possibility of inoculate mature plants, overcoming in this way the risk of toxic effects on regeneration and development.

- Reduced cost

- Fast results

- Useful for plants that are difficult to transform, such as soybean.

The Bean pod mottle virus (BPMV) is the viral pathogen that causes the bean pod mottle disease in legume species. BPMV has a bipartite positive-strand RNA genome consisting of RNA-1 (approximately 6.0 kb) and RNA-2 (approximately 3.6 kb). The RNA-1 carries the pathogenicity component by coding the proteins required for replication (Pflieger *et al*., 2014) (Figure 2 a). Whereas RNA-2 codes for a putative cell-to-cell movement protein (MP) and the two coat proteins (L-CP and S- CP) (Giesler *et al*., 2002) (Figure 2 b). Some modifications were made in RNA-2 sequence in order to

make it suitable for application in genetic studies of common bean (*Phaseolus vulgaris*), soybean (*Glycine Max*) and other legumes (Zhang *et al*., 2010). The nature of these modifications determine the vector purpose. If foreign sequences were inserted into the RNA2-encoded polyprotein ORF, specifically between the MP and the L-CP coding regions (Figure 2 c), the BPMV vector will act as an agent of heterologous proteins expression. However, the insertion of short sequences after the RNA-2 ORF stop codon will induce the gene silencing (Figure 2 d) (Zhang and Ghabrial, 2006).

3.1 Construction of BPMV clone for *GmWRKY106* silencing via VIGS

The pBPMV:*GmWRKY106* construct is designed to mediate the suppression of *GmWRKY106* by VIGS. It was obtained by following the instructions of Basic protocol 1 from Whitham *et al*. (2016). Briefly, a 248bp fragment of GmWRKY106 open reading frame was amplified using specific forward and reverse primers (Table 1q) which contain, respectively, XhoI and BamHI restriction sites in their 5´ ends. This fragment was then cloned into pGEM-Teasy vector (Promega, Madison, USA) and recombined into the pBPMV-IA-V2 vector (Zhang *et al*., 2010) using the BamHI and XhoI sites to generate the construct pBPMV:GmWRKY106 (Table 1q). The obtained construct will be used in future experiments to assess the effect of GmWRKY106 gene silencing in the soybean phenotype under stress conditions.

3.2 Construction of BPMV clones for GmWRKY56 transcripts expression

The GmWRKY56 gene has four alternative transcripts. The sequences of three of them were downloaded from Phytozome (http:// www.phytozome.net/). Then, modifications in the DNA sequence of these transcripts were made in order to avoid the gene silencing that can be caused by the identity shared between the BPMV clones and the soybean genes (Supplementary Material 1). Basically, one nucleotide was changed to each 15bp sequence stretch. These alterations were made into a form which the original amino acids were conserved even with the codons changes. Some additions were also made in the transcripts sequences (Supplementary Material 1): 1.The BamHI

and KpnI restriction sites in the 5’ 3’ ends of the sequences, respectively. 2. The HA-tag sequence in the 5´ end is designed to facilitate the detection and isolation of the proteins. 3. A 15bp fragment of the destination vector sequence in both transcript sequences enables the subsequent ligation through the Gibson assembly method (New England Biolabs, Ipswich, MA, USA). The altered sequences were then synthesized and ligated into BamHI and KpnI digested pBPMV-IA-D35 vector (Zhang *et al*., 2010) by the Gibson assembly method to generate the constructs pBPMV-D35:Gm56-1, pBPMV-D35:Gm56-2 and pBPMV-D35:Gm56-3 (Table 1r to t). The obtained construct will be used in future experiments to assess the effect of the GmWRKY156 transcripts overexpression in the soybean phenotype under stress conditions and identify possible differences among the functions of transcripts in the soybean stress response.

4. Protoplast Transactivation Assay

Several methods as Chromatin immunoprecipitation, DNA Immunoprecipitation, One-hybrid system and stable transformation of promoter:gene reporter systems have been employed to study the specific interactions between DNA and transcription factors (TF) in plant (Meng *et al*., 2005; Berger *et al.*, 2007). The One-hybrid system is based on the transcription factor ability to regulate the transcription through physical DNA-protein interactions. Two components comprise this technique, the effector and the reporter vector. In brief, in the effector (or prey) construct a strong promoter is fused to a TF coding sequence to keep its expression tuned to the highest level. While the reporter construct (or bait) is obtained from a promoter sequence of interest cloned upstream of a gene encoding a reporter protein that can be easily detected (Reece-Hoyes and Walhout, 2012; Brand *et al*., 2013). So, interactions between TF and promoter sequence are identified by the downstream reporter gene activation. Although efficient, the bacterial (Meng *et al*., 2005) and yeast one-hybrid (Fields and Song, 1989) assays have the drawback to be heterologous systems (Wehner *et al*., 2011)

Protoplast are plant cells from which the cell wall has been enzymatically or mechanically removed (Eeckhaut *et al*., 2013). Because of their high plasticity, they

have been used successfully in plant genetic research (Chupeau *et al*., 2013). Moreover, the use of protoplast as experimental material enables the analysis of a gene in isolation but still in its native cellular environment, increasing the reliability of the results. Therefore, the Protoplast One-hybrid Assay also called Protoplast Transactivation Assay (PTA) constitutes a great system to assess the interaction between TF and the promoter region of its potential target genes.

However, the PTA require smaller vectors than the “regular” ones used for stable transformations. Large vectors decrease the protoplast transfection rate impairing the achievement of robust and reproducible results (Berger *et al.*, 2007). Here, vectors proper to PTA were constructed. Effector vectors carrying *GmWRKY56* transcripts sequences and reporter vector carrying *GmWRKY106* promoter fragment are obtained in order to better understand the *GmWRKY56* and *GmWRKY106* genes similarity in the expression behavior under biotic (Bencke-Malato *et al*. 2014) and abiotic (Dias *et al*. 2016) stress conditions. These vectors will be applied in future experiments to elucidate if and which of the proteins encoded by *GmWRKY56* play as a *GmWRKY106* regulator.

4.1 Construction of vectors designed for the Protoplast Transactivation Assay

To obtain the reporter vector, the 1Kb promoter fragment upstream from TSS of *GmWRKY106* gene was isolated from EMBRAPA48 and BR16 soybeans cultivar by PCR using the specific primers (Table 1y and z), subcloned into the pENTR/D-TOPO vector (Invitrogen, Carlsbad, CA, USA) and then recombined into the pORGW vector (Figure 1 c) by Gateway recombination (Invitrogen, Carlsbad, CA, USA). In the resulting reporter constructs pOR-GW: Gm106.1BR and pOR-GW:Gm106.1EM, the promoter fragments are fused with the firefly luciferase gene. Regarding the effector vectors the cloning process was as follows: The *GmWRKY56* transcripts were isolated using specific primers (Table 1u to x) and the constructs pBPMV-D35:Gm56-1, pBPMV-D35:Gm56-2 and pBPMV-D35:Gm56-3 (Table 1r to t) were used as templates in the polymerase chain reaction. The amplified sequences were subcloned into the pENTR/D-TOPO vector (Invitrogen, Carlsbad, CA, USA) and introduced into pCSVMV-

GW vector (Pruneda-Paz *et al*. 2014) (Figure 1 d) by Gateway recombination (Invitrogen, Carlsbad, CA, USA)**.** In the resulting effector constructs pCSVMV-GW: Gm56-1, pCSVMV-GW:Gm56-2 and pCSVMV-GW:Gm56-3 (Table 1u to x), transcription of GmWRKY56 is directed by the Cassava vein mosaic virus (CsVMV) promoter, which is a strong constitutive promoter. These constructs will be used in future experiments to identify interactions between the GmWRKY56 proteins and the GmWRKY106 1Kbp promoter. The assessment of the differences among the interaction profiles of these proteins with the promoter fragment is also of interest to address regulation of stress responses by alternative splicing.

REFERENCES

Bencke-Malato, M., Cabreira, C., Wiebke-Strohm, B., Bücker-Neto, L., Mancini, E., Osorio, M.B., Homrich, Milena, S., Turchetto-Zolet, A., De Carvalho, M., Stolf, R., Weber, R., Westergaard, G., Castagnaro, A.P., Abdelnoor, R.V., Marcelino-Guimarães, F.C., Margis-Pinheiro, M., Bodanese-Zanettini, M. (2014) Genome-wide annotation of the soybean WRKY family and functional characterization of genes involved in response to *Phakopsora pachyrhizi* infection. BMC Plant Biology. 14 (1), 236

Berger, B., Stracke, R., Yatusevich, R., Weisshaar, B., Flügge, U. I., Gigolashvili, T. (2007) A simplified method for the analysis of transcription factor-promoter interactions that allows high-throughput data generation. *Plant Journal*. *50*(5), 911–916

Brand, L. H., Henneges, C., Schüssler, A., Kolukisaoglu, H. Ü., Koch, G., Wallmeroth, N., Hecker, A., Thurow, K., Zell, A., Harter, K., Wanke, D. (2013) Screening for Protein-DNA interactions by Automatable DNA-Protein Interaction ELISA. *PLoS ONE*. *8*(10), e75177

Chupeau, M. C., Granier, F., Pichon, O., Renou, J. P., Gaudin, V., Chupeau, Y. (2013) Characterization of the Early Events Leading to Totipotency in an Arabidopsis Protoplast Liquid Culture by Temporal Transcript Profiling. The Plant Cell. *25*(7), 2444–2463

Dawson, W. O. (2011) Finite element modeling of kirschner pin and bone thermal contact during drilling. Current Topics in Microbiology and Immunology*. 375, 1–18*

Dias, L. P., Oliveira-Busatto, L. A., Bodanese-Zanettini, M. H. (2016) The differential expression of soybean [Glycine max (L.) Merrill] WRKY genes in response to water deficit. Plant Physiology and Biochemistry, 107, 288–300

Eeckhaut, T., Lakshmanan, P. S., Deryckere, D., Bockstaele, E.V., Huylenbroeck, J.V. (2013). Progress in plant protoplast research. Planta. 238, 991–1003

Fields, S., and Song, O.K. (1989) A novel genetic system to detect protein-protein interactions. Nature. 340, 245–246

Giesler, L. J., Ghabrial, S. A., Hunt, T. E., & Hill, J. H. (2002) Bean pod mottle virus: A Threat to U.S. Soybean Production. Plant Disease. 86(12), 1280–1289

Gleba, Y. Y., Tusé, D., Giritch, A. (2013) Plant Viral Vectors for Delivery by Agrobacterium. Current Topics in Microbiology and Immunology. 375: 155–192

Guo, H-S., Li, W-X., Ding, S-W. (2000) Suppression of post-transcriptional gene silencing by a plant viral protein localized in the nucleus. The EMBO Journal, *19*(7), 1672–1680

Karimi, M., Inze, Depicker, A. (2002) Gateway vectors for Agrobacterium-mediated plant transformation. Trends in Plant Science. 7, 193–195

Meng, X., Brodsky, M. H., Wolfe, S. A. (2005) A bacterial one-hybrid system for determining the DNA-binding specificity of transcription factors. Nature Biotechnology. 23(8): 988–994

Nakagawa, T., Kurose, T., Hino, T., Tanaka, K., Kawamukai, M., Niwa, Y., Toyooka, K., Matsuoka, K., Jinbo, T., Kimura, T. (2007) Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. Journal of Bioscience and Bioengineering. 104,34–41

Pflieger, S., Blanchet, S., Meziadi, C., Richard, M. M. S., Thareau, V., Mary, F., Mazoyer, C., Geffroy, V. (2014) The “one-step” Bean pod mottle virus (BPMV)-derived vector is a functional genomics tool for efficient overexpression of heterologous protein, virus-induced gene silencing and genetic mapping of BPMV R-gene in common bean (Phaseolus vulgaris L.). BMC Plant Biology. 14, 232

Pruneda-Paz, J. L., Breton, G., Nagel, D. H., Kang, E., Bonaldi, K., Doherty, C.J., Ravelo, S., Galli, M., Ecker, J. R., Kay, S. A. (2014) A genome-scale resource for the functional characterization of Arabidopsis transcription factors. Cell Reports . 8(2), 622–632

Reece-Hoyes, J. S., Walhout, A.J. M. (2012) yeast one-hybrid assays: A historical and technical perspective. Methods. 57(4), 441–447

Todd, A. T., Liu, E.,Page, J. E. cDNA Libraries for Virus-Induced Gene Silencing. In: *Plant Epigenetics: Methods and Protocols, Methods in Molecular Biology.* Springer Science, *2010. Chapter 16, 221–236*

Vaucheret, H., Béclin, C., Fagard, M. (2001) Post-transcriptional gene silencing in plants by RNA. Journal of Cell Science. 114, 3083-3091

Wehner, N., Hartmann, L., Ehlert, A., Bottner, S., Onate-Sanchez, L., Droge-Laser, W. (2011) High-throughput protoplast transactivation (PTA) system for the analysis of Arabidopsis transcription factor function. The Plant Journal. 68, 560–569

Whitham, S. A., Lincoln, L. M., Chowda-Reddy, R. V., Dittman, J. D., O'Rourke, J. A., Graham, M. A. (2016) Virus-Induced Gene Silencing and Transient Gene Expression in Soybean *(Glycine max)* Using *Bean Pod Mottle Virus* Infectious Clones. Current Protocols in Plant Biology. *1*, 263–283

Zhang, C., Bradshaw, J. D., Whitham, S. A., Hill, J. H. (2010)The Development of an Efficient Multipurpose Bean Pod Mottle Virus Viral Vector Set for Foreign Gene Expression and RNA Silencing. Plant Physiology. *153*(1), 52–65

Zhang, C., Ghabrial, S. A. (2006) Development of Bean pod mottle virus-based vectors for stable protein expression and sequence-specific virus-induced gene silencing in soybean. Virology. 344(2), 401–411

Table 1. Primers sequences and constructs purposes

	Constructs / Primers Names	Primers sequence (5'→ 3')	Fragment amplified	Constructs Purpose
a	pHGWFS7: Gm46.5 BR	F: **CACC**TGGACCCTATCATTAACCAAT R: TTTTTGTGTTGTTTATAATATAGTCAAAAGG	500bp sequence upstream TSS of GmWRKY46 gene isolated from BR16 cultivar	Promoter activity analysis promoter fragment fused to gfp::gus
b	pHGWFS7: Gm46.5 EM	F: **CACC**TGGACCCTATCATTAACCAAT R: TTTTTGTGTTGTTTATAATATAGTCAAAAGG	500bp sequence upstream TSS of GmWRKY46 gene isolated from EMBRAPA48 cultivar	Promoter activity analysis promoter fragment fused to gfp::gus
c	pHGWFS7: Gm106.5 BR	F: **CACC**CGCCACTCCAGAGAAGCTAT R: GGCTTCCAATAAAATTCAAGCA	500bp sequence upstream TSS of GmWRKY106 gene isolated from BR16 cultivar	Promoter activity analysis promoter fragment fused to gfp::gus
d	pHGWFS7: Gm106.5 EM	F: **CACC**CGCCACTCCAGAGAAGCTAT R: GGCTTCCAATAAAATTCAAGCA	500bp sequence upstream TSS of GmWRKY106 gene isolated from EMBRAPA48 cultivar	Promoter activity analysis promoter fragment fused to gfp::gus
e	pHGWFS7: Gm56.5 EM	F: **CACC**TGATTTGATGAAGATGTGATATTTGA R: GTCTTCCAATAAAATTCAAAGAGG	500bp sequence upstream TSS of GmWRKY56 gene isolated from EMBRAPA48 cultivar	Promoter activity analysis promoter fragment fused to gfp::gus
f	pHGWFS7: Gm106.1 EM	F: **CACC**TCAAAATTCTAACATTGTTGATG R: GGCTTCCAATAAAATTCAAGCA	1Kbp sequence upstream TSS of GmWRKY106 gene isolated from EMBRAPA48 cultivar	Promoter activity analysis promoter fragment fused to gfp::gus
g	pHGWFS7: Gm56.1 EM	F: **CACC**TGTTACGCGTAGAAGATAGAAAG R: GTCTTCCAATAAAATTCAAAGAGG	1Kbp sequence upstream TSS of GmWRKY56 gene isolated from EMBRAPA48 cultivar	Promoter activity analysis promoter fragment fused to gfp::gus
h	pHGWFS7: 35S	F: **CACC**ACTAGAGCCAAGCTGAT R: ACATTACAATTTACTATTCTAGTCGA	35S CaMV promoter	Promoter activity analysis promoter fragment fused to gfp::gus
i	pGWB504:Gm46.5 BR	F: **CACC**TGGACCCTATCATTAACCAAT R: TTTTTGTGTTGTTTATAATATAGTCAAAAGG	500bp sequence upstream TSS of GmWRKY46 gene isolated from BR16 cultivar	Promoter activity analysis promoter fragment fused to gfp
j	pGWB504:Gm46.5 EM	F: **CACC**TGGACCCTATCATTAACCAAT R: TTTTTGTGTTGTTTATAATATAGTCAAAAGG	500bp sequence upstream TSS of GmWRKY46 gene isolated from EMBRAPA48 cultivar	Promoter activity analysis promoter fragment fused to gfp
k	pGWB504:Gm106.5 BR	F: **CACC**CGCCACTCCAGAGAAGCTAT R: GGCTTCCAATAAAATTCAAGCA	500bp sequence upstream TSS of GmWRKY106 gene isolated from BR16 cultivar	Promoter activity analysis promoter fragment fused to gfp
l	pGWB504:Gm106.5 EM	F: **CACC**CGCCACTCCAGAGAAGCTAT R: GGCTTCCAATAAAATTCAAGCA	500bp sequence upstream TSS of GmWRKY106 gene isolated from EMBRAPA48 cultivar	Promoter activity analysis promoter fragment fused to gfp
m	pGWB504:Gm106.1 BR	F: **CACC**TCAAAATTCTAACATTGTTGATG R: GGCTTCCAATAAAATTCAAGCA	1Kbp sequence upstream TSS of GmWRKY106 gene isolated from BR16 cultivar	Promoter activity analysis promoter fragment fused to gfp

n	pGWB504: Gm106.1 EM	F: **CACC**TCAAAATTCTAACATTGTTGATG R: GGCTTCCAATAAAATTCAAGCA	1Kbp sequence upstream TSS of GmWRKY106 gene isolated from EMBRAPA48 cultivar	Promoter activity analysis promoter fragment fused to gfp
o	pGWB504:Gm56.1 EM	F: **CACC**TGTTACGCGTAGAAGATAGAAAG R: GTCTTCCAATAAAATTCAAAGAGG	1Kbp sequence upstream TSS of GmWRKY56 gene isolated from EMBRAPA48 cultivar	Promoter activity analysis promoter fragment fused to gfp
p	pGWB504:35S	F: **CACC**ACTAGAGCCAAGCTGAT R: ACATTACAATTTACTATTCTAGTCGA	35S CaMV promoter	Promoter activity analysis promoter fragment fused to gfp
q	pBPMV:GmWRKY106	F: CCGCTCGAGCTGCAGAAAACAAGAAGTTGGCC R: CGCGGATCCTGCTTTGATGATTTCTTCCCTTGGT	ORF fragment of GmWRKY106 gene	GmWRKY106 gene silencing via VIGS
r	pBPMV-D35:Gm56-1	……. …….	……. …….	Overexpression of GlymaWRKY56.1 transcript BPMV vector
s	pBPMV-D35:Gm56-2	……. …….	……. …….	Overexpression of GlymaWRKY56.2 transcript BPMV vector
t	pBPMV-D35:Gm56-3	……. …….	……. …….	Overexpression of GlymaWRKY56.3 transcript BPMV vector
u	pCSVMV-GW: Gm56-1	F: CACCATGGATTGCTCATCATGG R: TTAATTATTGTGCAACAATCTACCTGAGAT	GmWRKY56.1 transcript	Protoplast Transactivation Assay CsVMV promoter fused to GmWRKY56.1 transcript sequence
v	pCSVMV-GW: Gm56-2	F: CACCATGGATTGTTCTTCATGGATT R: TTAGCTTGAAATCCATTTCCCTACGA	GmWRKY56.2 transcript	Protoplast Transactivation Assay CsVMV promoter fused to GmWRKY56.2 transcript sequence
x	pCSVMV-GW: Gm56-3	F: CACCATGGATTGTTCATCTTGGATT R: CTATGCTATGTCGAAGAACCAAAGG	GmWRKY56.3 transcript	Protoplast Transactivation Assay CsVMV promoter fused to GmWRKY56.3 transcript sequence
y	pOR-GW: Gm106.1 BR	F: **CACC**TCAAAATTCTAACATTGTTGATG R: GGCTTCCAATAAAATTCAAGCA	1Kbp sequence upstream TSS of GmWRKY106 gene isolated from BR16 cultivar	Protoplast Transactivation Assay promoter fragment fused to Firefly Luciferase
z	pOR-GW: Gm106.1 EM	F: **CACC**TCAAAATTCTAACATTGTTGATG R: GGCTTCCAATAAAATTCAAGCA	1Kbp sequence upstream TSS of GmWRKY106 gene isolated from EMBRAPA48 cultivar	Protoplast Transactivation Assay promoter fragment fused to Firefly Luciferase

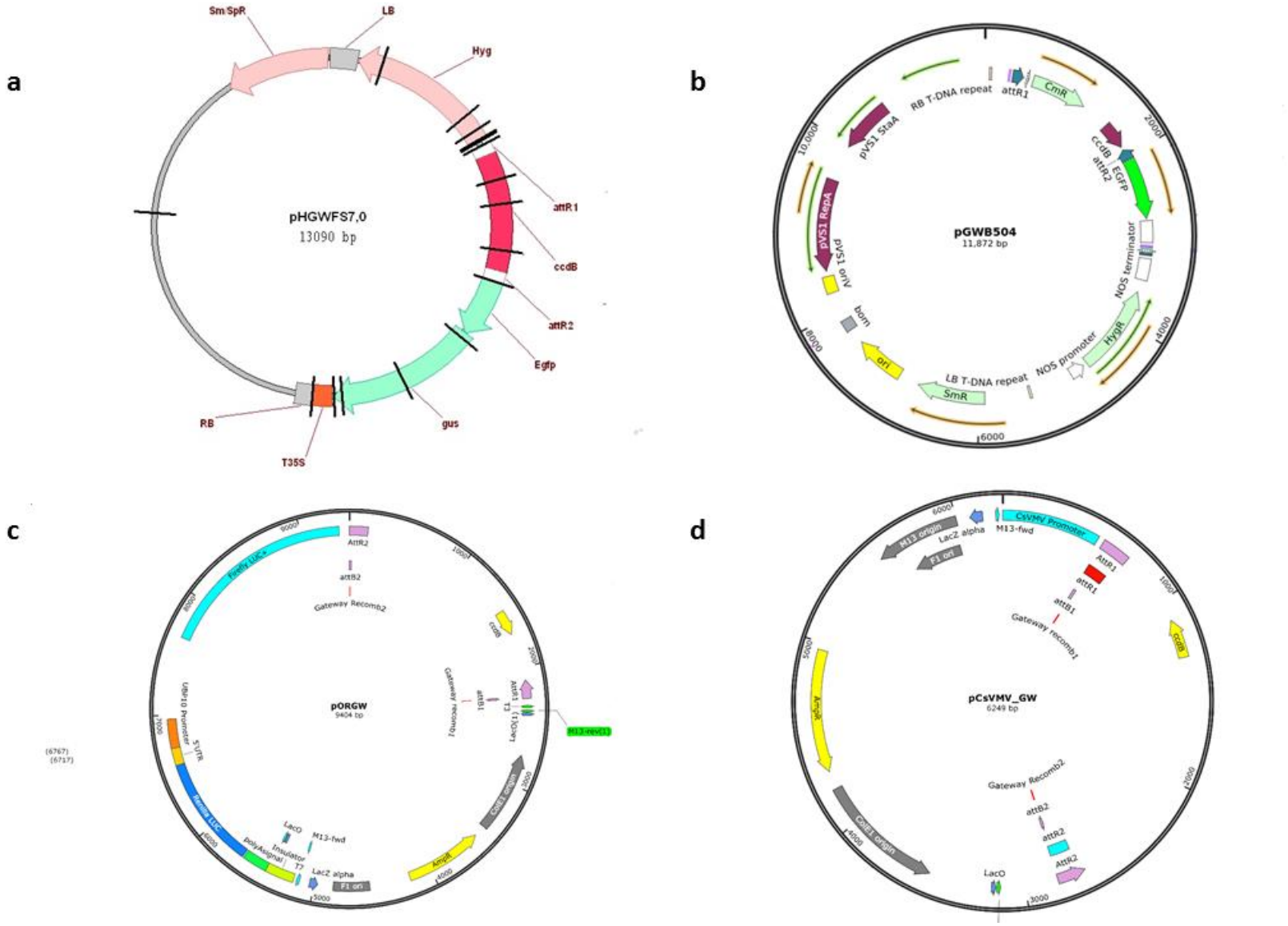

Fig 1. Maps of vectors used in the research proceedings

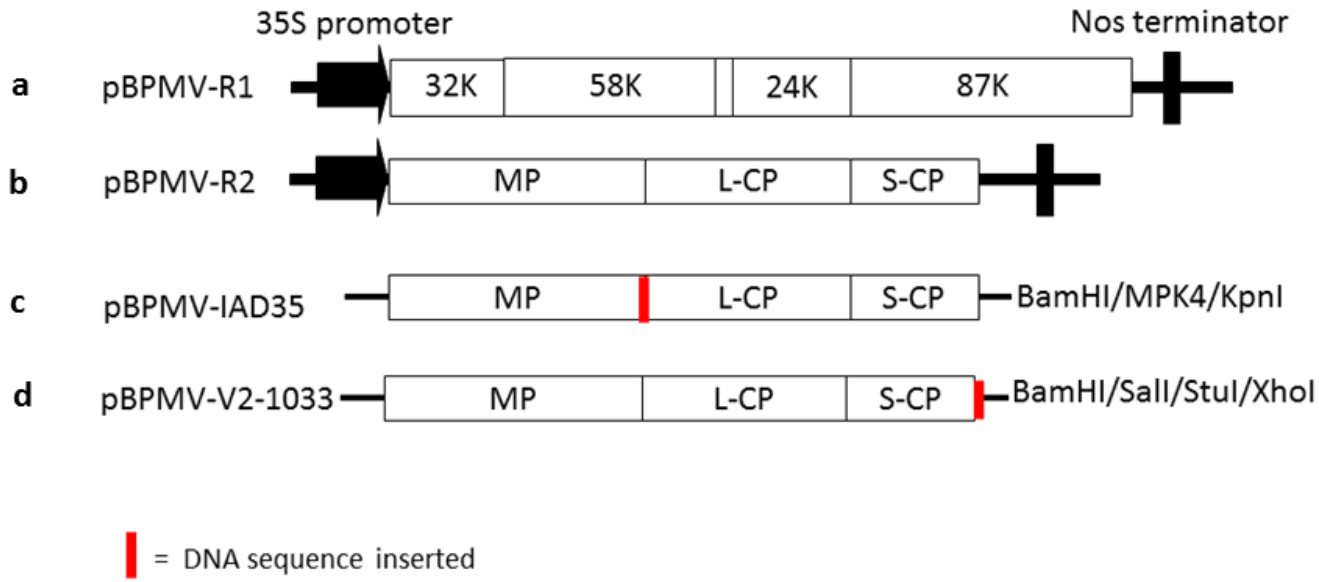

Fig 2. BPMV vectors

SUPPLEMENTARY MATERIAL 1

GmWRKY56 (Glyma.08G218600) TRANSCRIPTS SEQUENCE

→ Glyma.08G218600.1

>Glyma.08G218600.1 CDS (942 bp)

ATGGATTGTTCATCATGGATTAACACTTCCTTGGATCTCAGCATTAATCCTCGCAGAGTTCATGAAGAAGCTGT
TCCTAAGGTGGTAGAAAGCAAGCTTTTCTCTTTGGGAATGCCCAAGTTTAACGTCGAAGAAGAGTCTACTAGTG
ACTTGGAGGAGGAACTGAAGCGGGTGAGTGCAGAAAACAAGAAGTTGGCCGAAATGCTCTCAGTGGTGTGTGAG
AATTACAACACTTTGAGAAGCCATTTGATGGAATACATGAGGAAAAATGGCGAAAAGGAGGTCAGCCCAACATC
AAAGAAAAGAAAGTCTGAAAGCAGCAACAACAACAACAGTAATTTGATGGGAACTAACAATGGAAACTCAGAGA
GCAGTTCTACTGATGAAGAGTCTTGCAAGAAACCAAGGGAGGAAACCATCAAAGCAAAAATTTCAAGAGTTTAT
GTCAGGACTGAATCATCTGATACTAGCCTTATTGTGAAAGATGGATACCAATGGAGGAAATATGGACAAAAGGT
GACCAGAGATAACCCTTACCCTAGAGCATATTTCAAGTGCTCTTTTGCTCCAAGCTGCCCTGTCAAAAAGAAGG
TGCAAAGAAGTGTGGATGATCATTCTGTTCTGCTTGCTACTTATGAAGGGGAGCACAATCATCCTCAGGCTTCT
TCCCAAATGGAAGCAACATCAGGTTCTGGCCGTAGTGTGACCCTTGGTTCAGTGCCTTGTTCAGCATCTCTCAG
CACTTCCACTCCAACACTTGTTACCCTTGACTTGACAAAATCTAAGGGAAGCAACGATTCCAAGAGCACAAAAC
CTAAAGGAGATTCACCTAAAGTACCTCAGGTTTTGGTGGAACAGATGGCTACTTCTTTGACCACGGATCCTAAT
TTTAGAGCAGCACTTGTTGCTGCCATCTCAGGAAGATTGTTGCACAATAATTAA

TAA = Stop Codon

- **Glyma.08G218600.1 CDS SEQUENCE WITHOUT STOP CODON AND WITH BASES REPLACED (in green)**

GTTGAAACTCCTAAAGGATCCATGTACCCATACGATGTTCCAGATTACGCTGATTGCTCATCATGGATCAACAC
TTCCTTAGATCTCAGCATCAATCCTCGCAGAGTGCATGAAGAAGCTGTGCCTAAGGTGGTAGAGAGCAAGCTTT
TCTCATTGGGAATGCCCAAATTTAACGTCGAAGAGGAGTCTACTAGTGATTTGGAGGAGGAACTTAAGCGGGTG
AGTGCCGAAAACAAGAAGTTAGCCGAAATGCTCTCTGTGGTGTGTGAGAACTACAACACTTTGAGGAGCCATTT
GATGGAGTACATGAGGAAAAACGGCGAAAAGGAGGTGAGCCCAACATCAAAGAAGAGAAAGTCTGAGAGCAGCA
ACAACAATAACAGTAATTTAATGGGAACTAACAACGGAAACTCAGAAAGCAGTTCTACTGACGAAGAGTCTTGC
AAAAAACCAAGGGAGGAGACCATCAAAGCTAAAATTTCAAGAGTGTATGTCAGGACCGAATCATCTGATACAAG
CCTTATTGTGAAGGATGGATACCAGTGGAGGAAATATGGTCAAAAGGTGACCAGGGATAACCCTTACCCGAGAG
CATATTTCAAATGCTCTTTTGCTCCTAGCTGCCCTGTGAAAAAGAAGGTTCAAAGAAGTGTGGACGATCATTCT
GTTCTTCTTGCTACTTATGAGGGGGAGCACAATCACCCTCAGGCTTCTTCTCAAATGGAAGCAACTTCAGGTTC
TGGCCGCAGTGTGACCCTCGGTTCAGTGCCTTGCTCAGCATCTCTCAGTACTTCCACTCCAACTCTTGTTACCC
TTGATTTGACAAAATCAAAGGGAAGCAACGACTCCAAGAGCACAAAACCAAAAGGAGATTCTCCTAAAGTACCT
CAAGTTTTGGTGGAACAAATGGCTACTTCATTGACCACTGATCCTAATTTTAGGGCAGCACTTGTTGCAGCCAT
CTCAGGTAGATTGTTGCACAATAATAGGTACCCCGGGCCGCAGTATA

TACCCATACGATGTTCCAGATTACGCT= HA Tag

GTTGAAACTCCTAAA= PBMV-IA-D35 Vector sequence fragment

CCGGGCCGCAGTATA= PBMV-IA-D35 Vector sequence fragment

GGATCC= BamHI restriction site
GGTACC= KpnI restriction site

→ Glyma.08G218600.2

>Glyma.08G218600.2 CDS (663 bp)

ATGGATTGTTCATCATGGATTAACACTTCCTTGGATCTCAGCATTAATCCTCGCAGAGTTCATGAAGAAGCTGT
TCCTAAGGTGGTAGAAAGCAAGCTTTTCTCTTTGGGAATGCCCAAGTTTAACGTCGAAGAAGAGTCTACTAGTG
ACTTGGAGGAGGAACTGAAGCGGGTGAGTGCAGAAAACAAGAAGTTGGCCGAAATGCTCTCAGTGGTGTGTGAG
AATTACAACACTTTGAGAAGCCATTTGATGGAATACATGAGGAAAAATGGCGAAAAGGAGGTCAGCCCAACATC
AAAGAAAAGAAAGTCTGAAAGCAGCAACAACAACAACAGTAATTTGATGGGAACTAACAATGGAAACTCAGAGA
GCAGTTCTACTGATGAAGAGTCTTGCAAGAAACCAAGGGAGGAAACCATCAAAGCAAAAATTTCAAGAGTTTAT
GTCAGGACTGAATCATCTGATACTAGCCTTATTGTGAAAGATGGATACCAATGGAGGAAATATGGACAAAAGGT
GACCAGAGATAACCCTTACCCTAGAGCATATTTCAAGTGCTCTTTTGCTCCAAGCTGCCCTGTCAAAAAGAAGG
TAAATATATATAACCAAGTGCTAACTGGTTTGGTTATTCACATGTTTGTAGGGAAATGGATCTCAAGCTAA

TAA = Stop Codon

- **Glyma.08G218600.2 CDS SEQUENCE WITHOUT STOP CODON AND WITH BASES REPLACED (in green)**

GTTGAAACTCCTAAAGGATCCATGTACCCATACGATGTTCCAGATTACGCTGATTGTTCTTCATGGATTAACAC
CTCCTTGGATCTCAGTATTAATCCTCGTAGAGTTCATGAAGAGGCTGTTCCTAAGGTTGTAGAAAGCAAGCTCT
TCTCTTTGGGAATGCCTAAGTTTAACGTTGAAGAAGAGTCTACCAGTGACTTGGAGGAAGAACTGAAGCGGGTT
AGTGCAGAGAACAAGAAGTTGGCTGAAATGCTCTCTGTGGTGTGTGAGAACTACAACACTTTGAGGAGCCATTT
GATGGAGTACATGAGGAAGAATGGCGAAAAGGAAGTCAGCCCAACATCTAAGAAAAGAAAGTCAGAAAGCAGCA
ATAACAACAACAGTAACTTGATGGGAACCAACAATGGAAACTCTGAGAGCAGTTCAACTGATGAAGAATCTTGC
AAGAAACCAAGAGAGGAAACCATTAAAGCAAAAATTTCTAGAGTTTATGTTAGGACTGAATCATCAGATACTAG
CCTTATCGTGAAAGATGGATATCAATGGAGGAAATACGGACAAAAGGTGACAAGAGATAACCCATACCCTAGAG
CATACTTCAAGTGCTCTTTCGCTCCAAGCTGTCCTGTCAAAAAGAAAGTAAATATATATAATCAAGTGCTAACT
GGTTTAGTTATTCACATGTTCGTAGGGAAATGGATTTCAAGCAGGTACCCCGGGCCGCAGTATA

TACCCATACGATGTTCCAGATTACGCT= HA Tag

GTTGAAACTCCTAAA= PBMV-IA-D35 Vector sequence fragment

CCGGGCCGCAGTATA= PBMV-IA-D35 Vector sequence fragment

GGATCC= BamHI restriction site
GGTACC= KpnI restriction site

→ Glyma.08G218600.3

>Glyma.08G218600.3 CDS (522 bp)
ATGGATTGTTCATCATGGATTAACACTTCCTTGGATCTCAGCATTAATCCTCGCAGAGTTCATGAAGAAGCTGT
TCCTAAGGTGGTAGAAAGCAAGCTTTTCTCTTTGGGAATGCCCAAGTTTAACGTCGAAGAAGAGTCTACTAGTG
ACTTGGAGGAGGAACTGAAGCGGGTGAGTGCAGAAAACAAGAAGTTGGCCGAAATGCTCTCAGTGGTGTGTGAG
AATTACAACACTTTGAGAAGCCATTTGATGGAATACATGAGGAAAAATGGCGAAAAGGAGGTCAGCCCAACATC
AAAGAAAAGAAAGTCTGAAAGCAGCAACAACAACAACAGTAATTTGATGGGAACTAACAATGGAAACTCAGAGA
GCAGTTCTACTGATGAAGAGTCTTGCAAGAAACCAAGGGAGGAAACCATCAAAGCAAAAATTTCAAGAGTTTAT
GTCAGGACTGAATCATCTGATACTAGCCTTGTGAGTCTCTTTACTAGTACTCTCCTTTGGTTCTTTGACATAGC
ATAG

TAG = Stop Codon

- **Glyma.08G218600.3 CDS SEQUENCE WITHOUT STOP CODON AND WITH BASES REPLACED (in green)**

GTTGAAACTCCTAAAGGATCCATGTACCCATACGATGTTCCAGATTACGCTGATTGTTCATCTTGGATTAACAC
TTCTTTGGATCTCAGTATTAATCCTCGCAGGGTTCATGAAGAAGCAGTTCCTAAGGTGGTTGAAAGCAAGCTCT
TCTCTTTGGGTATGCCCAAGTTCAACGTCGAAGAAGAATCTACTAGTGACTTAGAGGAGGAACTTAAGCGGGTG
AGTGCTGAAAACAAGAAGTTAGCCGAAATGCTCTCAGTTGTGTGTGAGAATTATAACACTTTGAGGAGCCATTT
GATGGAGTACATGAGGAAGAATGGCGAAAAGGAAGTCAGCCCTACATCAAAGAAGAGAAAGTCTGAAAGTAGCA
ACAACAACAATAGTAATTTGATGGGTACTAACAATGGAAATTCAGAGAGCAGTTCAACTGATGAAGAATCTTGC
AAGAAGCCAAGGGAGGAAACTATCAAAGCAAAGATTTCAAGAGTGTATGTCAGGACCGAATCATCTGATACCAG
CCTTGTGAGCCTCTTTACTAGTACCCTCCTTTGGTTCTTCGACATAGCAAGGTACCCCGGGCCGCAGTATA

TACCCATACGATGTTCCAGATTACGCT= HA Tag

GTTGAAACTCCTAAA= PBMV-IA-D35 Vector sequence fragment

CCGGGCCGCAGTATA= PBMV-IA-D35 Vector sequence fragment

GGATCC= BamHI restriction site
GGTACC= KpnI restriction site