RAFAELA RIBEIRO REIS

ANÁLISE DO TRANSCRIPTOMA DE CULTIVARES DE SOJA COM NÍVEIS DE TOLERÂNCIA CONTRASTANTES EM CONDIÇÕES DE ESTRESSE POR DÉFICIT HÍDRICO

Londrina
2018

RAFAELA RIBEIRO REIS

ANÁLISE DO TRANSCRIPTOMA DE CULTIVARES DE SOJA COM NÍVEIS DE TOLERÂNCIA CONTRASTANTES EM CONDIÇÕES DE ESTRESSE POR DÉFICIT HÍDRICO

Tese apresentada à Universidade Estadual de Londrina, como requisito parcial para aquisição do título de doutora pelo programa de Pós-Graduação em Genética e Biologia Molecular.

Orientador: Prof. Dr. Alexandre Lima Nepomuceno

Coorientadora: Dra. Liliane Marcia Mertz Henning

Londrina
2018

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

Reis, Rafaela Ribeiro Reis.
ANÁLISE DO TRANSCRIPTOMA DE CULTIVARES DE SOJA COM NÍVEIS DE TOLERÂNCIA CONTRASTANTES EM CONDIÇÕES DE ESTRESSE POR DÉFICIT HÍDRICO / Rafaela Ribeiro Reis Reis. - Londrina, 2018.
129 f. : il.

Orientador: Alexandre Lima Nepomuceno.
Coorientador: Liliane Marcia Mertz Henning.
Tese (Doutorado em Genética e Biologia Molecular) - Universidade Estadual de Londrina, Centro de Ciências Biológicas, , 2018. Inclui bibliografia.

1. Fatores de transcrição - Tese. 2. Famílias AP2/EREBP e WRKY - Tese. 3. Respostas de tolerância aos estresse - Tese. 4. Cross-talk - Tese. I. Lima Nepomuceno, Alexandre . II. Mertz Henning, Liliane Marcia . III. Universidade Estadual de Londrina. Centro de Ciências Biológicas. . IV. Título.

RAFAELA RIBEIRO REIS

ANÁLISE DO TRANSCRIPTOMA DE CULTIVARES DE SOJA COM NÍVEIS DE TOLERÂNCIA CONTRASTANTES EM CONDIÇÕES DE ESTRESSE POR DÉFICIT HÍDRICO

Tese apresentada à Universidade Estadual de Londrina, como requisito parcial para aquisição do título de doutora pelo programa de Pós-Graduação em Genética e Biologia Molecular.

BANCA EXAMINADORA

_____

Orientador: Dr. Alexandre Lima Nepomuceno
Empresa Brasileira de Pesquisa Agropecuária - EMBRAPA

_____

Dra. Renata Fuganti Pagliarine
Autônoma - Londrina/PR

_____

Dr. André Luís Laforga Vanzela
Universidade Estadual de Londrina - UEL

_____

Dra. Michele Pires Rincão
Autônoma - Maringá/PR

_____

Dra. Renata Stolf Moreira
Universidade Estadual de Londrina - UEL

Londrina, 05 de março de 2018.

DEDICO

Ao meu pai Miguel,
Homem de honra, o qual me ensinou o poder da dedicação, esforço e amor à profissão, o valor do mérito e de cada gota de suor que escorre por dentro do jaleco branco.

OFEREÇO

À meu padrinho Marquinho (*in memoriam*),
Esse que sempre se orgulhou de mim a cada etapa cumprida, a cada encontro ao acaso, seus olhos brilhavam de admiração por me ver crescer profissionalmente, e hoje, com o título de doutora em Genética e Biologia Molecular, tenho plena convicção de que o céu está em festa e vibrando por mim nesta conquista, realizando em mim um sonho.

AGRADECIMENTOS

Ao Programa de Pós-graduação em Genética e Biologia Molecular.

À CAPES, UEL, EMBRAPA pelo apoio financeiro e incentivo à pesquisa. Ao corpo docente da UEL, por ampliar e diversificar os meus conhecimentos.

À Deus, pela implacável força e impulso de fé que sempre me acompanhou, dando forças, esperança e me reinventando a cada tropeço.

Ao meu fiel escudeiro, o qual me inspira à todos os instantes, que me aconselha e torce por mim incondicionalmente. Ele, um verdadeiro mestre na arte de ensinar e projetar pessoas, mestres e doutores. Com toda a humildade humana e acadêmica desbravou o mundo científico com muita criatividade e sabedoria, em um tempo em que até uma "simples cuba de eletroforese" era algo muito distante e novo em nosso país. Sabedoria essa que foi transmitida em gerações e venerada por todos. Hoje, recolhido em seu universo de missão cumprida e sucesso, recebe por onde passa o agradecimento daqueles que um dia foram os aprendizes e eu, uma eterna aprendiz deste homem, reverencio com muito orgulho com os dizeres: Pai, sou doutora em Genética e Biologia Molecular e este título, é mérito sobretudo seu!! (Miguel).

À minha mãe, Beatriz, pelo amor, orações com poderes sublimes! Por me apoiar nos momentos difíceis e sempre ter um gesto de carinho e injeção de ânimo.

Aos meus familiares, dinha (Virgínia) e irmão Osíris, que nunca saíram dos meus pensamentos e que sempre torcem por mim.

Ao Jean, meu esposo, companheiro, amor e amigo em todos os momentos, sou imensamente grata por Deus ter colocado você em meu caminho. Você é meu porto seguro, minha força para vencer, meu foco!! Me entende, me conforta me impulsiona, disfruta comigo o imenso dom de amar!! Amo-te!!

Aos meus amados sogros Tere e João e cunhados Lígia e Renato que sempre nos apoiam em todos os momentos e decisões.

À Agnes pela serenidade e ombro amigo nas horas de desabafo.

Às velhas e boas amizades (As + +), amigos da FUNED, PUC, UFLA, inesquecíveis que mesmo distantes estão sempre presentes em minhas memórias.

Às novas amizades, que se tornaram nossa família: Vivi, Lê, Kaka, Lolo, Lu, Isa, Ju e Danilo. Esses que compartilharam os anseios de um futuro brilhante, vivenciando cada angústia e incertezas que as vezes nos cerca. Mas, que acima de tudo, nos acolheram, foram nosso apoio à momentos difíceis e proporcionaram muita alegria, desabafo e gargalhadas.

À todos do Laboratório Biotecnologia Vegetal da Embrapa Soja, os quais dividiram

comigo seus conhecimentos, cada um com seu jeitinho e habilidades.

À Dra. Liliane Marcia Mertz Henning, por ser um exemplo de profissional e pessoa, por se destacar na sua maneira humilde de introduzir/conduzir suas opiniões e sugestões, por ser humana e paciente, mesmo quando se depara em um ambiente fortemente competitivo, enfim, por despertar em mim o gosto pela pesquisa.

Ao Dr. Alexandre Lima Nepomuceno, por acender em mim um espírito crítico e inovador que é fundamental na pesquisa, me impulsionando a enxergar além do horizonte, empreender, querer algo mais!

À Dra. Juliana Marcolino Gomes que foi fundamental para a realização e desenvolvimento deste trabalho.

À Dra. Alexandra Cardoso Moraes pelo enorme conhecimento transmitido e boa vontade em ensinar.

À Dra. Maria Cristina Neves de Oliveira, sempre disposta a ajudar.

Aos professores Dr. Luiz Filipe Pereira, Dra. Renata Stolf Moreira, Dra. Renata Fuganti Pagliarine Dr. André Luís Laforga Vanzela, Dra. Michele Pires Rincão, os quais contribuíram de forma imprescindível para o avanço e qualidade dos artigos desenvolvidos aqui.

Por fim, agradeço à todos que de alguma forma contribuíram para esta conquista vitoriosa que gerou esta tese!

REIS, Rafaela Ribeiro. **Análise do transcriptoma de cultivares de soja com níveis de tolerância contrastantes em condições de estresse por déficit hídrico**. 2018. 129 f. Thesis (Doutorado em Genética e Biologia Molecular) - Universidade Estadual de Londrina, Londrina, 2018.

RESUMO

A soja, será fortemente afetada pela mudança climática, que projeta uma contenção de área plantada de 41%, em todo o país para 2070. Dentre os estresses abióticos, o déficit hídrico é o principal fator que leva a perdas de rendimento em importantes culturas. O desenvolvimento de materiais tolerantes ao déficit hídrico é significativamente vantajoso em áreas sujeitas à ocorrência do estresse. Recentemente, alguns avanços foram obtidos na identificação de genes responsivos ao déficit hídrico, potencialmente capazes de aumentar a tolerância das plantas em condições de restrição hídrica. Compreender os mecanismos moleculares na resposta ao déficit hídrico é importante na definição de estratégias biotecnológicas e de melhoramento. As tecnologias genômicas, como o RNA-seq são ferramentas úteis para essa finalidade, podendo oferecer um melhor entendimento sobre os mecanismos ativados pelas plantas em resposta ao estresse. Objetivou-se avaliar o transcriptoma de duas cultivares de soja contrastantes na resposta de tolerância ao déficit hídrico, BR 16 (sensível) e Embrapa 48 (tolerante), em folhas e raízes submetidas a dois níveis de déficit hídrico (moderado e severo). Genes diferencialmente expressos, foram identificados e categorizados, sendo alguns selecionados para validação via RT-qPCR. Duas importantes famílias de fatores de transcrição (AP2/EREBP e WRKY) foram selecionadas para melhor entendimento de seus mecanismos de resposta e oscilações da expressão gênica em resposta ao déficit hídrico. Foram detectadas diferenças nos perfis de expressão entre os materiais avaliados, sendo possível inferir que a cultivar Embrapa 48 responde mais rapidamente ao estresse do que a cultivar BR 16, apresentando maior número de genes diferencialmente expressos (*up*-regulados) já no primeiro nível (moderado) de estresse, enquanto BR 16 exibiu um perfil de expressão predominantemente *down*-regulado no mesmo período. Esse comportamento persiste em todas as classes de genes relatados, sendo ainda mais evidente nas folhas, reforçando a ideia de que a resposta sob a influência do estresse ocorre especialmente nas folhas para ambas as cultivares. Os resultados mostram, que a cultivar Embrapa 48 além de exibir uma modulação gênica inicial mais acelerada, ainda mantém genes associados aos processos de desenvolvimento, inerente ao metabolismo primário, ativo por mais tempo. Adicionalmente, foram identificados e selecionados os 10 principais genes de cada uma das famílias de fatores de transcrição (AP2/EREBP e WRKY) na resposta de tolerância ao déficit hídrico.

Palavras-chave: AP2/EREBP. WRKY. Fatores de transcrição. ABA. Fisiologia do estresse abiótico.

REIS, Rafaela Ribeiro**. Transcriptome analysis of soybean cultivars with contrasting tolerance levels under water deficit stress conditions.** 2018. 129 p. Thesis (Doctate Degree in Genetics and Molecular Biology) – Universidade Estadual de Londrina, Londrina. 2018.

ABSTRACT

Soybean will be strongly affected by climate change, with a projected 41% restraint in cultivated land, in the whole country, for 2070. Among abiotic stresses, drought is the main factor leading to yield loss of important crops. The development of tolerant materials to water deficit is significantly advantageous in areas subject to stress. Recently, some advance has been made in the identification of genes responsive to drought, that are potentially capable of increasing plant tolerance to hydric restriction. The understanding of the molecular mechanisms involved in water deficit response is important for the definition of biotechnological and genetic improvement strategies. Genomic technologies, such as RNA-seq, are useful tools and can offer a better understanding on activated mechanisms in plants due to stress response. We aimed to assess the transcriptome of two soybean cultivars with contrasting drought response, BR 16 (sensitive) and Embrapa 48 (tolerant), in leaves and roots exposed to different levels of water deficit (moderate and severe). Differentially expressed genes were identified and categorized, and some were selected for validation with RT-qPCR. Two important transcription factor families (AP2/EREBP e WRKY) were singled out for a more detailed analysis of its response mechanisms and gene expression fluctuations under water deficit conditions. Differences were detected in expression profiles between samples, and we infered cultivar Embrapa 48 responds more rapidly to stress than cultivar BR 16, presenting a vast number of differentially expressed genes (up-regulated) at the first stress level (moderate), while BR 16 exhibited a predominantly down-regulated profile in the same level. This behavior persists in all reported gene classes, being even more evident in leaves, reinforcing the idea that stress response occurs especially in leaves, in both cultivars. Results show cultivar Embrapa 48 not only exhibits a faster initial genic modulation but also maintains genes associated to developmental processes, inherent to primary metabolism, active for longer. Additionally, we identified and selected the top ten genes from each transcription factor family (AP2/EREBP and WRKY) in water deficit tolerance response.

Keywords: AP2/EREBP. WRKY. Transcription factors. ABA. Abiotic stress physiology.

LISTA DE FIGURAS

CAPÍTULO 1

CAPÍTULO 2

MATERIAL SUPLEMENTAR

CAPÍTULO 3

MATERIAL SUPLEMENTAR

LISTA DE TABELAS

CAPÍTULO 2

MATERIAL SUPLEMENTAR

CAPÍTULO 3

MATERIAL SUPLEMENTAR

LISTA DE ABREVIATURAS E SIGLAS

A. thaliana	*Arabidopsis thaliana*
ABA	Ácido abiscísico
AP2/EREB	APETALA2/ethylene-responsive element-binding
AREB	*ABA - responsive element binding*
ATP	Adenosina trifosfato
CBF	*C-repeat binding factors*
CDPK	*Calcium-dependent protein kinases*
CNA	Confederação da Agricultura e Pecuária do Brasil
CO_2	Dióxido de carbono
CONAB	Companhia Nacional de Abastecimento
DHN	*Dehidrinas*
DNA	Ácido desoxirribonucleico
EAR	*Ethylene-responsive element binding factor-associated amphiphilic repression*
ERO	Espécies reativas de oxigênio
FT	Fator de transcrição
GR	Glutationa redutase
HSP	*Heat shock protein*
HSS	Hibridização subtrativa supressiva
ICE	*Inducer of CBF expression*
KDa	Kilodalton
LEA	*Late embryogenesis abundant protein*
MAPK	*Mitogen activated protein kinases*
MPSS	*Massive parallel signature sequencing*
NADPH	*Nicotinamide adenine dinucleotide phosphate*
O_2	Oxigênio diatômico/oxigênio molecular
PIB	Produto interno bruto nacional
POX	Peroxidases
SAGE	*Serial analysis of gene expression*
SNF	*Sucrose non-fermenting*
SOD	Enzimas superóxido dismutase
STZ/Zat10	*Recombinant zinc finger protein*

USDA	United States Department of Agriculture
WRKY	Triptofano (W), arginina (R), lisina (K) e tirosina (Y)

SUMÁRIO

CAPÍTULO 1 – CONSIDERAÇÕES GERAIS

1 INTRODUÇÃO

A soja (*Glycine max* (L). Merr.) é a leguminosa de maior importância econômica e agrícola mundial. O Brasil é o segundo maior produtor de soja ficando atrás apenas dos Estados Unidos (EUA). Os Estados Unidos, Brasil e Argentina, juntos, são responsáveis por 82,62 % da safra no mundo. Devido ao favorável clima em todas as fases da cultura, a produtividade média da safra brasileira 2016/17 foi acima da dos últimos anos, atingindo 3.364 kg/ha (COMPANHIA NACIONAL DE ABASTECIMENTO - CONAB, 2017a). Assim como outras culturas, a soja é afetada por estresses bióticos e abióticos. Desse modo, com a expansão das áreas agrícolas e oscilações climáticas, estabelece-se novas exigências relacionadas a fatores que afetam diretamente a produtividade, como o ataque de pragas e doenças (estresses bióticos) ou ocorrência de secas, inundações, altas e baixas temperaturas, alta salinidade do solo, entre outros (estresses abióticos).

As mudanças climáticas previstas para as próximas décadas, resultante do aquecimento global, são uma ameaça eminente à produção agrícola nacional. Uma vez que a agricultura é uma atividade fortemente sujeita aos fatores ambientais, estudos preveem que, o aumento da temperatura no país irá reduzir a área propícia aos cultivos de soja, café, algodão, arroz, feijão e milho (AGRITEMPO, 2018). Várias estratégias podem ser utilizadas para mitigar os efeitos do ambiente, por exemplo o déficit hídrico, desde o plantio direto, manejo do solo, rotação de culturas, até o uso de ferramentas em biotecnologia. As tecnologias genômicas são ferramentas úteis para um melhor entendimento sobre os efeitos do estresse nas plantas, bem como, para auxiliar no desenvolvimento de estratégias biotecnológicas que visem a obtenção de plantas com maior tolerância aos estresses ambientais. Plantas geneticamente modificadas por meio da transgenia ou por meio das novas ferramentas de edição gênica representam uma oportunidade de minimizar os danos causados pela seca (entre outros fatores abióticos/bióticos) com o intuito de maximizar a produção em áreas marginais ou tradicionais afetadas pelo clima.

Produtos gênicos envolvendo respostas ao déficit hídrico podem ser classificados em dois grupos: aqueles oriundos de genes regulatórios e funcionais, que são aquelas que atuam diretamente contra os danos causados pelo estresse. A classe primária de proteínas regulatórias compreende os fatores de transcrição (FT) *ABA responsive element binding* (AREB), *APETALA2/ethylene responsive element binding protein* (AP2/EREBP), NAC,

Basic leucine zipper (bZIP), MYC, MYB e as proteínas cinases: *Mitogen activated protein kinases* (MAPK), *Calcium dependent protein kinases* (CDPK), receptor de proteínas cinases, proteínas ribossomais cinases, regulador transcricional de proteínas cinases e fosfatases, as quais sincronizam o sinal de transdução e a expressão gênica durante a resposta ao estresse (WANI et al., 2013).

Dentre os genes que codificam elementos regulatórios existem diversas classes e famílias, as quais já foram bem descritas e avaliadas em inúmeras espécies. Um exemplo, foi o estudo realizado por Paul et al. (2015), os quais investigaram plantas de *Oryza sativa* transformadas com um membro da família de FT AP2/EREBP, o gene *DREB1/CBF* de *Arabidopsis thaliana (A. thaliana)*, em que a sua superexpressão induziu tolerância à desidratação dos eventos transgênicos. Outro trabalho nesta mesma linha, foi desenvolvido em tabaco, o qual por meio da transgenia foi dirigida uma superexpressão do gene de *Brachypodium distachyon*, *BdWRKY36*, pertencente à família de FT WRKY, em que obtiveram uma resposta de tolerância à seca nos eventos avaliados (SUN et al., 2015). Em *A. thaliana*, aproximadamente 6% do proteoma é classificado como FT (RAYKO et al., 2010), tais proteínas possuem importante papel em rotas metabólicas e regulatórias, bem como um envolvimento na detoxificação, biossíntese de osmólitos, proteólises de substratos celulares, canais de água, transporte de íons, *Heat shock protein* (HSP) e *Late embryogenesis abundant protein* (LEA) (JOSHI et al., 2016b).

De modo geral, desenvolver e utilizar-se de ferramentas biotecnológicas/genômicas na identificação/caracterização de promotores, genes, rotas metabólicas entre outros fatores que respondem ao déficit hídrico induzindo tolerância, torna-se fundamental para definição dos melhores alvos a serem alterados por meio da engenharia genética de plantas. Ao cruzar as fronteiras da incompatibilidade genética entre os indivíduos, as ferramentas de engenharia genética possibilitam gerar intercâmbio gênico entre espécies e incorporam características de difícil introgressão na cultura (por ser de natureza complexa e que envolve a modulação/resposta de inúmeros genes em conjunto). Ou ainda, todo o entendimento da função dos genes, atua na viabilidade da modificação de forma precisa e pontual do genoma alvo, conferindo maiores chances na obtenção de uma planta com características de interesse agrícola.

Em soja, a caracterização molecular aprofundada para tolerância à seca, pode ser amplamente explorada. Devido à complexidade que contempla as respostas de tolerância e levando-se em consideração sua identidade poligênica, deixa margem para um vasto universo gênico ainda não explorado. Apesar dos esforços desprendidos por Rodrigues et al. (2012),

por exemplo, os quais empregaram a Hibridização subtrativa supressiva (HSS), técnica que permitiu a investigação de genes diferencialmente expressos e do perfil/modulação das classes gênicas em condições de seca, das cultivares BR 16 e Embrapa 48, não foi possível o retrato mais informativo de genes diferencialmente expressos como nos dados gerados por RNA-seq, apresentados aqui neste estudo.

Oya et al. (2004) determinaram as cultivares BR 16 e Embrapa 48 como sensível e tolerante ao déficit hídrico respectivamente. Nesse estudo, foi mensurada a suscetibilidade de dez cultivares de soja fundamentando-se na resposta de rendimento das mesmas condições de déficit hídrico. Baseando-se nesta classificação dos materiais BR 16 (sensível) e Embrapa 48 (tolerante), o objetivo deste trabalho foi avaliá-los determinando as principais respostas moleculares de diferentes tecidos (folha e raiz), sob desidratação, a partir da análise do transcriptoma utilizando a plataforma de RNA-seq.

1.1 OBJETIVO GERAL

Obter dados de transcriptoma de soja sob déficit hídrico em cultivares com sensibilidade contrastante BR 16 (sensível) e Embrapa 48 (tolerante), com ênfase no metabolismo de categorias gênicas funcionais e em duas famílias de FT, AP2/EREBP e WRKY.

1.2 OBJETIVO ESPECÍFICO

a) Detectar genes diferencialmente expressos e seus respectivos níveis de dinâmica de expressão sob os tratamentos (estresse por déficit hídrico moderado e severo), de acordo com o tecido (folha e raiz) e genótipos (Embrapa 48 e BR 16);
b) Realizar uma análise comparativa da oscilação da expressão gênica entre as duas cultivares (Embrapa 48 e BR 16);
c) Selecionar os 10 genes, classificados como pertencentes às famílias de FT (AP2/EREBP e WRKY), mais relevantes na resposta de tolerância ao déficit hídrico.
d) Obter informações da região promotora destes genes “Top 10” das duas famílias de transcrição.

1.3 HIPÓTESES LEVANTADAS

a) A Embrapa 48 responde mais fortemente ao déficit hídrico exibindo um maior número de genes diferencialmente expressos em relação à BR 16;
b) A Embrapa 48 apresenta uma resposta de tolerância mais rápida, exibindo uma expressão gênica diferencial deste o nível moderado de estresse, comparada à BR 16;
c) A Embrapa 48 exibe maior número de genes classificados como FT comparado com a BR 16;
d) A Embrapa 48 apresenta maior número de genes up-regulados pertencentes à família gênica AP2/EREBP em relação à BR 16;
e) Espera-se uma modulação gênica mais intensa em folhas quando comparada a raízes.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 IMPORTÂNCIA ECONÔMICA DA SOJA NO BRASIL E NO MUNDO

Em um contexto mundial a soja é uma das *commodities* mais importantes economicamente. Essa relevância está fortemente associada ao comércio do complexo soja, as quais lideram a linha das exportações do Brasil com tendência à permanência desta elevação ainda por um longo período, visto que a demanda de soja continua crescendo e com satisfatório valor de mercado (CONAB, 2017b). Segundo estimativa da CONAB (2017b), para a safra 2017/18 da oleaginosa, o EUA permanece liderando com notáveis 33,60 %, seguido por Brasil representando 31,04 % e finalmente Argentina, com sua parcela de 16,54 % da safra mundial.

Com a demanda crescente pela leguminosa e por seus derivados, as transações também seguem no mesmo ritmo. Segundo o United States Department of Agriculture (USDA, 2018), 150,4 milhões de toneladas de soja em grãos devem ser movimentadas mundialmente, 4,2 % a mais que na safra 2016/17. A China é o principal país em que as importações devem aumentar, com 97 milhões de toneladas (+3,75 %), seguida pela União Europeia, com 14 milhões de toneladas (+4,6 %), México (+4,2 %), Japão (+3,9 %), Tailândia (+2,4 %) e Egito (+32,4 %). Diante destes dados, observa-se que a demanda mundial por soja se mantém estável, ainda sustentada por sua aplicabilidade na geração de farelo e óleo. A oferta da oleaginosa na safra 2017/18 também permanece sólida, uma vez que há uma preferência dos agricultores em cultivar a soja frente a outros grãos e cereais (CEPEA, 2018).

Com relação ao Brasil, segundo Pereira (2017), na safra 2017/18, será o maior exportador de soja do mundo, com 64 milhões de toneladas de soja em grãos exportadas, aumento de 4,07 %, em relação às exportações da safra 2016/17 que foram de 61,50 milhões de toneladas. O complexo soja, composto pela soja em grãos e seus derivados, foi o principal produto exportado em 2016, representando 13,72 % de toda a exportação brasileira, ou seja, US$ 25,42 bilhões, ficando à frente de produtos importantes como minérios, petróleo e combustíveis (CONAB, 2017b). Assim como em outros países produtores do grão, no Brasil, a área cultivada com soja nas principais regiões agricultáveis continua aumentando, um esforço necessário para atender a progressiva demanda pelo produto. Proporcionalmente à área cultivada com outros grãos em escala mundial, a de soja tem sido a que mais cresceu no decorrer das últimas décadas (DALL'AGNOL, 2018).

O ramo agroindustrial da soja tem expressiva importância socioeconômica no país, pois movimenta um amplo número de agentes e organizações ligados aos mais diversos setores econômicos, como empresas de pesquisa e desenvolvimento, fornecedores de insumos, indústrias de máquinas e equipamento, produtores rurais, cooperativas agropecuárias, cooperativas agroindustriais, processadoras, produtores de óleo, fabricantes de ração e usinas de biodiesel, dentre outras. Em outros termos, o complexo soja é um importante gerador de riquezas, empregos e divisas, se transformando em um dos principais vetores de desenvolvimento regional do país (HIRAKURI; LAZZAROTTO, 2014). O setor do agronegócio representa quase 23 % do total produto Interno Bruto Nacional (PIB), sendo que o PIB de cadeias no setor da soja, de janeiro a junho de 2017, teve aumento de 2,22 % com valor estimado em R$ 103,2 bilhões (CEPEA; CONFEDERAÇÃO DA AGRICULTURA E PECUÁRIA DO BRASIL - CNA, 2017).

No Brasil, a soja é a principal cultura em extensão de área e volume de produção. Tanto no Brasil, como nos Estados Unidos, a produtividade estimada para a safra 2016/17 foi muito superior à média normal das últimas safras, assim, e apesar do aumento de área estimado em 2,35 %, a CONAB (2018) prevê uma redução de produtividade para o Brasil e, consequentemente, uma redução na produção para a safra 2017/18, passando de 114 milhões de toneladas para 111,56 milhões de toneladas. Os maiores Estados produtores de soja nacional, hoje, estão localizados na Região Centro-Sul do Brasil (Mato Grosso, Rio Grande do Sul, Paraná, Goiás e Mato Grosso do Sul) que concentram 75 % da área cultivada (25,8 milhões de hectares) e 77 % da produção nacional (79,25 milhões de toneladas) (CONAB, 2017a).

A cultura da soja é amplamente difundida devido à sua versátil utilização em diferentes segmentos. Em 2016, 40,20 milhões de toneladas de grãos de soja foram esmagadas, produzindo, aproximadamente, 7,83 milhões de toneladas de óleo da leguminosa. Esse, empregado, especialmente, para o consumo humano e para fabricação de biodiesel. Além disso, significativa parte da produção do farelo de soja, juntamente com o milho, são os principais componentes da ração animal, o que torna a produção de carnes um importante elo na cadeia produtiva de grãos (CONAB, 2017a).

2.2 ESTRESSES ABIÓTICOS QUE COMPROMETEM A CULTURA DA SOJA – PERDAS RELACIONADAS AO ESTRESSE HÍDRICO

No Brasil, dentre os fatores ambientais adversos (déficit hídrico, alta salinidade, baixas e altas temperaturas, alagamento, poluentes) o déficit hídrico é o que mais afeta de maneira negativa o crescimento, metabolismo e produtividade da soja. Assim, seus efeitos nas plantas incluem redução nas taxas de assimilação de CO_2, alteração do tamanho das células foliares (epiderme diminui, parênquima esponjoso aumenta), taxa de transpiração, potencial de água nas plantas, taxa de crescimento e condutância estomática (LAWLOR, 2013), tudo isso refletindo na redução da produtividade e, também da qualidade final do produto. Nos últimos anos no país, segue uma tendência à expansão de áreas plantadas, contudo, as regiões que mais sofrem com estas ocorrências de longos períodos sem água, regiões áridas e semi-áridas, não conseguem avançar no aumento de área para cultivo, além de prejuízos econômicos em consequência da perda da produtividade (CONAB, 2014).

Os estresses abióticos, na safra 2015/16, foram os principais responsáveis pela quebra de produção ocorrida na cultura da soja, especialmente as elevadas temperaturas e os períodos de estresse hídrico (falta ou excesso de chuva) enfrentado pelas plantas. A quebra na produção nacional de soja chegou a 9,7 % totalizando 186 milhões de toneladas, por influência do fenômeno *El Niño* (CONFEDERAÇÃO DA AGRICULTURA E PECUÁRIA DO BRASIL, 2017; HIRAKURI, 2016). Apesar desse percentual aparentemente não ter relevância, a cultura da oleaginosa ocupa uma vasta área produtiva (mais de 33 milhões de hectares) o que reforça os impactos negativos, tornando-os bastante marcantes. As perdas econômicas ocorridas na safra 2015/16 foram superiores às ocorridas na safra 2013/14 (R$ 9,2 bilhões), ano em que também foi marcada por intervenções climáticas nos estados sojicultores (HIRAKURI, 2016).

Isso pode revelar um indício de que as alterações ambientais estão sendo cada vez mais impactantes na agricultura brasileira, agregando força quando se observa as más práticas de manejo, como a mobilização excessiva do solo, a baixa diversificação de espécies cultivadas, a nutrição desequilibrada do solo e o cultivo em sistema de plantio direto com pouca palha, as quais, muitas vezes não estão baseadas em critérios técnicos adequados e se transformam em um facilitador para a propagação de pragas e doenças em seus cultivos, dentre os quais a lavoura é exposta (HIRAKURI, 2016).

2.3 REMODELAÇÃO CELULAR/MOLECULAR EM CONDIÇÕES DE DÉFICIT HÍDRICO

O déficit hídrico é um agravante quando se trata da queda na produtividade das culturas, acarretando grandes perdas no mundo todo. Os danos celulares (metabólicos, bioquímicos e fisiológicos) provocados pelo estresse, comprometem o desenvolvimento da planta e isso dependerá da intensidade, duração e, principalmente, da fase do ciclo de vida em que ocorre. Tais danos podem tornar-se irreparáveis e irreversíveis podendo levar à morte vegetal (CHAVES; OLIVEIRA, 2004). Estresses abióticos, no geral, são condições ambientais comuns que afetam o crescimento e o rendimento das mais variadas culturas. O desenvolvimento adaptativo ao estresse é dependente da ativação de uma cascata de redes moleculares envolvidas na percepção do estresse, sinal de transdução, expressão de genes estresse-regulados e metabólitos.

As plantas têm respostas adaptativas estresse-específicas, assim como respostas de proteção a mais de uma interferência ambiental. Os sinais de percepção e rotas de sinalização podem ser específicos ou sofrerem *cross-talk*, sendo ativas em períodos de estresse distintos, podendo até mesmo haver estas vias cruzadas entre fatores estressantes no âmbito biótico e abiótico ao mesmo tempo. A cascata de ativação coordena e controla respostas fisiológicas e bioquímicas adaptativas necessárias (HUANG et al., 2012). Alguns ajustes na fisiologia da planta, como enrolamento e permanência de área verde foliar, deposição de cera epicuticular, fechamento estomático, aumento no crescimento de raízes e mecanismos na remodelagem bioquímica, bem como, redução da fotossíntese, resistência à fotoinibição, ajuste osmótico e estabilização de membrana em nível celular, retratam uma performance comum e marcante da planta perante a condições de estresse (KHAZAEI et al., 2013).

Após a percepção e transmissão de sinais, em decorrência do déficit hídrico nas plantas, há um aumento na concentração de ácido abscísico (ABA) no xilema, o qual regula FTs que induzem inúmeras respostas adaptativas ao estresse, em que auxiliam na redução do crescimento foliar, na alteração de crescimento e perfil das raízes, induzem o fechamento estomático limitando a taxa de transpiração, participam da ativação de genes ligados ao ajuste osmótico, compartimentalização iônica, e à condutividade hidráulica (SALISBURY; ROSS, 2013; VERSLUES; ZHU, 2005). O fechamento estomático é uma das linhas de defesa a ser ativada em resposta à deficiência hídrica em plantas. Este mecanismo resulta na menor absorção de dióxido de carbono (CO_2) e somado a baixa concentração de água na célula, leva uma redução das taxas fotossintéticas, contribuindo para menor acúmulo de biomassa. Com o fechamento estomático e consequente redução da assimilação de CO_2, há acúmulo de ATP e

do poder redutor (NADPH), além da redução no conjunto do aceptor final da cadeia de transporte de elétrons ($NADP^+$). O excesso da atividade redutora pode induzir uma limitação exagerada da cadeia de transporte de elétrons. Neste processo, pode haver escape de elétrons, que reagem com o oxigênio molecular (O_2), formando, portanto, as espécies reativas de oxigênio (EROs) (CAVATTE et al., 2011).

As EROs são produzidas quando os elétrons da molécula de água são transferidos ao oxigênio e essas, causam danos nas membranas dos tilacoides. Os cloroplastos são as organelas mais suscetíveis ao estresse oxidativo, principalmente em razão da elevada concentração de oxigênio no interior dos mesmos (CORNIC; BRIANTAIS, 1991; CORNIC; MASSACCI, 1996). De acordo com estudos de Miller et al. (2010) e Tuteja, Singh e Tuteja (2011), as enzimas superóxido dismutase (SOD), glutationa redutase (GR), peroxidases (POX) e outras redutases são ativadas em resposta ao estresse oxidativo induzido pelo déficit hídrico. O efeito negativo das EROs pode ser atenuado pela presença de compostos antioxidantes (e.g. ascorbato, glutationa, tocoferol e carotenoides) (DESIKAN et al., 2004).

O conteúdo de carotenóides foliares, por exemplo, aumenta a partir do avanço do estresse por déficit hídrico, minimizando danos ao aparato fotossintético o que caracteriza uma estratégia para dissipação do excesso de energia luminosa, utilizado pelas plantas, principalmente, mediante às condições impostas pela deficiência hídrica, na qual os carotenóides desempenham papel fotoprotetor (ROJAS et al., 2012). Os carotenóides absorvem espectro de luz visível de 400 – 700 nm, comportando-se como um filtro à radiação ultravioleta, a qual atua em torno de 200 – 400 nm. Portanto, reduz os possíveis danos celulares oriundos da incidência intensa de luz, bem como exerce função fotoprotetora a partir da ágil extinção dos estados excitados da clorofila, preservando, assim, as células da fotooxidação (TAIZ; ZEIGER, 2009).

A queda na taxa fotossintética nem sempre é resultado da redução da condutância estomática (MANSFIELD; DAVIES, 1981). As mudanças na atividade fotossintética podem estar associadas aos danos de membrana nas células do mesofilo, a diminuição no teor de clorofila ou, ainda, alteração na síntese e transporte de assimilados (CORNIC; MASSACCI, 1996). Este decréscimo na taxa de fotossíntese líquida sob estresse, também está relacionado com danos bioquímicos de natureza não-estomática, causada pela oxidação dos lipídios encontrados no cloroplasto, especialmente na membrana dos tilacóides e às mudanças na estrutura dos pigmentos e proteínas. Além do comprometimento da operação do ciclo de Clavin, como por exemplo baixa atividade da rubisco, ou ainda incapacidade de regenerar a ribulose- 1,5- bifosfato, devido a queda de NADPH e ATP (TAIZ; ZEIGER, 2013).

Os efeitos primários de plantas sob déficit hídrico são por exemplo, redução potencial de água e desidratação celular, os quais alteram diretamente as propriedades físicas e bioquímicas das células, que por sua vez, podem provocar efeitos secundários. Sendo esses, a alteração da atividade metabólica, citotoxidade iônica e produção EROs, as quais iniciam e aceleram o rompimento da integridade celular, podendo levar a morte celular (TAIZ; ZEIGER, 2009). O ácido abiscísico (ABA) regula a expressão de vários genes sob condições de dessecação e estresse osmótico (INGRAM; BARTLES, 1996). Já a resposta na raiz é distinta, ao contrário do que ocorre na parte aérea, há expansão do sistema radicular, pois o ABA inibe a síntese de etileno local (BLUM, 1997; HANSON; HITZ, 1982; HSIAO, 1973). Na porção aérea da planta, ocorre um aumento das taxas de citocininas e auxinas, protegendo o aparelho fotossintético e cloroplastos, amenizando o efeito do déficit hídrico, permitindo a recuperação mais rápida da planta frente ao estresse (KRAMER; BOYER, 1995; YORDANOV; VELIKOVA; TSONEV, 2000).

Em plantas submetidas ao estresse, o ajuste osmótico pode ocorrer a partir do acúmulo de solutos orgânicos de baixo peso molecular, aminoácidos livres e íons inorgânicos. O ajuste osmótico, é um processo pelo qual o potencial de água da planta pode ser reduzido sem que haja decréscimo no turgor celular, pois as mudanças no potencial de água resultam em mudanças no potencial osmótico, ou seja, o ajuste osmótico é o aumento líquido na concentração de soluto na célula. A manutenção do turgor possibilita a continuação da expansão celular e favorece manutenção de altas condutâncias estomáticas sob potenciais de água menores, sugerindo que o ajuste osmótico aumenta a tolerância à seca. Durante o ajuste osmótico, ocorre compartimentalização de íons no vacúolo, prevenindo enzimas citosólicas e outras organelas dos efeitos nocivos dos íons. Com isso, solutos compatíveis (prolina, manitol, glicina-betaína) são acumulados no citoplasma. Como eles são compostos orgânicos osmoticamente ativos, as moléculas de água formam uma camada de solvatação espessa, que favorece a redução do potencial osmótico, tornando o potencial de água mais negativo na planta. Portanto, o ajuste osmótico mantém o gradiente decrescente de potencial de água entre o solo e a planta durante condições de déficit hídrico, permitindo que as plantas continuem a absorver água do solo (TAIZ; ZEIGER, 2009).

Os osmólitos mais comumente encontrados são os carboidratos solúveis (sacarose, frutose e glicose) e prolina, apesar dos íons inorgânicos também contribuírem para o ajuste osmótico (Na^+, NO_3^-, $SO4^-$, K^+, Cl^-, Ca^+) (CHEN; JIANG, 2010). O acúmulo de prolina é uma característica relevante em muitas espécies sob deficiência hídrica (BANDURSKA, 2000; BANDURSKA; GO´RNY; ZIELEZIN´SKA, 2008; COSTA; HUANG, 2006; JAVADI;

ARZANI; EBRAHIMZADEH, 2008). Este aminoácido regula o acúmulo de nitrogênio utilizável, é osmoticamente ativo e contribui para a estabilidade de membranas (OZTURK; DEMIR, 2002). Além disso, a prolina pode agir como uma molécula sinalizadora e reguladora capaz de ativar múltiplas respostas, que são elementos do processo de adaptação e tolerância. Outra função importante é a capacidade antioxidante, estabilizando EROs e ainda a ativação de enzimas de desintoxicação de vias alternativas (MATYSIK et al., 2002; RUGGIERO et al., 2004; SZABADOS; SAVOURE, 2010).

Além dos osmoprotetores citados acima, existem moléculas cuja função é igualmente importante na proteção da estrutura celular ao mais severo nível de estresse, são o grupo das chaperonas que inclui as proteínas: *Late embryogenesis abundant proteins* (LEA) e *Dehidrinas* (DHN). As proteínas LEA desempenham um papel de proteção, ativam o sistema de reparo de danos das membranas celulares. Estas proteínas são hidrofílicas e ligam-se fortemente à água. Sua função pode estar associada à capacidade de reter água e deste modo evitar a cristalização de proteínas e outras moléculas importantes durante períodos de seca. Em relação as LEA, têm-se como exemplo as proteínas de choque térmico *Heat shock proteins* (HSP), as quais atuam na ativação do sistema de reparo das membranas, além de participarem da síntese, degradação, transporte, maturação e proteção de proteínas em vários tecidos e em estádios de maturidade também distintos. Podem ainda recuperar a forma original de proteínas inativadas pela dessecação (BOHNERT et al., 2001; SEKI et al., 2001). Já as DHN, estabilizam macromoléculas permitindo o crescimento celular em raízes sob déficit hídrico e assim como outras chaperonas atuam na proteção de proteínas contra agentes desnaturantes, mediante a captura de íons e reestabelecimento das pontes de hidrogênio, o que auxilia na estabilidade de membrana (CHAVES; OLIVEIRA, 2004).

2.4 FATORES DE TRANSCRIÇÃO E SEU PAPEL NA TOLERÂNCIA A ESTRESSES

Os FT participam na cascata de ativação gênica de diversas vias metabólicas relacionadas a estresses bióticos e abióticos. O estresse em plantas inicia-se a partir de uma complexa via de respostas, começando com a percepção do estresse, o qual desencadeia uma cascata de eventos moleculares, sendo finalizada em vários níveis de respostas fisiológicas, metabólicas e de desenvolvimento (BRAY, 1993). Mudanças de expressão gênicas, induzidas por estresse, por sua vez podem participar na geração de fitohormônios tais como ABA, ácido salicílico e etileno. Estas moléculas podem amplificar os sinais e induzir uma segunda etapa de sinalização podendo seguir a mesma via metabólica ou usar outra via metabólica. Inúmeros

FTs atuam na resposta ao estresse, alguns são dependentes de ABA e outros não, tais vias são denominadas ABA-dependente ou ABA-independente quando assim são ativadas. Existem elementos que são induzidos por ambas as vias, podendo ocorrer respostas cruzadas entre estas diferentes vias de sinalização (LEUNG; MERLOT; GIRAUDAT, 1997; WANG et al., 2008).

Fatores de transcrição são ativados ou reprimidos por proteínas cinases ou fosfatases, e regulam diretamente a expressão de uma variedade de genes abaixo da cascata de sinalização (*downstream*), interagindo com o *cis*-elemento específico na região promotora. Adicionalmente, os próprios FTs são regulados por outros componentes acima da cascata de sinalização (*upstream*) (HIRAYAMA; SHINOZAKI, 2010) e após várias modificações à nível pós-traducional, como ubiquitinação e sumoilação, eles formam uma rede regulatória complexa que modula a expressão de genes sensíveis ao estresse, ativando vários processos fisiológicos e metabólicos (MIZOI et al., 2013). A ubiquitinação é uma reação rápida e reversível que consiste na adição de uma proteína, a ubiquitina, a um resíduo de lisina do substrato através de uma ligação isopeptídica (CIECHANOVER et al., 1980). Grande parte das ubiquitinas conhecidas são membros da família do *Small ubiquitin related modifier* (SUMO), essas, são pequenas com peso cerca de cerca de 12 kDa (SARACCO et al., 2007). A adição de SUMO (ou sumoilação) é uma conjugação reversível e que necessita de energia, este processo atua na regulação das atividades biológicas cruciais. A SUMOilação é importante na diversificação de respostas ao estresse das plantas (CASTRO et al., 2012). As enzimas da via do SUMO são homólogas às da via da ubiquitinação, porém são específicas do processo (MEULMEESTER; MELCHIOR, 2008) (Figura 1).

Nas últimas décadas, foram realizadas pesquisas importantes para a identificação e caracterização de diferentes FTs que contribuem para a resposta ao estresse por seca (JOSHI et al., 2016b). Em plantas, grande parte dos genes (até 10 %) são potencialmente codificados em FTs (FRANCO-ZORRILLA et al., 2014), os quais são categorizados dentro de diferentes famílias gênicas, bem como: *AREB, AP2/EREBP, MYB, WRKY, NAC, HD-Zip, Zinc Finger Proteins, bHLH, ARF/Aux-IAA* e *bZIP* e são classificadas de acordo com a diferença da estrutura do seu domínio de ligação ao DNA (GOLLDACK; LÜKING; YANG, 2011; JIN et al., 2014). São inúmeros os FTs existentes que atuam nos estresses abiótico e biótico, dentre eles o *C-repeat binding factors - CBF (CBF1, CBF2, CBF3, CBF4, CBF2-1, CBF2-2)* (CHINNUSAMY; ZHU; ZHU, 2007; KUME et al., 2005; XIONG; FEI, 2006), membros da família AP2/EREBP, que atuam, principalmente, em baixas temperaturas (frio) ativando genes *Cold-regulated genes COR (COR78/RD29A; COR47; COR15a, COR* 6.6)

(STOCKINGER; GILMOUR; THOMASHOW, 1997; YAMAGUCHI-SHINOZAKI; SHINOKAKI, 1994); *DREB1/DREB2 (DREB1a, DREB1b, DREB1c, DREB2A, DREB2B*); que atuam em situações de déficit hídrico, frio e alta salinidade; *NAM, ATAF, and CUC* (*NAC) (NAC2, NAC019, NAC055, NAC072, NAC6, NAC1, NAC2, NAC23)* (HU, 2006; NAKASHIMA et al., 2007) e *bZIP (ABF1, ABF2, ABF3, ABF4, bZIP 44, bZIP 62, bZIP78, bZIP132*) que respondem a condições de estresse abiótico e biótico (como defesa contra patógenos) (DEPPMANN; ALVANIA; TAPAROWSKY, 2006; LIAO et al., 2008a); *MYC*/*MYB (MYB2, MYB76, MYB92, MYB177)* (AGARWAL et al., 2006; LIAO et al., 2008b) que, também, são induzidos pelos estresses relacionados ao DREB (Figura 1).

Alguns elementos específicos induzem pelo seu sinal particular a expressão de FT alvo, é o caso do gene *Inducer of CBF expression (ICE) (ICE1, ICEr, ICErl)* que ativam a expressão dos membros CBFs (ZARKA et al., 2003) (Figura 1). Vários trabalhos relacionam uma tolerância ao estresse em plantas ao desempenho dos FT. A superexpressão do *AtCBF1* ou *AtCBF3* tanto em construções estresse-induzidas como constitutivas em plantas transgênicas de tomate *(Solanum lycopersicum)* (HSIEH et al., 2002); tabaco *(Nicotiana tabacum)* (KASUGA et al., 2004); trigo *(Triticum)* e arroz *(Oryza sativa)* (OH et al., 2005) proporcionam tolerância ao frio. Em outros estudos relacionados ao mesmo gene, a superexpressão e respostas em arroz *(Oryza sativa)* (DUBOUZET et al., 2003); milho *(Zea mays)* (QIN et al., 2004) também foram semelhantes.

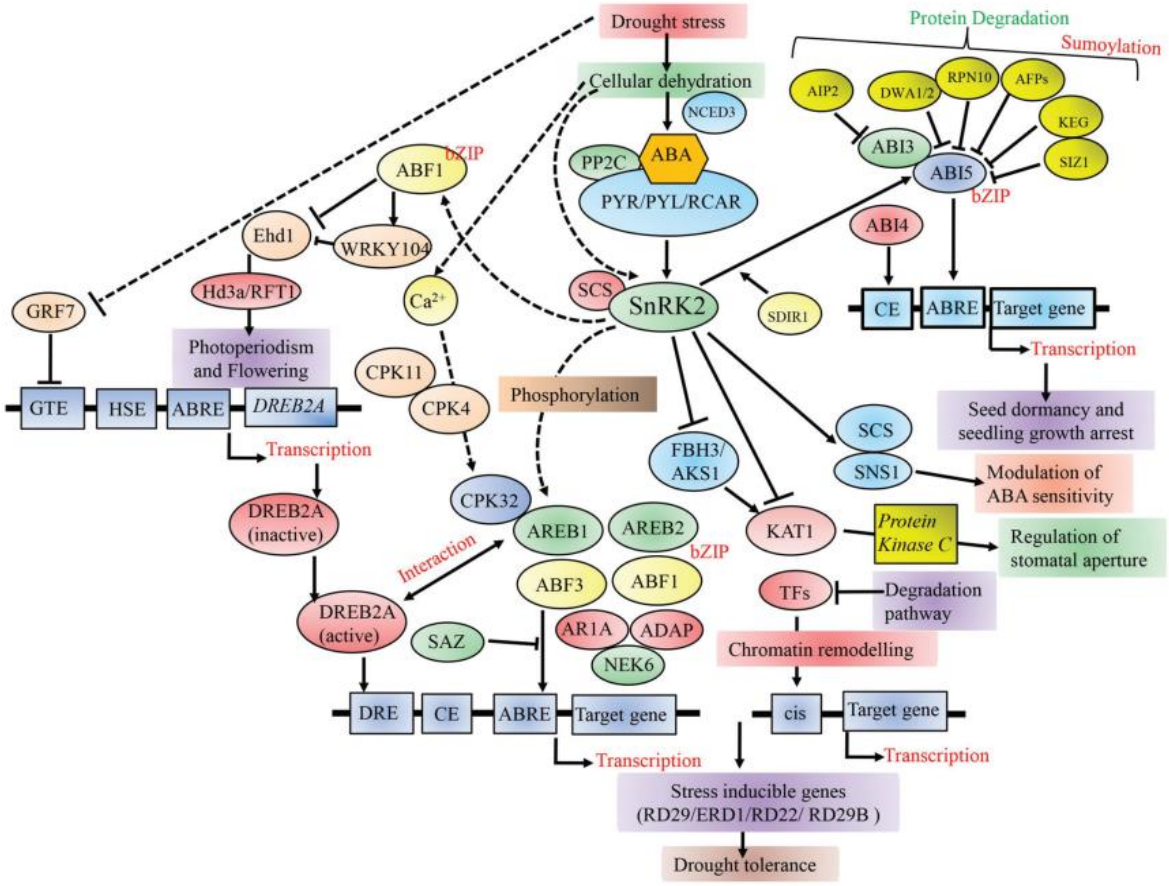
Já os FTs DREB e outros como AREB não possuem esta característica, são acionados por inúmeras moléculas e respostas. Os DREB desempenham papel importante na ativação de genes funcionais relacionados à tolerância à seca, salinidade, frio e calor (NAKASHIMA et al., 2007; SAKUMA et al., 2006b). Em *A. thaliana DREB1A* mutante, por exemplo, foi possível observar a indução constitutiva da expressão dos genes COR (alvos do CBF) (CHINNUSAMY; ZHU; ZHU, 2006). O gene *AtDREB2A*, foi inserido em plantas e essas apresentaram tolerância ao calor, desidratação e salinidade (SAKUMA et al., 2006). Análises moleculares amplas mostraram que o ABA regula a expressão de inúmeros genes sob condições de estresse osmótico, e o elemento ABA-responsivo (ABRE) é o principal *cis*-elemento para a expressão de genes sensíveis ao ABA.

Os FTs são reguladores-chave da expressão gênica (Figura 1). A proteína ABRE-ligada e o FT ABRE-ligado controlam a expressão gênica em uma via ABA-dependente. As proteínas cinases relacionadas com *Sucrose non-fermenting* (SNF1 e 2), proteínas fosfatases do tipo do grupo A e 2C e os receptores ABA controlam a via de sinalização ABA (NAKASHIMA; KAZUKO; KAZUO, 2014). As vias de sinalização independentes de ABA,

como os FTs DREB e NAC também estão envolvidas em respostas de tolerância, incluindo seca, calor e frio. Ainda de acordo com Nakashima, Kazuko e Kazuo (2014), existem interações entre a principal via de sinalização ABA e outros fatores de sinalização na resposta ao estresse, além disso, o controle de ABA ou a expressão dos fatores de sinalização do desequilíbrio ambiental podem favorecer tolerância a condições adversas. A superexpressão de FTs estresse-específicos ampliam a tolerância ao déficit hídrico e rendimento de grãos em comparação com as plantas controles no campo (FUGANTI-PAGLIARINE et al., 2017).

Zhou et al. (2008), relataram a resposta de tolerância ao frio de plantas de *A. thaliana* transgênicas contendo o gene *GmWRKY2*. Enquanto, neste mesmo estudo também em plantas de *A. thaliana* mutantes, o gene *GmWRKY54* conferiu tolerância à seca e à salinidade, o que foi atribuído a uma possível regulação do *DREB2A* e *Recombinant zinc finger protein (STZ/Zat10).* A superexpressão do gene *GmWRKY13* em um dos eventos transgênicos avaliados, por outro lado, apresentou aumento na sensibilidade ao estresse salino e osmótico (mannitol), mas reduziu a sensibilidade ao ABA quando comparado às plantas selvagens. Adicionalmente, as plantas transgênicas contendo o gene *GmWRKY13* mostraram um aumento em raízes laterais. Tais resultados indicam que, os três genes *GmWRKY* avaliados possuem papeis distintos na resposta de tolerância ao estresse abiótico. A compreensão de FTs envolvidos na regulação das respostas da planta que contribuem para a tolerância ao déficit hídrico através de suas vias/atividade e *cross-talk*, é, portanto, uma fonte potencial em estratégias para programas de melhoramento de culturas (JOSHI et al., 2016a).

Figura 1 - Modelo esquemático da regulação transcricional de diferentes FTs na resposta de tolerância ao déficit hídrico em plantas.



Fonte: Adaptada de Joshi et al. (2016a).

2.5 FATORES DE TRANSCRIÇÃO DA FAMÍLIA WRKY E AP2/EREBP

A família de fator de transcrição WRKY é uma das mais importantes em plantas, sendo amplamente distribuída entre todos os organismos vivos variando de eucariota unicelular a plantas superiores (RUSHTON et al., 2010). Estes FTs são parte integrante de muitas vias de sinalização envolvidas na regulação de diversos processos em plantas, podem atuar como repressores ou ativadores e participam da sinalização e regulação transcricional por meio das interações com uma variedade de proteínas, incluindo MAP cinases, calmodulina, histonas desacetilases, proteínas de resistência entre outras (KHAN et al., 2017). Os FTs pertencentes à família WRKY são marcados pelo domínio de DNA-ligante WRKY com cerca de 60 aminoácidos com sequências invariantes de WRKY (triptofano, arginina, lisina e tirosina) no N-terminal e uma atípica estrutura *Zinc-finger* no C-terminal. A estrutura *Zinc-finger* é CX4-5CX22-23HXH ou Cx7Cx23HxC (EULGEM et al., 2000; RUSHTON et al., 2010). FT WRKY interagem com a região *W-box* (com motif principal TTGACC/T) do

promotor de seus genes-alvo. Em soja já foram identificados 182 genes classificados como pertencentes a família WRKY (BENCKE-MALATO et al., 2014).

Mutantes de *A. thaliana* contendo o gene *OsWRKY08* demonstraram-se tolerantes ao estresse osmótico por manitol, através do aumento do número de raízes laterais e indução do crescimento de raízes principais durante a germinação das sementes (SONG; JING; YU, 2009). Ainda em *A. thaliana*, sob altas temperaturas, os genes *WRKY25, WRKY26, WRKY33* e *WRKY39* foram diferencialmente expressos, em que Li et al. (2011) sugerem uma regulação cruzada entre os genes WRKY25, WRKY26 e WRKY33 refletindo na interação sinergética entre estes três genes durante o estresse por altas temperaturas. Outro membro da família WRKY, o gene *OsWRKY74* está envolvido nas respostas de tolerância contra o frio em arroz (RAMAMOORTHY et al., 2008). Os FT *WRKY18, WRKY40* e *WRKY60* de *A. thaliana* participam de rotas-chave na resposta a estresses bióticos e abióticos, formando uma rede reguladora altamente interativa que orquestra a expressão de genes de defesa em plantas (CHEN et al., 2010).

A família AP2/EREBP inclui um grande grupo de FT específicos de plantas e é caracterizada pela presença do domínio de ligação *APETALA2/ethylene responsive factor* (AP2/ERF) altamente conservado, constituído por 58 a 60 aminoácidos. Esta classe de FT interage diretamente com o elemento GCC-box e/ou *Dehydration-responsive element/C-repeat element* (DRE/CRT) *cis*-elemento ativado na região promotora (SAKUMA et al., 2002; XU et al., 2008). A superfamília AP2/EREBP é dividida de acordo com o número e similaridade dos domínios AP2/ERF em quatro subfamílias principais: AP2, ERF, DREB e RAV (relatada por ABI3/VP1) (SAKUMA et al., 2002). Em uma análise de genômica foi identificado um vasto número de membros da família *AP2/ERFBP* em várias espécies, com 145 representantes em *A. thaliana* (RIECHMANN; MEYEROWITZ, 1998), 170 em arroz *(Oryza sativa)* (RASHID et al., 2012) 178 em sorgo *(Sorghum bicolor)* (SRIVASTAV et al., 2010), 200 em álamo *(Populus)* (ZHUANG et al., 2008), 291 em repolho chinês *(Brassica rapa subsp. Pekinensis)* (SONG; LI; HOU, 2013), 171 em milheto *(Pennisetum glaucum)* (LATA et al., 2014), 116 em bambu *(Bambusa vulgaris)* (WU et al., 2015) e 148 em soja *(Glycine max)* (ZHANG et al., 2008).

As subfamílias ERF e DREB/C-repeat binding factor (CBF) foram amplamente estudadas devido ao envolvimento em respostas a estresses bióticos e abióticos em plantas (NAKASHIMA; KAZUKO; KAZUO, 2014; RASHID et al., 2012; SHARONI et al., 2011). Os fatores de transcrição de DREB/CBF desempenham um papel importante na regulação dos genes, respondendo a vários estresses, como frio, seca, alta salinidade e calor, em diferentes

espécies (KIDOKORO et al., 2015; NAKASHIMA; ITO; YAMAGUCHI-SHINOZAKI, 2009; ZONG et al., 2016). O *DREB1A/CBF e DREB2*, com 11 e 6 genes identificados em arroz *(Oryza sativa)* respectivamente, atuam, nessa mesa ordem, em resposta ao estresse por frio e déficit hídrico. Ambos atuam em rotas ABA-independente (NAKASHIMA; KAZUKO; KAZUO, 2014; SRIVASTAV et al., 2010). Os genes *OsDREB1A, OsDREB1B, OsDREB1C, e OsDREB1D* de arroz *(Oryza sativa)*, foram caracterizados por Dubouzet et al. (2003). Os FTs *DREB1/CBF*, interagem especificamente com DRE/CRT e regulam a expressão de importantes genes estresse responsivos. Diferentes estudos confirmam que a expressão de genes *DREB1/CBF* é diferencialmente regulada por outros genes como *ICE1, High expression of osmotically responsive genes 1 (HOS1), MYB15, SAP e Miz 1 (SIZ1), Phytochrome-interacting factor* (*PIF7), (Calmodulin-binding transcription activator 3 (CAMTA3)* (DONG; FARRÉ; THOMASHOW, 2011; QIN; SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2011).

A superexpressão do gene *DREB1A/CBF* induziu expressão de inúmeros genes estresse-responsivos proporcionando melhor tolerância em relação ao déficit hídrico, incluindo tomate *(Solanum lycopersicum)*, crisântemo *(Chrysanthemum)*, batata *(Solanum tuberosum)* (IWAKI et al., 2013), soja *(Glycine max)* (DE PAIVA ROLLA et al., 2014), arroz *(Oryza sativa)* (DATTA et al., 2012; NAKASHIMA; ITO; YAMAGUCHI-SHINOZAKI, 2014; PAUL et al., 2015), tabaco *(Nicotiana tabacum)* (PHUONG et al., 2015), cana-de-açúcar *(Saccharum officinarum)* (AUGUSTINE et al., 2015), amendoim *(Arachis hypogaea)* (BHATNAGAR-MATHUR et al., 2014; TIWARI et al., 2015) e trigo *(Triticum)* (SHAVRUKOV et al., 2016). O gene *OsDREB1F* é reportado como induzido por déficit hídrico e ABA exógeno (WANG et al., 2008). A superexpressão do gene *OsDREB1G* também é conhecida por promover tolerância ao déficit hídrico (CHEN et al., 2008).

Mais recentemente, plantas de tabaco *(Nicotiana tabacum)* transgênico com elevado nível de tolerância à seca e salinidade foi obtido através da superexpressão do gene *SsCBF4* da suculenta *Suaeda salsa* (ZHANG et al., 2015). A expressão estresse-induzida e constitutiva dos genes *OsDREB2A e OsDREB2B* em plantas de arroz transgênicas tem resultado no aumento da tolerância ao estresse osmótico (CUI et al., 2011; MIZOI et al., 2013). Plantas de *A. thaliana* superexpressando *OsDREB2B* exibiram ativação ampliada dos genes alvo do DREB2A e consequente aumento na tolerância à seca (MATSUKURA et al., 2010). O mesmo foi observado em soja transgênica, em que a superexpressão do *DREB2A* ampliou a condição de tolerância a este mesmo estresse (ENGELS et al., 2013).

Sob condições normais, o *DREB2A* é controlado por dois mecanismos: (I) *Growth regulating factor 7* (GRF7), o qual inibe a expressão do *DREB2A* ligando-se à uma curta porção da região promotora (SINGH; LAXMI, 2015) e (II) degradação direcionada pela proteólise mediada por proteossoma 26S auxiliado pelas proteínas *DREB2A- interacting protein 1 and 2* (DRIP1) e (DRIP2) (MORIMOTO et al., 2013; SINGH; LAXMI, 2015) (Figura 1). Assim, a expressão estresse-induzida do *DREB2A* sob condições de déficit hídrico, atua como um mecanismo de minimizar a perda ineficiente de energia (SINGH; LAXMI, 2015). Além dos membros *DREB1A e DREB2*, vários outros FTs do tipo AP2/EREBP também foram caracterizados em arroz com o objetivo de melhorar a tolerância ao estresse abiótico. Entre eles, a superexpressão do gene *HARDY* (um subgrupo da família A-4 AP2/EREBP) obtido a partir de plantas de *A. thaliana* em *Trifolium* levou à tolerância ao estresse osmótico (ABOGADALLAH et al., 2011). *OsDERF1* foi estabelecido como um modulador negativo no ajuste osmótico interagindo diretamente com uma região GCC-box localizada no promotor do gene *OsERF3* (WAN et al., 2011). Além disso, o *OsERF3* contém um domínio *Ethylene-responsive element binding factor-associated amphiphilic repression* (EAR), que reprime a produção de etileno ao nível severo de estresse por déficit hídrico (ZHANG et al., 2013). No entanto, Joo et al. (2013) descreveram o gene *OsERF4a* como um regulador positivo do crescimento das plantas e tolerância à seca em arroz. Além disso, a expressão ectópica do gene *JERF3* em tomate *(Solanum lycopersicum),* libera açúcares solúveis e prolina com aumento da tolerância à seca nesta mesma espécie (ZHANG et al., 2010).

2.6 Análise de transcriptoma e a ferramenta RNA-seq

Algumas tecnologias vêm trazendo avanços nos estudos de transcriptoma nos últimos anos. O sequenciamento do RNA (RNA-seq), por exemplo, possui uma vasta aplicação, porém, o protocolo de análise nem sempre é o mesmo para todos os casos (CONESA et al., 2016). O RNA-seq se resume na combinação de uma metodologia de sequenciamento de alto rendimento com métodos computacionais para capturar e quantificar transcrições presentes em um extrato de RNA (OZSOLAK; MILOS, 2011). As sequências de nucleotídeos geradas são tipicamente cerca de 100 pb de comprimento, mas podem variar de 30 pb a mais de 10.000 pb, dependendo do método de sequenciamento utilizado. O RNA-Seq se dispõe da profunda amostragem do transcriptoma, com muitos fragmentos leves das sequências alvo, para permitir a reconstrução computacional do RNA original, alinhando as sequências (reads)

geradas com um genoma de referência, para assim, fazer a montagem final dos transcritos (WANG; SNYDER, 2009). O intervalo dinâmico típico de cinco ordens de grandeza para RNA-seq é uma das vantagens-chave em relação ao transcriptoma oriundo de metodologias como a de *Massive Parallel Signature Sequencing (MPSS), DNA microarray, Serial Analysis of gene expression. (SAGE), Differential display (RT-PCR), EST sequencing* (LOWE et al., 2017).

Adicionalmente, a quantidade de RNA utilizado é muito menor para aplicação do RNA-seq (concentração em nanograma), enquanto em outras metodologias, há uma necessidade maior de moléculas de RNA, ficando limitado à microgramas. Deste modo, o RNA-seq permitiu um estudo mais refinado das estruturas celulares, chegando a análises em uma única célula quando combinado com amplificação linear do cDNA (HASHIMSHONY et al., 2012). Supostamente, não há limite de quantificação em RNA-seq para transcritos longos e que são supexpressos em relação ao material comparado, porém, fragmentos pequenos, menores do que 100 pb, em regiões não repetitivas, a quantificação pode ser comprometida (OZSOLAK; MILOS, 2011).

O RNA-seq pode ser uma ferramenta na identificação de genes que compõe um genoma, além da possibilidade de discriminar aqueles genes ativos/reprimidos nas condições ambientais e metabólicas exatas em que a planta se encontra no momento da coleta de tecidos para análise. A contagem de sequências (reads) pode ser usada para presumir o nível de expressão gênica relativo (TACHIBANA, 2015). Há um avanço constante na metodologia do RNA-seq, principalmente pelo desenvolvimento e inovação das tecnologias de sequenciamento de DNA, aumentando o rendimento, a precisão e o comprimento das sequências (TACHIBANA, 2015). As estratégias transcriptômicas têm ampla aplicação em diversas áreas de pesquisa biomédica, incluindo diagnóstico e perfil de doenças (OZSOLAK; MILOS, 2011). Estudos também foram realizados em soja, em que foi possível obter o perfil de expressão de uma família de FT (GmAP2/EREB-like) em plantas sob condições de déficit hídrico (MARCOLINO-GOMES et al., 2013), análise do transcriptoma ao longo das oscilações diárias da modulação gênica em plantas sob déficit hídrico (RODRIGUES et al., 2015), e ainda em Chen et al. (2016) em que identificaram e compararam genes diferencialmente expressos em folhas sob condições de déficit hídrico e encharcamento. Mais recentemente, Janiak et al. (2018), observaram que a tolerância à seca pode ser atribuída aos padrões de expressão que existem antes da ocorrência de estresse em cevada *(Hordeum vulgare).*

Em suas variadas abordagens, o RNA-seq permitiu a identificação, em grande escala, dos sítios de início da transcrição e descoberta do uso de promotores alternativos, além das novas alterações de *splicing* (COSTA et al., 2013). O RNA-seq também pode identificar polimorfismos de nucleotídeos únicos associados à doença (SNP), expressão específica do alelo e fusão de genes, contribuindo para a compreensão de variantes causais da doença (KHURANA et al., 2016). Os retrotransposons são elementos transponíveis que proliferam nos genomas eucarióticos através de um processo que envolve a transcrição reversa, o RNA-seq, neste caso, pode fornecer informações sobre a transcrição de retrotransposons endógenos que podem influenciar a transcrição de genes vizinhos por vários mecanismos epigenéticos que levam à doença alvo (SLOTKIN; MARTIENSSEN, 2007).

Análises transcriptômicas permitem a identificação de genes e rotas que respondem e enfrentam estresses ambientais bióticos e abióticos, viabilizando o reconhecimento de novas redes transicionais em sistemas complexos. Um exemplo dessa aplicabilidade foi demonstrado na análise comparativa de uma variedade de grão-de-bico *(Cicer arietinum)* em diferentes estádios de desenvolvimento, em que foram observados perfis transcricionais distintos associados aos estresses de seca e salinidade, incluindo a identificação do papel das isoformas dos transcritos do AP2/EREBP (GARG et al., 2016). A Transcriptômica revolucionou a compreensão de como os genes são modulados/expressos, oscilando e adequando-se às condições ambientais. Ao longo das últimas três décadas, as novas tecnologias redefiniram e reinventaram o que é possível investigar. A integração com outras plataformas de informação está gerando uma visão cada vez mais interativas das complexidades celulares (LOWE et al., 2017).

REFERÊNCIAS

ABOGADALLAH, G. M. et al. Overexpression of HARDY, an AP2/ERF gene from Arabidopsis, improves drought and salt tolerance by reducing transpiration and sodium uptake in transgenic *Trifolium alexandrinum* L. **Planta**, Berlin, v. 233, p. 1265-1276, 2011.

AGARWAL, M. et al. A R2R3 type MYB transcription factor is involved in the cold regulation of CBF genes and in acquired freezing tolerance. **Journal of Biological Chemistry**, Baltimore, v. 281, p. 37636-37645, 2006.

AGRITEMPO. **Aquecimento global e a produção agrícola do Brasil**. Disponível em: <http://www.agritempo.gov.br/climaeagricultura/>. Acesso em: 10 fev. 2018.

AUGUSTINE, S. M. et al. Overexpression of EaDREB2 and pyramiding of EaDREB2 with the pea DNA helicase gene (PDH45) enhance drought and salinity tolerance in sugarcane (*Saccharum* spp. hybrid). **Plant Cell Reports**, Berlin, v. 34, p. 247-263, 2015.

BANDURSKA, H. Does proline accumulated in leaves of water deficit stressed barley plants confine cell membrane injury?: I., free proline accumulation and membrane injury index in drought and osmotically stressed plants. **Acta Physiologiae Plantarum**, Copenhagen, v. 22, p. 409-415, 2000.

BANDURSKA, H.; GO´RNY, A. G.; ZIELEZIN´SKA, M. Effects of water deficit on relative water content, proline accumulation and injury of cell membranes in leaves of old and modern cultivars of winter wheat. **Zeszyty Problemowe Poste¸po´w Nauk Rolniczych**, Warszawa, v. 524, p. 115-126, 2008.

BENCKE-MALATO, M. et al. Genome-wide annotation of the soybean WRKY family and functional characterization of genes involved in response to *Phakopsora pachyrhizi* infection. **BMC Plant Biology**, London, v. 14, p. 236, 2014.

BHATNAGAR-MATHUR, P. et al. Transgenic peanut overexpressing the DREB1A transcription factor has higher yields under drought stress. **Molecular Breeding**, Dordrecht, v. 33, p. 327-340, 2014.

BLUM, A. Crop responses to drought and the interpretation of adaptation. In: BELHASSEN, E. (Ed.). **Drought tolerance in higher plants:** genetical, physiological and molecular biological analysis. Dordrecht: Kluwer Academic, 1997. p. 57-70.

BOHNERT, H. J. et al. A genomics approach towards salt stress tolerance. **Plant Physiology and Biochemistry**, Paris, v. 39, p. 1-10, 2001.

BRAY, E. A. Molecular responses to water deficit'. **Plant Physiology**, Riverside, v. 103, p. 1035-1040, 1993.

CASTRO, P. H. et al. SUMO, a Heavyweight player in plant abiotic stress responses. **Cellular and Molecular Life Sciences**, Basel, v. 69, p. 3269-3283, 2012.

CAVATTE, P. C. et al. A fisiologia dos estresses abióticos. In: FRITSCHE-NETO, R.; BORÉM, A. (Ed.). **Melhoramento de plantas para condições de estresses abióticos**. Visconde do Rio Branco: Suprema, 2011. p. 39-79.

CENTRO DE ESTUDOS AVANÇADOS EM ECONOMIA APLICADA. **Soja/PERSPEC 2018:** oferta se estabiliza, mas alto estoque deve evitar forte oscilação de preço. Disponível em: <https://www.cepea.esalq.usp.br/br/soja-perspec-2018-oferta-se-estabiliza-mas-alto-estoque-deve-evitar-forte-oscilacao-de-preco.aspx>. Acesso em: 10 fev. 2018.

CENTRO DE ESTUDOS AVANÇADOS EM ECONOMIA APLICADA; CONFEDERAÇÃO DA AGRICULTURA E PECUÁRIA DO BRASIL. **PIB de cadeias agropecuárias**. 2017. Disponível em: <https://www.cepea.esalq.usp.br/br/pib-de-cadeias-agropecuarias.aspx>. Acesso em: 10 fev. 2018.

CHAVES, M. M. Effects of water stress on carbon assimilation. **Journal of Experimental Botany**, Lancaster, v. 42, p. 1-16, 1991.

CHAVES, M. M.; OLIVEIRA, M. M. Mechanisms underlying plant resilience to water defecits: prospects for water-saving agriculture. **Journal of Experimental Botany**, Lancaster, v. 55, n. 407, p. 2365-2384, 2004.

CHEN, H. et al. Roles of Arabidopsis WRKY18, WRKY40 and WRKY60 transcription factors in plant responses to absCisic acid and abiotic stress. **BMC Plant Biology**, London, v. 10, p. 281, 2010.

CHEN, H.; JIANG, J. G. Osmotic adjustment and plant adaptation to environmental changes related to drought and salinity. **Environmental Review**, Denver, v. 18, p. 309-319, 2010.

CHEN, J. Q. et al. Overexpression of OsDREB genes lead to enhanced drought tolerance in rice. **Biotechnology Letters**, Dordrecht, v. 30, p. 2191-2198, 2008.

CHEN, W. et al. Identification and comparative analysis of differential gene expression in soybean leaf tissue under drought and flooding stress revealed by RNA-Seq. **Frontiers in Plant Science**, Lausanne, v. 7, p. 1044, 2016.

CHINNUSAMY, V.; ZHU, J.; ZHU, J. K. Cold stress regulation of gene expression in plants. **Trends in Plant Science**, Oxford, v. 12, p. 444-451, 2007.

_____. Gene regulation during cold acclimation in plants. **Physiology Plant**, Bethesda, v. 126, p. 52-61, 2006.

CIECHANOVER, A. et al. ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protei degradation. **Proceedings of the National Academy of Sciences USA**, Washington, DC, v. 77, n. 3, p. 1365-1368, 1980.

COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da safra brasileira de grãos, v. 1:** safra 2013/14, n. 12: décimo segundo levantamento, set. Brasília, DF, 2014.

_____. **Acompanhamento da safra brasileira de grãos v. 4:** safra 2017/18, n. 2: segundo levantamento, novembro. Brasília, DF, 2017a.

_____. **A produtividade da soja:** análise e perspectivas. Brasília, DF, 2017b. (Compêndio de estudos Conab, 10).

_____. **Séries históricas de produção de grãos**. Disponível em: <http://www.conab.gov.br/conteudos.php?a=1252&Pagina_objcmsconteudos=1#A_objcmsconteudos>. Acesso em: 10 fev. 2018.

CONESA, A. et al. A survey of best practices for RNA-seq data analysis. **Genome Biology**, London, v. 17, p. 13, Jan. 2016.

CONFEDERAÇÃO DA AGRICULTURA E PECUÁRIA DO BRASIL. **El Niño impõe cautela ao início da safra de soja**. Disponível em: <http://www.cnabrasil.org.br/noticias/el-nino-impoe-cautela-ao-inicio-da-safra-de-soja>. Acesso em: 10 dez. 2017.

CORNIC, G.; BRIANTAIS, J. M. Partitioning of electron flow between CO2 and O2 reduction in a C3 leaf (*Phaseolus vulgaris* L.) at different CO2 concentration and during drought stress. **Planta**, Berlin, v. 183, p. 178-184, 1991.

CORNIC, G.; MASSACCI, A. Leaf photosyntesis under dorught stress. In: BAKER, N. R. (Ed.). **Photosyntesis and the environment**. The Netherlands: Kluwer Academic, 1996. p. 347-366.

COSTA, M. da; HUANG, B. R. Osmotic adjustment associated with variation in bentgrass tolerance to drought stress. **Journal of the American Society for Horticultural Science**, Alexandria, v. 131, n. 3, p. 338-344, 2006.

COSTA, V. et al. RNA-Seq and human complex diseases: recent accomplishments and future perspectives. **European Journal of Human Genetics**, London, v. 21, p. 134-142, 2013.

CUI, M. et al. Induced over-expression of the transcription factor OsDREB2A improves drought tolerance in rice. **Plant Physiology and Biochemistry**, Paris, v. 49, p. 1384-1391, 2011.

DALL'AGNOL, A. **Soja lidera pauta de exportação**. Disponível em: <http://blogs.canalrural.com.br/embrapasoja/2017/05/12/soja-lidera-pauta-de-exportacao/>. Acesso em: 10 fev. 2018.

DATTA, K. et al. Overexpression of Arabidopsis and rice stress genes inducible transcription factor confers drought and salinity tolerance to rice. **Plant Biotechnology Journal**, Oxford, v. 10, p. 579-586, 2012.

DE PAIVA ROLLA, A. A. et al. Phenotyping soybean plants transformed with rd29A: AtDREB1A for drought tolerance in the greenhouse and field. **Transgenic Research**, London, v. 23, p. 75-87, 2014.

DEPPMANN, C. D.; ALVANIA, R. S.; TAPAROWSKY, E. J. Cross-species annotation of basic leucine zipper factor interactions: Insight into the evolution of closed interaction networks. **Molecular Biology and Evolution**, Chicago, v. 23, p. 1480-1492, 2006.

DESIKAN, R. et al. ABA, hydrogen peroxide and nitric oxide signaling in stomatalguard cell. **Journal of Experimental Botany**, Oxford, v. 55, p. 205-212, 2004.

DONG, M. A.; FARRÉ, E. M.; THOMASHOW, M. F. Circadian clock-associated 1 and late esevereated hypocotyl regulate expression of the C-repeat binding factor (CBF) pathway in Arabidopsis. **Proceedings of the National Academy of Science USA**, Washington, DC, v. 108, p. 7241-7246, 2011.

DUBOUZET, J. G. et al. OsDREB genes in rice, Oryza sativa L.; encode transcription activators that function in drought-, high-salt-, and cold-responsive gene expression. **The Plant Journal**, Oxford, v. 33, p. 751-763, 2003.

ENGELS, C. et al. Introduction of the rd29A: AtDREB2A CA gene into soybean (*Glycine max* L. Merril) and its molecular characterization in leaves and roots during dehydration. **Genetics and Molecular Biology**, Ribeirão Preto, v. 36, p. 556-565, 2013.

EULGEM, T. et al. The WRKY superfamily of plant transcription factors. **Trends in Plant Science**, Oxford, v. 5, p. 199-206, 2000.

FRANCO-ZORRILLA, J. M. et al. DNA-binding specificities of plant transcription factors and their potential to define target genes. **Proceedings of the National Academy of Science USA**, Washington, DC, v. 111, p. 2367-2372, 2014.

FUGANTI-PAGLIARINI, R. et al. Characterization of soybean genetically modified for drought tolerance in field conditions. **Frontiers in Plant Science**, Lausanne, v. 8, p. 448, 2017.

GARG, R. et al. Transcriptome analyses reveal genotype- and developmental stage-specific molecular responses to drought and salinity stresses in chickpea. **Scientific Reports**, London, v. 6, p. 19228, 2016.

GOLLDACK, D.; LÜKING, I.; YANG, O. Plant tolerance to drought and salinity: stress regulating transcription factors and their functional significance in the cellular transcriptional network. **Plant Cell Reports**, Berlin, v. 30, p. 1383-1391, 2011.

HANSON, A. D.; HITZ, W. D. Metabolic responses of mesophytes to plant water deficits. **Annual Review of Plant Physiology**, Palo Alto, v. 33, p. 163-203, 1982.

HASHIMSHONY, T. et al. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. **Cell Reports**, New York, v. 2, p. 666-673, 2012.

HIRAKURI, M. H. **Impactos econômicos de estresses na produção de soja da safra 2015/16**. Londrina: EMBRAPA Soja, 2016. 7 p. (EMBRAPA Soja. Circular técnica, 125).

HIRAKURI, M. H.; LAZZAROTTO, J. J. **O agronegócio da soja nos contextos mundial e brasileiro**. Londrina: EMBRAPA Soja, 2014. 70 p. (EMBRAPA Soja. Documentos, 349).

HIRAYAMA, T.; SHINOZAKI, K. Research on plant abiotic stress responses in the post-genome era: past, present and future. **Plant Journal**, Oxford, v. 61, p. 1041-1052, 2010.

HSIAO, T. C. Plant response to water stress. **Annual Review of Plant Physiology**, Palo Alto, v. 24, p. 519-770, 1973.

HSIEH, T. H. et al. Heterology expression of the Arabidopsis C-repeat/dehydration response element binding factor 1 gene confers elevated tolerance to chilling and oxidative stresses in transgenic tomato. **Plant Physiology**, Bethesda, v. 129, p. 1086-1094, 2002.

HUANG, G. T. et al. Signal transduction during cold, salt, and drought stresses in plants. **Molecular Biological Reports**, London, v. 39, p. 969-987, 2012.

INGRAM, J.; BARTELS, D. The molecular basis of dehydration tolerance in plants. **Annual Review of Plant Physiology and Plant Molecular Biology**, Palo Alto, v. 47, p. 377-403, 1996.

IWAKI, T. et al. Metabolic profiling of transgenic potato tubers expressing Arabidopsis dehydration response element-binding protein 1A (DREB1A). **Journal of Agriculture and Food Chemistry**, Easton, v. 61, p. 893-900, 2013.

JANIAK, A. et al. No time to waste: transcriptome study reveals that drought tolerance in barley may be attributed to stressed-like expression patterns that exist before the occurrence of stress. **Frontiers in Plant Science**, Lausanne, v. 8, p. 2212, 2018.

JAVADI, T.; ARZANI, K.; EBRAHIMZADEH, H. Study of proline, soluble sugar, and chlorophyll a and b changes in nine Asian and one European pear cultivar under drought stress. **Acta Horticulturae**, The Hague, v. 769, p. 241-246, 2008.

JIN, J. et al. Plant TFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. **Nucleic Acids Research**, Oxford, v. 2014, p. 1182-1187, 2014.

JOO, J. et al. A transcriptional repressor of the ERF family confers drought tolerance to rice and regulates genes preferentially located on chromosome 11. **Planta**, Berlin, v. 238, p. 155-170, 2013.

JOSHI, R. et al. Salt stress signalling pathways: specificity and crosstalk. In: WANI, S. H.; HOSSAIN, M. A. (Ed.). **Managing salinity tolerance in plants:** molecular and genomic perspectives. Boca Raton: CRC, 2016a. p. 51-78.

_____. Transcription factors and plants response to drought stress: current understanding and future directions. **Frontiers in Plant Science**, Lausanne, v. 7, p. 1029, 2016b.

KASUGA, M. et al. A combination of the Arabidopsis DREB1A gene and stress-inducible RD29A promoter improved drought-and low temperature stress tolerance in tobacco by gene transfer. **Plant and Cell Physiology**, Oxford, v. 45, p. 346-350, 2004.

KAZUO, N.; KAZUKO, Y. S.; KAZUO, S. The transcriptional regulatory network in the drought response and its crosstalk in abiotic stress responses including drought, cold, and heat. **Frontiers in Plant Science**, Lausanne, v. 5, p. 1-170, 2014.

KHAN, Z. H. et al. Transcriptional regulatory network of cis-regulatory elements (Cres) and transcription factors (Tfs) in plants during abiotic stress. **International Journal of Plant Biology & Research**, San Diego, v. 5, n. 2, p. 1064, 2017.

KHAZAEI, H. et al. The FIGS (focused identification of germplasm strategy) approach identifies traits related to drought adaptation in *Vicia faba* genetic resources. **PLoS ONE**, San Francisco, v. 8, p. e63107, 2013.

KHURANA, E. et al. Role of non-coding sequence variants in cancer. **Nature Reviews Genetics**, London, v. 17, p. 93-108, 2016.

KIDOKORO, S. et al. Soybean DREB1/CBF-type transcription factors function in heat and drought as well as cold stress-responsive gene expression. **Plant Journal**, Oxford, v. 81, p. 505-518, 2015.

KRAMER, P. J.; BOYER, J. S. **Water relations of plants and soils**. New York: Academic, 1995.

KUME, S. et al. Differential and coordinated expression of Cbf and Cor/Lea genes during severe-term cold acclimation in two wheat cultivars showing distincti levels of freezing tolerance. **Genes & Genetics System**, Shizuoka, v. 80, p. 185-197, 2005.

LATA, C. et al. Genome-wide investigation and expression profiling of AP2/ERF transcription factor superfamily in foxtail millet (*Setaria italica* L.). **PLoS ONE**, San Francisco, v. 9, p. e113092, 2014.

LAWLOR, D. W. Genetic engineering to improve plant performance under drought: physiological evaluation of achievements, limitations, and possibilities. **Journal of Experimental Botany**, Oxford, v. 63, p. 695-709, 2013.

LEUNG, J.; MERLOT, S.; GIRAUDAT, J. The Arabidopsis ABSCISIC ACID INSENSITIVE2 (ABI2) and ABI1 genes encode homologous protein phosphatases 2c involved in abiscisic acid signal transduction. **The Plant Cell**, Rockville, v. 9, p. 759-771, 1997.

LI, S. et al. Arabidopsis thaliana WRKY25, WRKY26, and WRKY33 coordinate induction of plant thermotolerance. **Planta**, Berlin, v. 233, p. 1237-1252, 2011.

LIAO, Y. et al. Role of soybean GmbZip132 under abscisic acid and salt stresses. **International Journal of Plant Biology**, Pavia, v. 50, p. 221-230, 2008a.

_____. Soybean GmMYB76, GmMYB92, and GmMYB177 genes confer stress tolerance in transgenic Arabidopsis plants. **Cell Research**, London, v. 18, p. 1047-1060, 2008b.

LOWE, R. et al. Transcriptomics technologies. **PLoS Computational Biology**, San Francisco, v. 13, n. 5, p. e1005457, 2017.

MANSFIELD, T. A.; DAVIES, W. J. Stomata and stomatal mechanisms. In: PALEG, L. G.; ASPINALL, D. (Ed.). **The physiology and biochemistry of drought resistance in plants**. New York: Academic, 1981. p. 315-346.

MARCOLINO-GOMES, J. et al. Expression patterns of GmAP2/EREB-Like transcription factors involved in soybean responses to water deficit. **PLoS ONE**, San Francisco, v. 8, n. 5, p. e62294, 2013.

MATSUKURA, S. et al. Comprehensive analysis of rice DREB2-type genes that encode transcription factors involved in the expression of abiotic stress-responsive genes. **Molecular Genetics and Genomics**, Berlin, v. 283, p. 185-196, 2010.

MATYSIK, J. et al. Molecular mechanisms of quenching of reactive oxygen species by proline under stress in plants. **Current Science**, Columbus, v. 82, p. 525-532, 2002.

MEULMEESTER, E.; MELCHIOR, F. Cell biology: SUMO. **Nature**, London, v. 452, n. 7188, p. 709-711, 2008.

MILLER, G. et al. Reactive oxygen species homeostasis and signalling during drought and salinity stresses. **Plant, Cell & Environment**, Oxford, v. 33, p. 453-467, 2010.

MIZOI, J. et al. GmDREB2A: 2, a canonical DEHYDRATION-RESPONSIVE ELEMENT-BINDING PROTEIN2-type transcription factor in soybean, is posttranslationally regulated and mediates dehydration-responsive element-dependent gene expression. **Plant Physiology**, Bethesda, v. 161, p. 346-361, 2013.

MORIMOTO, K. et al. Stabilization of Arabidopsis DREB2A is required but not sufficient for the induction of target genes under conditions of stress. **PLoS ONE**, San Francisco, v. 8, p. e80457, 2013.

NAKASHIMA, K. et al. Functional analysis of a NAC-type transcription factor OsNAC6 involved in abiotic and biotic stress-responsive gene expression in rice. **Plant Journal**, Oxford, v. 51, p. 617-630, 2007.

NAKASHIMA, K.; ITO, Y.; YAMAGUCHI-SHINOZAKI, K. Transcriptional regulatory networks in response to abiotic stresses in Arabidopsis and grasses. **Plant Physiology**, Bethesda, v. 149, p. 88-95, 2009.

OH, S. J. et al. Arabidopsis CBF3/DREB1A and ABF3 in transgenic rice increased tolerance to abiotic stress without stunting growth. **Plant Physiology**, Bethesda, v. 138, p. 341-351, 2005.

OYA, T. et al. Drought tolerance characteristics of Brazilian Soybean cultivars: evaluation and characterization of drought tolerance of various Brazilian soybean cultivars in the field. **Plant Production Science**, London, v. 7, n. 2, p. 129-137, 2004.

OZSOLAK, F.; MILOS, P. M. RNA sequencing: advances, challenges and opportunities. **Nature Reviews Genetics**, London, v. 12, p. 87-98, 2011.

OZTURK, L.; DEMIR, Y. In vivo and in vitro protective role of proline. **Plant Growth Regulation**, Dordrecht, v. 38, p. 259-264, 2002.

PAUL, S. et al. Dissecting root proteome of transgenic rice cultivars unravels metabolic alterations and accumulation of novel stress responsive proteins under drought stress. **Plant Science**, Shannon, v. 234, p. 133-143, 2015.

PEREIRA, N. **Cautela é a palavra de ordem na safra 2017/2018**. Goiânia: FAEG, 2017. Disponível em: <http://sistemafaeg.com.br/noticias/16416-cautela-e-a-palavra-de-ordem-na-safra-2017-2018>. Acesso em: 10 jan. 2018.

PHUONG, N. D. et al. Identification and characterization of a stress-inducible gene OsNLI-IF enhancing drought tolerance in transgenic tobacco. **Current Science**, Columbus, v. 109, p. 541-551, 2015.

QIN, F. et al. Cloning and functional analysis of a novel DREB1/CBF transcription factor involved in coldresponsive gene expression in Zea mays L. **Plant and Cell Physiology**, Oxford, v. 45, p. 1042-1052, 2004.

QIN, F.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Achievements and challenges in understanding plant abiotic stress responses and tolerance. **Plant and Cell Physiology**, Oxford, v. 52, p. 1569-1582, 2011.

RAMAMOORTHY, R. et al. A comprehensive transcriptional profiling of the WRKY gene family in rice under various abiotic and phytohormone treatments. **Plant Cell Physiology**, Oxford, v. 49, n. 6, p. 865-879, 2008.

RASHID, M. et al. AP2/ERF transcription factor in rice: genome-wide canvas and syntenic relationships between monocots and eudicots. **Evolutionary Bioinformatics Online**, London, v. 8, p. 321-355, 2012.

RAYKO, E. et al. Transcription factor families in ferred from genome sequences of photosynthetics tramenopiles. **New Phytologist**, Cambridge, v. 188, p. 52-66, 2010.

RIECHMANN, J. L.; MEYEROWITZ, E. M. The AP2/EREBP family of plant transcription factors. **Biological Chemistry**, Baltimore, v. 379, p. 633-646, 1998.

RODRIGUES, F. A. et al. Daytime soybean transcriptome fluctuations during water deficit stress. **BMC Genomics**, London, v. 16, n. 1, p. 505, 2015.

_____. Subtractive libraries for prospecting differentially expressed genes in the soybean under water deficit. **Genetics and Molecular Biology**, Ribeirão Preto, v. 35, p. 304-314, 2012.

ROJAS, A. et al. Physiological response of gmelina (Gmelina arborea Roxb.) to hydric conditions of the colombian Caribbean. **Agronomía Colombiana**, Bogota, v. 30, n. 1, p. 52-58, 2012.

RUGGIERO, B. et al. Uncoupling the effects of ABA on plant growth and water relations: analysis of sto1/nced3, BA deficient salt stress tolerant mutant in Arabidopsis thaliana. **Plant Physiology**, Bethesda, v. 136, p. 3134-3147, 2004.

RUSHTON, P. J. et al. WRKY transcription factors. **Trends in Plant Science**, Oxford, v. 15, p. 247-258, 2010.

SAKUMA, Y. et al. DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydrationand cold-inducible gene expression. **Biochemical and Biophysical Research Communications**, Orlando, v. 290, p. 998-1009, 2002.

_____. Dual function of an Arabidopsis transcription factor DREB2A in waterstress-responsive and heat-stress-responsive gene expression. **Proceedings of the National Academy of Sciences USA**, Washington, DC, v. 103, p. 18828-18833, 2006.

SALISBURY, F. B.; ROSS, C. W. **Fisiologia das plantas**. Rio de Janeiro: Cengage Learning, 2013.

SARACCO, S. A. et al. Genetic analysis of SUMOylation in Arabidopsis: conjugation of SUMO1 and SUMO2 to nuclear proteis is essential. **Plant Physiology**, Bethesda, v. 145, p. 119-134, 2007.

SEKI, M. et al. Monitoring the expression pattern of 1300 Arabidopsis genes under drought and cold stresses by using a full-length cDNA microarray. **Plant Cell**, Rockville, v. 13, p. 61-72, 2001.

SHARONI, A. M. et al. Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. **Plant and Cell Physiology**, Oxford, v. 52, p. 344-360, 2011.

SHAVRUKOV, Y. et al. The TaDREB3 transgene transferred by conventional crossings to different genetic backgrounds of bread wheat improves drought tolerance. **Plant Biotechnology Journal**, Oxford, v. 14, p. 313-322, 2016.

SINGH, D.; LAXMI, A. Transcriptional regulation of drought response: a tortuous network of transcriptional factors. **Frontiers in Plant Science**, Lausanne, v. 6, p. 895, 2015.

SLOTKIN, R. K.; MARTIENSSEN, R. Transposable elements and the epigenetic regulation of the genome. **Nature Reviews Genetics**, London, v. 8, p. 272-285, 2007.

SONG, X.; LI, Y.; HOU, X. Genome-wide analysis of the AP2/ERF transcription factor superfamily in Chinese cabbage (Brassica rapa ssp. pekinensis). **BMC Genomics**, London, 14, p. 573, 2013.

SONG, Y.; JING, S.; YU, D. Overexpression of the stress-induced OsWRKY08 improves osmotic stress tolerance in Arabidopsis. **Chinese Science Bulletin**, Beijing, v. 54, p. 4671-4678, 2009.

SRIVASTAV, A. et al. Over-represented promoter motifs in abiotic stress-induced DREB genes of rice and sorghum and their probable role in regulation of gene expression. **Plant Signaling & Behaviour**, London, v. 5, p. 775-784, 2010.

STOCKINGER, E. J.; GILMOUR, S. J.; THOMASHOW, M. F. Arabidopsis thaliana CBF1 encodes an AP2 domain-containing transcription activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. **Proceedings of the National Academy of Sciences USA**, Washington, DC, v. 94, p. 1035-1040, 1997.

SUN, J. et al. The Brachypodium distachyon BdWRKY36 gene confers tolerance to drought stress in transgenic tobacco plants. **Plant Cell Reports**, Berlin, v. 34, n. 1, p. 23-35, Jan. 2015.

SZABADOS, L.; SAVOURE´, A. Proline: a multifunctional amino acid. **Trends in Plant Science**, Oxford, v. 15, p. 89-97, 2010.

TACHIBANA, C. Transcriptomics today: microarrays, RNA-seq, and more. **Science**, New York, v. 349, n. 6247, p. 544-546, July 2015.

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

_____. _____. 5. ed. Porto Alegre: Artmed, 2013. 918 p.

TIWARI, V. et al. Introgression of the SbASR-1 gene cloned from a halophyte salicornia brachiate enhances salinity and drought endurance in transgenic groundnut (Arachis hypogaea) and acts as a transcription factor. **PLoS ONE**, San Francisco, v. 10, p. e0131567, 2015.

TUTEJA, N.; SINGH, G. S.; TUTEJA, R. **Omics and plant abiotic stress tolerance**. New Delhi: Bentham & Biiks, 2011.

UNITED STATES DEPARTMENT OF AGRICULTURE. **World agricultural production**. Washington, DC, 2018. (Circular series WAP1-18).

VERSLUES, P. E.; ZHU, J. K. Before and beyond ABA: upstream sensing and internal signals that determine ABA accumulation and response under abiotic stress. **Biochemical Society Transactions**, London, v. 33, p. 375-379, 2005.

WAN, L. et al. Transcriptional activation of OsDERF1 in *OsERF3* and *OsAP2-39*Negatively modulates ethylene synthesis and drought tolerance in rice. **PLos ONE**, San Francisco, v. 6, n. 9, p. e25216, 2011.

WANG, Q. et al. Over expression of a rice OsDREB1F gene increases salt, drought, and low temperature tolerance in both Arabidopsis and rice. **Plant Molecular Biology**, Dordrecht, v. 67, p. 589-602, 2008.

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

WANI, S. H. et al. Engineering abiotic stress tolerance in plants: extricating regulatory gene complex. In: MALIK, C. P.; SANGHERA, G. S.; WANI, S. H. (Ed.). **Conventional and non-conventional interventions in crop improvement**. New Delhi: CABI, 2013. p. 1-19.

WU, H. et al. Genome-wide analysis of the AP2/ERF transcription factors family and the expression patterns of DREB genes in moso bamboo (Phyllostachys edulis). **PLoS ONE**, San Francisco, v. 10, p. e0126657, 2015.

XIONG, Y.; FEI, S. Z. Functional and phylogenetic analysis of a DREB/CBF-like gene in perennial ryegrass (Lolium perenne L.) **Planta**, Berlin, v. 224, p. 878-888, 2006.

XU, D. Q. et al. Overexpression of a TFIIIA-type zinc finger protein gene ZFP252 enhances drought and salt tolerance in rice (Oryza sativa L.). **FEBS Letters**, Amsterdam, v. 582, p. 1037-1043, 2008.

YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress. **Plant Cell**, Rockville, v. 6, p. 251-264, 1994.

YORDANOV, I.; VELIKOVA, V.; TSONEV, T. Plant response to drought, acclimation, and stress tolerance. **Photosynthetica**, Prague, v. 38, p. 171-186, 2000.

ZARKA, D. G. et al. Cold induction of Arabidopsis CBF genes involves multiple ICE (inducer of CBF expression) promoter elements and a cold regulatory circuit that is desensitized by low temperature. **Plant Physiology**, Bethesda, v. 133, n. 2, p. 910-918, 2003.

ZHANG, G. et al. Phylogeny, gene structures, and expression patterns of the ERF gene family in soybean (Glycine max L.). **Journal of Experimental Botany**, Oxford, v. 59, n. 15, p. 4095-4107, 2008.

ZHANG, H. et al. EAR motif mutation of rice OsERF3 alters the regulation of ethylene biosynthesis and drought tolerance. **Planta**, Berlin, v. 237, p. 1443-1451, 2013.

_____. Functional analyses of ethylene response factor JERF3 with the aim of improving tolerance to drought and osmotic stress in transgenic rice. **Transgenic Research**, London, v. 19, p. 809-818, 2010.

ZHANG, X. et al. The SsDREB transcription factor from the succulent halophyte Suaeda salsa enhances abiotic stress tolerance in transgenic tobacco. **International Journal of Genomics**, London, v. 2015, p. 875497, 2015.

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

ZHU, J. K. Abiotic stress signaling and responses in plants. **Cell**, Cambridge, v. 167, n. 2, p. 313-324, Oct. 2016.

_____. Cell signaling under salt, water and cold stresses. **Plant Biology**, Stuttgart, v. 4, p. 401-406, 2001.

ZHUANG, J. et al. Genome-wide analysis of the AP2/ERF gene family in Populus trichocarpa. **Biochemical and Biophysical Research Communications**, Orlando, v. 371, p. 468-474, 2008.

ZONG, W. Feedback regulation of ABA signaling and biosynthesis by a bZIP transcription factor targets drought-resistance-related genes. **Plant Physiology**, Bethesda, v. 171, p. 2810-2825, 2016.

CAPÍTULO 2 - OSCILLATIONS OF GENE EXPRESSION IN LEAVES AND ROOTS OF SOYBEAN GENOTYPES WITH CONTRASTING TOLERANCE PROFILE UNDER WATER DEFICIT STRESS

Artigo preparado de acordo com as normas para submissão do periódico

Frontiers in Plant Science.

OSCILLATIONS OF GENE EXPRESSION IN LEAVES AND ROOTS OF SOYBEAN GENOTYPES WITH CONTRASTING TOLERANCE PROFILE UNDER WATER DEFICIT STRESS

Rafaela Ribeiro Reis*[1], Juliana Marcolino-Gomes[2], Fabiana Aparecida Rodrigues[3], Silvana Rockenbach Marin[4]; Liliane Marcia Mertz-Henning[5], Alexandre Lima Nepomuceno[6]

[1] Londrina State University, Rodovia Celso Garcia Cid, Km 380. CEP 86051-900, Londrina, PR, Brazil.

[2] CNPq SEPN 507 BL B, - lt-2 tér- Asa Norte - Brasília, DF - CEP 70740-522

[3] Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso do Sul Rua Salime Tanure, s/n°. CEP 79400-000. Fone: (67) 3291-9600. Coxim - MS

[4] Embrapa Soja, Rodovia Carlos João Strass, Distrito de Warta, CEP 86001-970 PR, Brazil.

[5] Embrapa Soja, Rodovia Carlos João Strass, Distrito de Warta, CEP 86001-970 PR, Brazil.

[6] Embrapa Soja, Rodovia Carlos João Strass, Distrito de Warta, CEP 86001-970 PR, Brazil.

* Corresponding Author: rafaelarireis@gmail.com

number of words: 5206

number of figures: 9

number of tables: 2

Abstract

Water deficit stress is considered the major threat to soybean production worldwide. Considering that the tolerance to water deficit stress is a genetical complex trait, the identification of genes and the understanding of the respective involved routes in plants is necessary. To study the oscillations of gene expression in leaves and roots of soybean, two contrasting soybean genotypes, BR 16 (sensitive) and Embrapa 48 (tolerant), under different stress levels (moderate and severe) were evaluated. Additionally, a large difference in the expression pattern between both genotypes was detected. The differentially expressed genes in Embrapa 48 represented approximately 47% of confronting among all 54,175-soybean genes, while BR 16 represented 42%, for the experimental conditions evaluated. Moreover, when the analysis was restricted to only eight gene categories (Transcription factor (TF), amino acid, ABA precursor, ABA metabolism, development, drought, peroxidase and sucrose), it was observed that Embrapa 48 showed 770 more up-regulated genes than BR 16, totaling 2,456 and 1,882 expressed genes, respectively. In general, the response in leaves is more active than roots, in which was possible to define the differences in gene profile between Embrapa 48 and BR 16. According to the results, it is possible to infer that Embrapa 48 responds to stress faster than BR 16, presenting a larger number of genes expressed, while the sensitive genotype exhibited a down-regulated expression activity. This behavior persisted throughout all functional and biologic reported categories. Embrapa 48 exhibits initial gene modulation directed for dangers stress, maintaining genes related to active development processes, in other words, to basic functions. Additionally, our study allowed the identification of important candidate genes for drought tolerance. As example, those genes expressed exclusively in the drought tolerant cultivar Embrapa 48, belonging to the category of dehydration responsive genes, or genes with contrasting expression pattern between Embrapa 48 and BR 16 (Glyma13g35970), which possibly act in the drought tolerance response.

Keywords: abiotic stress, drought-tolerance response, abscisic acid (ABA), osmoregulation.

1 Introduction

Soybean, considered a worldwide commodity, is affected by biotic and abiotic stresses. In this way, due to climate changes and expansion of agricultural areas, new requirements related to factors that directly affect agricultural productivity have been established (Haggag et al., 2015). Therefore, the development of biotechnological tools for identification/characterization of promoters, genes, and other factors that add tolerance to abiotic stresses, such as drought, has great importance in the productive aspect (Abdallah et al., 2015; Cardi et al., 2017). Crop yield is highly affected by water deficit stress, in which the plants present physiological strategies, biochemical and morphological modulation as an adaptive/defense method under stress conditions.

Oya et al. (2004), evaluating different parameters to determine the contrast sensitivity profile of soybean genotypes under water deficit stress, found high yield at Embrapa 48 and therefore, was considered tolerant, while BR 16 was classified as sensitive when based in yield characteristics of the evaluated plants under water deficit conditions. However, molecular characterization, which is a highly informative analyze, capable of demonstrating the reason for this difference in performance between sensitive and tolerant cultivars, have not been performed. The molecular tolerance responses are considered complex traits controlled by many genes (Molina et al., 2008). An initial response at the cellular level is activated by inducing a complex network of signal transduction orchestrated by a sequence of molecular events. Some genes are considered principal codes for proteins that act directly (functional genes), avoiding the possible damages caused by stress, or has

regulatory roles, acting as activators and/or deactivators of proteins, in metabolic routes and in the activation of the gene cascades (Khazaei et al., 2013).

Several transcriptomic studies have already been performed in different species and experimental conditions. In soybean, research was performed using RNA sequencing (RNA-seq) as a platform for evaluating gene expression (Rodrigues et al., 2015), in which it was observed the daytime transcriptome fluctuations during water deficit stress. Furthermore, Marcolino-Gomes et al. (2015) characterized the expression profile of the transcription factor (TF) family GmAP2/EREBP. Similarly to this work, Rodrigues et al. (2012) employed Suppressive Subtractive Hybridization (SSH) for transcriptome analysis in soybean under drought conditions, and found differentially expressed genes and profile/modulation of gene classes in a both evaluated cultivars (Embrapa 48 and BR 16). However, a more informative description of differentially expressed genes up and down regulated in data generated by RNA-seq was not evaluated. Recent data shows that the deficit water stress has resulted in strong losses in soybean productivity. An example occurred in January 2018 in Argentina, in which the Buenos Aires stock exchange reduces the soybean harvest estimate to 37 million tons (down almost 30% from the previously prediction) (Agro-DBO, 2018). Thus, the refined and more complete analysis of the soybean transcriptome in response to water deficit stress, using RNA-seq as a potential tool in the more detailed reporting of genes and their respective expression profiles, would be very useful in the direction of strategies for genetic engineering and genome editing.

Therefore, in order to expand the knowledge and reveal new perspectives about the oscillation of the soybean transcriptome in stress conditions, in this study, gene expression profiles of contrasting soybean genotypes (BR 16 and Embrapa 48) was compared by RNA-seq, allowing the evaluation of the main molecular peculiarities that differentiate these cultivars in both tissues (leaf and root), under different stress levels (moderate and severe).

2 Results

2.1 Comparative differentially expressed genes in leaf and root in response to moderate and severe stress library

The transcriptomic analysis by RNA-seq platform was used to identify differential expressed genes in two soybean cultivars (BR 16 and Embrapa 48), under water deficit conditions. Samples from roots and leaves of plants submitted to moderate (25 and 50 min under stress) and severe stress (100 and 150 min under stress). The differential expression was calculated based on calibration with control samples (not stressed). A total of 47,177,642 reads were obtained after adapter removal, it was reduced to 23,246,624 reads (Table 1), when mapping considering only sequences with a maximum two mismatches in the first 32 bases. Furthermore, were obtained 55,787 mRNAs annotated, whose, 51,322 were hit by at least 1 mRNA and 39,951 pass the low counts filter.

Differentially expressed genes under water deficit conditions were distributed between up and down-regulated genes (Table. 2). Based in Figure 1 in S1 and Figure 1 in S2, it can be inferred that the stress responses are more evident in leaves, where genes differentially expressed (up and down-regulated) respond more intensively when compared to roots in both cultivars. Additionally, Embrapa 48 showed more up-regulated genes at moderate stress in leaves, different from BR 16 (Table 2 and Figure 1 in S1), which presented similar number of genes differentially expressed distributed in up and down-regulated, considering the same tissue and level of stress (Table 2).

In the moderate stress, the number of up-regulated genes from Embrapa 48 (4693 in leaf) exceeded the amount of BR 16 expressed genes (2009 in leaf), maintaining this structure when analyzing the total number of differentially expressed genes (up and down). Embrapa 48 evidenced more genes (±11%) than BR 16, obtaining the sum 25203 differentially

expressed genes, whereas BR 16 had only 22517 (Table 2). Embrapa 48 presented a number of differentially expressed genes greater than that exhibited by BR 16 (Figure 1A, C and D) for almost all treatments (except for leaf on severe stress). Regarding the treatment of leaf in severe stress library, it was possible to observe a large number of up-regulated genes (4125) exclusive of BR 16 modulating, while Embrapa 48 presented a few less (3588) (Figure 1B). Considering the roots, the moderate stress BR 16 showed 2389 exclusive genes, increasing this number to 3267 exclusive genes in next stress level. This was also observed for Embrapa 48 (3591 in moderate stress) and (3494 in severe stress). However, the increase in the number of genes from one level of stress to the other was not as high as in BR 16 (Figure 1C and D). In general, BR 16 showed a higher gene activation in severe stress, reaching almost double differentially expressed genes when compared to the number of genes obtained in the moderate stress (Figure 1). On the other hand, Embrapa 48 has demonstrated a more active gene modulation since the moderate stress library (Figure 1A and C).

Among the genes that presented the highest differential expression obtained by RNA-seq, the Glyma09g31740 gene (Dehydrin) identified in leaf of Embrapa 48 under severe stress, stood out, presenting the highest expression level among all up-regulated genes and a fold-change of 1594.81 (Log_2= 10.63) (red arrow shown in Figure 2D). Moreover, the gene Glyma20g29770 achieved a high expression level in leaf of BR 16 in the severe stress, with fold-change 1186.29 (Log_2= 10.21) (no annotation) (red arrow shown in Figure 2B).

In roots, the difference in the general expression profile was not as much evident between cultivars (Table 2 and Figure 1 in S2). Embrapa 48 and BR 16 showed more number of differential expression gene in severe stress, considering the moderate stress (Table 2). Embrapa 48 obtained more down-regulated genes than BR 16 (Table 2). Considering the intensity of gene modulation based on both profile expression (up and down), Embrapa 48 maintained a higher pattern in the moderate stress, obtaining 3575 genes, whereas BR 16 presented only 2431 exclusive genes (Figure 1C). In the roots of Embrapa 48, the gene Glyma03g26310 (*AP2* domain) was highlighted, obtaining high level of fold-change 544.18 (Log_2= 9.08), when compared to the others genes in same treatment (red arrow shown in Figure 3D).

Analyzing just eight categories of the twelve selected in this study, for leaf and root, Embrapa 48 showed approximately 770 more up-regulated genes when compared to BR 16, totaling 2456 and 1882 expressed genes respectively, whereas to down-regulated genes occurred, totaling 1686 in Embrapa 48 and 2408 in BR 16 (Figure 4).

2.2 Analysis of patterns expression and functional roles of leaf under water deficit stress

RNA-seq transcriptome data of Embrapa 48 and BR 16 genotypes showed significantly different expression patterns to leaves, according to the different metabolism and biological process (Figure 5). Twelve categories were selected based on the results of MapMan 3.6 ORC bioinformatics tool: transcript factor, aminoacid, tricarboxylic cicle acid (TCA), sucrose, photorespiration, light reactions, lipids, cell wall, acid abscisic (ABA) precursor, ABA metabolism, drought/Salt, peroxidase, development.

It is possible to verify the impact of stress, being the greatest difference between cultivars observed in the leaf (Figure 5 and Figure 1 in S1), in which the higher contrast was observed in the modulation of the gene expression pattern. Considering the results obtained for leaf, in the moderate stress library, it is observed that in the tolerant cultivar Embrapa 48, most of the genes presented an up-regulation pattern in all categories, indicating gene modulation against the damages of stress, besides maintaining the basic plant metabolism processes such as photosynthesis, respiration, growth and development, primary metabolism genes such as sugars and amino acids, as well as for secondary metabolism (flavonoids, terpenes, phenylpropanoids and phenolics and waxes) (Figure 5). This same profile was not

observed in the sensitive cultivar, which in the moderate period of stress, presented a strong profile of down regulation of the genes compared to Embrapa 48, which intensified as the stress increase occurred (Figure 5).

As regards the functional category, the tolerant cultivar Embrapa 48 presented a pattern with a higher number of up-regulated genes compared to cultivar BR 16 for both periods. With increased stress, the number of down-regulated genes increases for both cultivars, but with higher intensity in BR 16 sensitive (Figure 5). It was observed that in both cultivars, the highest number of expressed genes related to the amino acid category occurred at the libraries from leaf, mainly in the severe stress library. Embrapa 48 presented 69 genes and BR 16 86 genes. However, Embrapa 48 showed, in proportion to the number of genes observed, a higher concentration of up-regulated genes in relation to BR 16 (Figure 5). In addition, the analysis of the gene group classified in sucrose, lipids, cell wall, light reaction, photorespiration, tricarboxylic cicle acids (TCA) and secondary metabolism categories and in the leaf transcriptome, revealed that Embrapa 48 also showed up-regulated profile genes expressed in both periods of stress (moderate and severe), when compared with BR 16 (Figure 5).

Genes responsive to drought stress classes were predominantly up-regulated in Embrapa 48 (25 genes) when compared to BR 16 (11 genes), in both libraries (moderate and severe), considering leaf (Figure 6). Some genes, in this same class, such as Glyma09g40090 (dehydration-responsive protein-related), were differentially expressed (up-regulated) in moderate stress, however, it has a higher fold-change (Fc) in Embrapa 48 (5.3) than in BR 16 (3.6). Also in moderate stress, the genes Glyma14g06200, Glyma13g35970, Glyma18g53780, Glyma02g05840, Glyma16g32180, Glyma11g35590, Glyma01g07020, Glyma09g26650, Glyma01g07020 were exclusively expressed in Embrapa 48. Twenty genes stood out among the others due to their remarkable performances in the tolerant cultivar and expressive levels of expression. Among those differentially expressed, in the Embrapa 48 leaf under moderate stress, the Glyma03g04920 (Fc= 34.36), Glyma13g22420 (Fc=22.26), Glyma10g03640 (Fc=21.02), Glyma11g13940 (Fc=19.32), Glyma13g43970 (Fc=18.55), Glyma01g26230 (Fc=17.75), Glyma12g05910 (Fc=16.96), Glyma15g05350 (Fc=15.9), Glyma17g07070 (Fc=15.66), Glyma11g15060 (Fc=15.63) genes exceed their standard of expression compared to the 693 exclusive of this treatment. On the other hand, Glyma03g29440 (Fc=192.2), Glyma04g00710 (Fc= 63.66), Glyma07g38580 (Fc=60.31), Glyma11g05530 (Fc=56.55), Glyma17g11160 (Fc=43), Glyma16g04440 (Fc=37.74), Glyma11g21420 (Fc=36.85), Glyma08g18900 (Fc=35.51), Glyma07g06620 (Fc=34.34), Glyma18g52700 (Fc=32.39), of the 921 differentially expressed genes on Embrapa 48 leaf under severe stress had high expression profiles. Some gene are notable for being detected in a contrasting behavior in Embrapa 48 and BR 16, which possibly act decisively in the drought tolerance response, which is a characteristic reported in Embrapa 48. An example, in moderate stress, is Glyma13g35970, described as RD22 nutrient reservoir, which presented high expression in Embrapa 48, reaching a fold-change of 31.8. While in severe stress presented fold-change of 11.17 in Embrapa 48 and -6.02 (Fc) in BR 16.

In the peroxidase category, the vast majority of differentially expressed genes occurred in leaf in the severe stress library, for both cultivars (Figure 6). BR 16 overcame the number of expressed genes, prevailing those with down-regulated profile when compared to Embrapa 48 (predominance of up-regulated genes). This pattern was maintained for both libraries (moderate and severe stress) (Figure 6). For BR 16 and Embrapa 48, the highest number of peroxidase genes was found in severe stress library, for both tissues. In leaf (Figure 6) the most common gene profile was down-regulation, characteristic reproduced in moderate and severe libraries when compared with root. The highest number of abscisic acid (ABA)

genes in leaf from both cultivars was differentially expressed in moderate stress library as up-regulated profile.

Some ABA-related genes were exclusive from one of the cultivars evaluated. In the moderate stress library, the genes Glyma15g40070 (Nine-cis-epoxycarotenoid dioxygenase 3 (NCED3)) and Glyma08g10190 (Nine-cis-epoxycarotenoid dioxygenase (NCED5)) occurred only in Embrapa 48, in both categories (precursors and genes linked to ABA) (Table 1 in S3). By analyzing the moderate stress in leaf, the ABA-precursor category was common to BR 16 and Embrapa 48. Furthermore, some genes were found only in the metabolism map and not in ABA precursors, for both cultivars. Interestingly, some genes linked to ABA metabolism occurred in both cultivars, highlighting the Glyma15g40070, Glyma03g37140 and Glyma20g23500 (Table 1 in S3).

For leaves from both cultivars, the largest number of differentially expressed genes categorized as TF was presented in the severe stress library, although most of them were suppressed by the stress treatment. For the tolerant genotype, contrary to the severe stress library, the most differentially expressed TF genes in the moderate stress library presented up-regulated profile (Figure 6). For leaf, genes responsive to drought stress classes were predominantly up-regulated in Embrapa 48 when compared to BR 16, in both libraries (moderate e severe) (Figure 6).

2.3 Comparative analysis of pattern expression and functional roles of root under water deficit stress

Considering the results obtained from the MapMan 3.6 ORC1 bioinformatics tool, in the moderate stress library for roots, it was observed that in the tolerant cultivar, Embrapa 48, most of the genes presented an up-regulation pattern in sucrose and TCA categories. However, in the severe stress, BR 16 genes expressed to sucrose was mostly up-regulated in relation Embrapa 48 (Figure 7). Aminoacid category genes did not show induction as that found in leaves (Figure 7). Additionally, the analysis of the gene group classified in sucrose, lipids, cell wall, and secondary metabolism categories, in the root transcriptome, revealed that there was no such visible difference between the gene expression (Figure 7).

Regarding genes responsive to drought stress in root, for Embrapa 48 it was up-regulated, in both moderate and severe stress (Figure 8). In BR 16 and Embrapa 48, the highest number of peroxidase genes was found in the severe stress library, for both tissues. In this category, up-regulated genes predominated (Figure 8). In the ABA gene class, this scenario was inverted to Embrapa 48, when comparing its performance in leaf. BR 16 in the moderate and severe stress libraries presented a major number up-regulated gene (Figure 8). The largest number of differentially expressed genes related to TF was observed in the severe stress library, for both cultivars, predominantly the down-regulated expression profile (Figure 8).

2.4 Validation of genes expression

Using the RT-qPCR method, the primer Glyma17g17860 (5´AAAGGCACAGAGTGATGAAT 3´ (forward) - 3´ CTTGATGACCTTGTGTACCA 5´(reverse)) was used to validate the RNA-seq data, from the selection of the target genes, was considered a gene commonly described as responsive to water deficit stress, indicating the response of plant Late Embryogenesis Abundant protein 18 (*LEA 18*). The RNA-seq values showed low variation compared with the respective values from the RT-qPCR analyses (Figure 9A). In general, the linear equation demonstrated a high and good correlation coefficient value between RNA-seq and qPCR (R^2=0.9895) (Figure 9B). Additionally, genes from *LEA18* (Glyma17g17860) were highly up-regulated in severe stress of leaf treatment,

especially in Embrapa 48 reaching the value highest level (fc= 131.3) compared to BR 16 (fc= 70.34) (Figure 9C).

3 Discussion

The stress response in plants is a process coordinated by physiologic and biochemical reactions, such as gene expression control, gene activation or inactivation, protein synthesis, post-translational modification and crosstalk in different metabolic pathways. Therefore, it has become essential to study the responses of drought sensitive and tolerant plants at a genetic level, identifying the molecular differences between these two types of cultivars (Neves-Borges et al., 2012). Herein, soybean genotypes demonstrated different molecular responses under water deficit, pointing Embrapa 48 as tolerant and BR 16 as sensitive. Similar phenotyping was obtained by Oya et al. (2004).

The *LEA 18* genes was selected to this validation study. LEA proteins are a large and highly diverse gene family present in plant species. LEAs are proposed to play a role in various stress tolerance responses (Gao and Lan, 2016). Olvera-Carrillo et al. (2010) demonstrated that overexpression of a representative member of the LEA proteins confers tolerance to severe drought in *A. thaliana* plants. LEA 18 is not suggested as a membrane stabilizing protein, as observed for other LEA proteins. Instead, a possible function of LEA 18 could be the composition-dependent modulation of membrane stability, e.g., during signaling or vesicle-mediated transport (Hundertmark et al., 2011). The up-regulation of the *LEA18* gene was more intense in Embrapa 48, evidencing this cultivar as superior for drought tolerance responses, when compared to BR 16. According to observations, especially in leaves, Embrapa 48 responds to stress quickly, presenting a higher number of up-regulated genes expressed in moderate stress libraries when compared to BR 16. This behavior persists throughout the gene classes, both in amino acid and sucrose related genes as seen in Figure 5. In words, it exhibits a faster initial gene modulation. Both classes cited may be related to the osmoregulation mechanisms of the plant, such as proline, trehalose, mannitol, ectoine, glycine and betaine, providing tolerance to cellular dehydration (Ha et al., 2015). Therefore, increase in stress tolerance due to water deficiency occurs mainly through osmotic adjustment and osmoprotective character (Conde et al., 2008).

Proline is an amino acid that acts on the osmoregulation in plants under stress conditions, providing tolerance to cellular dehydration (Varshney et al., 2011; Wani and Gosal, 2011). The accumulation of organic solute in response to drought is an important mechanism to maintain cellular turgor, contributing to a reduction of water potential (Ψw), which increases the water absorption capacity of plants (Silva et al., 2009). A major concentration of proline genes was showed in Embrapa 48 (Table 1 in S3). Thus, osmoregulation such as tolerance response, acts efficiently, contributing to positive response, delaying the damages caused by the low water content inside the plant cells. Osmoregulation is considered one of the main and most important tolerance mechanism, being able to characterize/determine a plant with a high degree of tolerance or susceptibility (Pantuwan et al., 2002). Sucrose has great metabolic importance in plants, being fundamental in tissues such as roots. According to studies conducted in maize (*Zea mays*) by Ogawa et al. (2005), sugars in roots contributed to both initiation and elongation, collaborating for a better performance of the plant in drought conditions. Auxin and sugars have been found to play an important role in the initiation of lateral roots (Takahashi et al., 2003). Responses observed in these studies corroborate with results obtained in this study, which was possible to detect the predominance of sucrose-related among the up-regulated genes in Embrapa 48, considering the moderate stress library in leaf and root to the detriment of the cultivar sensitiveness to water deficit stress (BR 16) (Figure 5 and 7).

Gargallo-Garriga et al. (2014), in metabolite analyses, suggest that leaves and roots present a contrasting metabolism in response to changes of environmental conditions, presenting different physiological mechanisms and function in adapting to stress. Many metabolic products and soluble sugars are produced in leaves but allocated and used in roots, due to root energy requirements for the assimilation of soil resources and growth. However, other molecules, such as terpenes and metabolites related to anti-stress mechanisms, are increased in leaves under drought. Thus, it is possible that the sucrose levels higher in leaves since the first level of stress observed in Embrapa 48, provides energy to primary metabolism and defense responses, inducing osmoregulation. Furthermore, solutes are accumulated to prepare cells to translocation of metabolites from leaves to root (Figure 5). On the other hand, BR 16 showed predominance only in up-regulated genes in root under severe stress (Figure 7). The metabolism of roots also appears to be strongly homeostatic and conserved than leaves, as also suggested by Gargallo-Garriga et al. (2014).

The reduction of stomatal conductance is a way of defense against cell dehydration. Rodrigues et al. (2012) developed similar tests to the ones performed in the present study and verify superior stomatal conductance in Embrapa 48 to all levels of stress, compared to BR 16. Thus, BR16 has probably entered oxidative stress precociously due stomatal closure and consequent reduction of CO_2 assimilation. It leads to accumulation of ATP, energy reduction (NADPH) and reduction of final acceptor set of electron transport chain ($NADP^+$). The excessive reduced activity of NADPH can induce an exaggerated limitation of the electron transport chain. In this process, electrons can escape and react with molecular oxygen (O_2), forming the reactive oxygen species (ROS) (Cavatte et al., 2011). This disordered accumulation of ROS causes oxidative stress.

Down-regulation of TCA (Tricarboxylic Acid) cycle and amino acid biosynthesis apparently acts to avoid energy loss under oxidative stress conditions (Urbanczyk-Wochniak et al., 2003). In this way, BR 16 was possibly under oxidative stress since the first stress situation, presenting higher number of down-regulated genes, whereas Embrapa 48 presented a down-regulated gene profile in the severe stress library. The high level of up-regulated peroxidases found in Embrapa 48 since the moderate stress in leaves is another characteristic that reinforces that hypothesis. Peroxidases act by combating the ROS, thus, it can be inferred that Embrapa 48 did not come to collapse due to oxidative stress (Figure 6). In addition, assuming that Embrapa 48 has a higher degree of tolerance than BR 16, it is evident that the predominance of up-regulated activity of abscisic acid (ABA) in Embrapa 48 (Figure 6) has led to whole response, activating signalized molecules of the stress, such as ROS, and at the same time recruiting the synthesis of peroxidases (Figure 6), providing equilibrium to the system. For example, Jiang and Zhang (2002) reported that a significant increase in ABA content precedes ROS increase, followed by activities of antioxidant enzymes increase. Shinozaki and Yamaguchi-Shinozaki and Shinozaki (2006) also concluded that water deficit stress induce the accumulation of ABA in plants. ABA accumulation promotes changes in gene expression and in stomata closing, leading to concomitant reduction in transpiration, carbon assimilation decreases and water loss.

The better performance of Embrapa 48 under water deficit is supported by the strong presence of up-regulated genes related to TF and drought (Figure 6). Stress-induced transcription factors are considered powerful targets, such as a natural starting point for mechanisms regulation of expression of several genes, and key to genetic transformation strategies (Woodrow et al., 2012). Interestingly, Embrapa 48 releases energy for the stimulation (of more focused) genes to combat the effects of stress, and conservation secondary metabolism (Figure 6). Furthermore, it is able to preserve genes of primary metabolism activated with basic function in plants, as well as, development, light reactions, photorespiration equilibrium and energy suppliers (Figure 5 and 6).

Changes in photosynthetic activity can be associated with membrane damage in mesophyll cells, decrease in chlorophyll content, or alteration in the synthesis and transport of assimilates. This decrease in the rate of liquid photosynthesis under stress is also related to biochemical damage of non-stomatal nature caused by the oxidation of lipids found in chloroplast, especially in the thylakoid membrane, and changes in pigment and protein structure. However, Embrapa 48 possibly maintained the photosynthetic activity on leaf, since the stimulation of gene classified in lipid and cell wall metabolism remaining up-regulated on both levels of stress compared to BR 16 (Figure 5).

Among those twenty differentially expressed genes selected as the Top genes, in Embrapa 48 leaf under moderate and severe stress, the genes that excelled in their expression patterns exclusive of these two treatments, are annotated in several classes, histone H2A variant 1-related, histone H3 (H3) , glutathione s-transferase u1-related, histone H3 (H3), histone H4 (H4), histone H2B (H2B), raffinose synthase, proline-rich protein 4, Protein kinase domain (Pkinase), zinc finger protein-related, 3 , 4-dihydroxy-2-butanone-4-phosphate synthase and G-box-binding factor 2-related. Some of these categories are linked to the processes of osmoregulation, energy balance of cells and gene regulation.

In summary, the top genes can be applied as potential candidates for future studies aiming at drought tolerance, since these genes were exclusive in the tolerant cultivar, presenting high levels of differential expression and also participate important roles in the tolerance response. A more intense gene modulation in Embrapa 48 was observed under moderate stress, maintained the basal metabolism, which was not observed in BR 16. Additionally, the gene modulation response under water deficit stress was predominantly in leaves to both cultivars. Furthermore, a confident correlation between RNA-seq and RT-qPCR was obtained.

4 Methods

4.1 Plant Materials and Experimental Design in hydroponic solution

4.1.1 Plant growth and drought stress treatment for Transcriptome analysis

Based on the method and experimental design proposed by Martins et al., 2008, soybean transcriptomic data under water deficit conditions using hydroponic system was evaluated. Seeds of the soybean cultivars BR 16 and Embrapa 48, classified by Oya et al. (2004) as sensitive and tolerant to water deficit, respectively, were germinated on germination paper during four days in a growth chamber at 25 ± 1 ºC and 100% relative humidity. Seedlings were placed in 36 L boxes containing Hoagland's solution 50% (without modifications) (Hoagland and Arnon, 1950), which was continuously aerated and replaced on a weekly basis. These boxes were then transferred to a greenhouse under a natural photoperiod of approximately 12/12 h light/dark cycle, temperature of 30 ± 5 ºC and relative humidity (RH) of 60 ± 10%, where the plants were allowed to grow until the V4 stage (Fehr et al., 1971). The experimental design was a randomized complete block design, 2x5 factorial, with three replications. The factors were two cultivars (BR 16 and Embrapa 48) and five periods of dehydration (0, 25, 50, 125 and 150 min). The stress was applied by removing the plants out of the hydroponic solution and leaving them in boxes without nutrient solution for up to 150 min under ambient-air exposure. For each time of water deficit, roots from 15 plants were collected, pooled and frozen in liquid nitrogen before storage at -80 ºC. The same procedure was performed for leaf samples.

4.2 Library construction and sequencing run

The total RNA was extracted from leaves and root samples from BR 16 and Embrapa 48 plants, using the Trizol Reagent (Invitrogen). Bulks of total RNA was made using all samples of same tissue (25 and 50 min - moderate stress library) and (125 and 150 min -

severe stress library), and the control. . Additionally, DNase treatment (Life Technologies Grand Island, NY, USA) with high-quality total RNA was used to analyze the transcripts. The RNA-seq libraries were built using the Nugen-OvationH kit according to the manufacturer's instructions (NuGEN Technologies Inc., San Carlos, CA, USA). The libraries obtained were subjected to sequencing by Illumina HiSeq2000 (Illumina, San Diego, CA, USA).

4.3 Mapping of readings and functional classification

Mapping of the readings was performed with the soybean genome (Phytozome Glycine max v1.1) using the GeneSifter platform (http://www.geospiza.com/Products/AnalysisEdition.shtml). To compare gene expression between different times and conditions, $\log^{2-}$ was used to transform the normalized readings by mapped million (RPM). Then a test analysis was run (t-test for two group comparisons). Libraries tagged with a barcode produced thousands of readings from each library. Contig sequences were submitted to the non-redundant protein database NCBI through Blast X (Altschul et al., 1997) in order to search for similarity with known proteins. In addition, sequences were analyzed by the moderateware AutoFact (Koski et al., 2005), which is an automated annotation tool that assigns biological information for a given sequence by comparing different databases. We used the UniRef90 and UniRef100, KEGG (Kanehisa and Goto, 2000), Pfam (Finn et al., 2010), Smart (Schultz et al., 1998) databases. In order to establish the GO (Gene Ontology) terms (Ashburner et al., 2000), we employed the Blast2Go program to classify the sequences according to the molecular function and biological process described to the respective protein (Götz et al., 2008; Carbon et al., 2009).

4.4 Analysis of differential gene expression

Using the GeneSifter platform (http://www.geospiza.com/Products/AnalysisEdition.shtml), we applied a pairwise comparison between the control and water deficit treatment using all libraries synthesized from plants of the same time period/treatment. In the pairwise analysis, we only used genes with more than 20 mapped readings to compare gene expression using the edgeR statistical test (Robinson et al., 2010). A ratio of expression (fold-change) was performed by dividing values of gene expression under water deficit and control conditions. The statistical test was combined with the multiple-hypothesis-testing correction method of Benjamini and Hochberg (Benjamini and Hochberg, 1995), which calculates the False Discovery Rate (FDR) to qualify statistically significant and differentially expressed genes by avoiding inflation of type-1 errors. Differential gene expression was considered significant at an adjusted p-value ≤ 0.01, and down- and up-regulation was established in the range of $\leq$ -2 to $\geq$ 2fold-change (fc), respectively. We also applied a stringent statistical significance cutoff (adjusted p-value $\leq$0.001) to improve confidence.

4.5 Classification and analyse of functional categories

Analyses through MapMan 3.6 ORC1 (http://mapman.gabipd.org/web/guest/mapman-version-3.6.0) allowed us to visualize differentially expressed genes, which was calculated based on calibration with time 0 of stress (control) and were classified by functional categories in several pathways. Twelve categories were mapped and selected for analysis (transcript factor, aminoacid, tricarboxylic cicle acid (TCA), sucrose, photorespiration, light reactions, lipids, cell wall, acid abscisic (ABA) precursor, ABA metabolism, drought/Salt, peroxidase, and development).

4.6 Validation of genic expression

4.6.1 Gene Selection

The selection of the gene for validation was performed through the search of data deposited in the GeneSifter - Analysis Edition (GSAE) System Requirements, an online moderateware analysis that assists in the rapid visualization of the amount of data generated by Next Gen Sequencing and Microarray Analysis, besides explore them by biological definition (Table 1 in S4). Differential gene expression was considered significant at an adjusted p-value ≤ 0.01, and down- and up-regulation was established in the range of $\leq$ -2 to ≥ 2 fold-change (fc), respectively.

4.6.2 Expression analysis

Relative expression levels of the target genes were measured in root and leaf samples from Embrapa 48 and BR 16 plants for each bulk time point (0, 25- 50, 125-150 min under water deficit), generating (control, moderate and severe) libraries, with three biological replicates, each with three technical replicates, were analyzed. After the DNAse treatment (Invitrogen/Life Technologies, Grand Island, NY, USA), high quality total RNA was used to synthesize cDNA strands (Superscript II First Strand Synthesis, Invitrogen/Life Technologies, Grand Island, NY, USA), and cDNA quality was verified using a standard PCR reaction with an actin primer that spanned an intronic region. Additionally, the genes were amplified by qPCR using a StepOne RT-qPCR Thermocycler (Applied Biosystems/Life Technologies, Grand Island, NY, USA) with the following cycling FAST parameters: 95 ºC for 20 seconds, 40 cycles at 95 ºC for 3 seconds, 60 ºC for 30 seconds and Melt Curve 95 ºC for 15 seconds and 60 ºC for 1 minute.

Data were collected during the extension phase, and dissociation curves were performed by heating each amplicon from 60 to 95 ºC and taking readings at one-degree intervals to verify the specificity of the primers. The Rest2009 moderateware package (Pfaffl et al., 2002) was used to evaluate the data because this program provided a more robust statistical analysis. The normalization of the real-time quantitative RT-qPCR was performed by taking the geometric average of the selected endogenous genes (FYVE and B-act, according with analyses of Marcolino-Gomes et al. (2015), and the control plants (0 min under stress) were used to normalize the relative expression. Hypothesis testing was used to determine whether the differences between the control and treatment conditions were significant (Pfaffl et al., 2002).

4.6.3 Primer design and efficiency analysis

Primers for the target genes were designed based on the GeneModels of the selected genes using the program Primer Express 3.0 (Applied Biosystems/Life Technologies, Grand Island, NY, USA), resulting in (5´AAAGGCACAGAGTGATGAAT 3´ (forward) - 3´ CTTGATGACCTTGTGTACCA 5´(reverse)). Primer sequences were determined for the 3'end of each gene, and the amplicons spanned up to 150 base pairs (bp). Primer sequences were BLASTed against the soybean genome (Phytozome database v1.0, http://www.phytozome.net/search.php) to verify the specificity of each primer, and standard curves were produced from serial dilutions of a cDNA pool to estimate the efficiency of the PCR amplification reactions.

5 Final considerations

It was observed an abundant set of information comparing the different functions and metabolic activity of differentially expressed genes in leaves and roots of two sensitivity genotypes that are contrasting to water stress. According to shown data, the difference in the response evidenced in leaves and roots is remarkable, which connect each other to better

performance of the plants under stress conditions. In fact, the leaves orchestrate a greater number of up-regulated genes, which may or may not be active on leaves. In addition, Embrapa 48 (tolerant genotype) anticipates the timing of gene activity from the first signs of stress (moderate stress). Furthermore, the top 10 genes might be potential study tools as candidates for future work aiming at drought tolerance.

References

Abdallah, N. A., Prakash, C. S., and Mchughen, A. G. (2015). Genome editing for crop improvement: challenges and opportunities. GM Crops Food. 6, 183–205. doi: 10.1080/21645698.2015.1129937

Agro-DBO (2018). Soja: projeção da safra argentina é reduzida. Rev. Agro. DBO. Disponível em: <http://www.portaldbo.com.br/Agro-DBO/Noticias/Soja-projecao-da-safra-argentina-e-reduzida/24163>. Acesso em: 15 mar. 2018.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic. Acids Res. 25, 3389-33402. doi: https:// doi.org/10. 1093/nar/25.17.3389

Andrade, J. L., Larque-Saavedra, A., and Trejo, C. L. (1995). Proline accumulation in leaves of four cultivars of Phaseolus vulgaris L. with different drought resistance. Phyton. 57, 149–157. doi: https://doi.org/10.1016/S0254-6299(15)30239-8

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D, Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat. Genet. 25, 25-29. doi: 10.1038/75556

Augustine, S. M., Ashwin-Narayan, J., Syamaladevi, D. P., Appunu, C., Chakravarthi, M., Ravichandran, V., et al. (2015). Overexpression of EaDREB2 and pyramiding of EaDREB2 with the pea DNA helicase gene (PDH45) enhance drought and salinity tolerance in sugarcane (Saccharum spp. hybrid). Plant Cell. Rep. 34, 247–263. doi: 10.1007/s00299-0141704-6

Babitha, K. C.; Ramu, S. V.; Pruthvi, V., Mahesh, P., Nataraja, K. N., and Udayakumar, M. (2013). Co-expression of AtbHLH17 and AtWRKY28 confers resistance to abiotic stress in Arabidopsis. Transgenic. Res. 22, 327–341. doi: 10.1007/s11248-012-9645-8

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Statist. Soc. B., 57, 289–300. URL: http://www.jstor.org/stable/2346101

Besseau, S., Li, J., and Palva, E. T. (2012). WRKY54 and WRKY70 co-operate as negative regulators of leaf senescence in Arabidopsis thaliana. J. Exp. Bot. 63, 2667–2679. doi: 10.1093/jxb/err450

Cai, R., Zhao, Y., Wang, Y., Lin, Y., Peng, X., Li, Q., et al. (2014). Overexpression of a maize WRKY58 gene enhances drought and salt tolerance in transgenic rice. Plant Cell. Tiss. Org. Cult. 119, 565–577. doi: 10.1007/S212 40-0140556-7

Campanhia Nacional de Abastecimento – CONAB. (2017). Observatório agrícola. Acomp. Safra Bras. Grãos 4, 1-170.

Carbon, S., Ireland, A., Mungall, C. J., Shu, S., Marshall, B., and Lewis, S. (2009). AmiGO: online access to ontology and annotation data. Bioinformatics. 25, 288-289. doi: 10.1093/bioinformatics/btn615

Cardi, T., D'Agostino, N., and Tripodi, P. (2017). Genetic transformation and genomic resources for next-generation precise genome engineering in vegetable crops. Front. Plant Sci. 8, 241. doi: 10.3389/fpls.2017.00241

Carvalho, M. H. C. de (2008). Drought stress and reactive oxygen species. Plant Signal. Behav. 3, 156-165.

Cavatte, P. C., Martins, S. C. V., Morais, L. E., Silva, P. E. M., Souza, L. T., and Matta, F. M. (2011). A fisiologia dos estresses abióticos, in Melhoramento de Plantas para Condições de Estresses Abióticos, R. Fritsche-Neto, e A. Borém (Viçosa: Ed. UFV), 39-45.

Chen, H., Liu, L., Wang, L., Wang, S., and Cheng, X. (2015). VrDREB2A, a DREB-binding transcription factor from Vigna radiata, increased drought and high-salt tolerance in transgenic Arabidopsis thaliana. J. Plant Res. 129, 263-273. doi: 10.1007/S20265-015-0773-0

Chen, L. G., Zhang, L. P., and Yu, D. Q. (2010). Wounding-induced WRKY8 is involved in basal defense in Arabidopsis. Mol. Plant. Microbe. Interact. 23, 558–565. doi: 10.1094/MPMI-23-5-0558

Chu, X., Wang, C., Chen, X., Lu, W., Li, H., Wang, X., et al. (2015). The cotton WRKY gene GhWRKY41 positively regulates salt and drought stress tolerance in transgenic Nicotiana benthamiana. PLoS One 10, e0143022. doi: 10.1371/ journal.pone.0143022

Conde, C., Delrot, S., and Gerós, H. (2008). Physiological, biochemical and molecular changes occurring during olive development and ripening. J. Plant Physiol. 165, 1545-1562. doi: 10.1016/j.jplph.2008.04.018

Das, K., and Roychoudhury, A. (2014). Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. Front. Environ. Sci. 2, 53. doi: 10.3389/fenvs.2014.00053

Dias, L. P., Oliveira-Busatto, A. de, and Bodanese-Zanettini, M. H. (2016). The differential expression of soybean [Glycine max (L.) Merrill] WRKY genes in response to water deficit. Plant Physiol. Biochem. 107, 288-300. doi: 10710. 1016/j.plaphy. 2016.06.018

Ding, Z. J., Yan, J. Y., Li, C. X., Li, G. X., Wu, Y. R., and Zheng, S. J. (2015). Transcription factor WRKY46 modulates the development of Arabidopsis lateral roots in osmotic/salt stress conditions via regulation of ABA signaling and auxin homeostasis. Plant J. 84, 56–69. doi: 10.1111/tpj.12958

Dossa, K., Wei, X., Li, D., Fonceka, D., Zhang, Y., Wang, L., et al. (2016). Insight into the AP2/ERF transcription factor superfamily in sesame and expression profiling of DREB subfamily under drought stress. BMC Plant Biol. 16, 171. doi: 10.1186/s12870-016-0859-4.

Elstner, E. F. (1987). Metabolism of activated oxygen species, in: Biochemistry of Plants, ed. D. D. Davies (London: Academic Press), 253-315.

Elstner, E. F., Osswald, W. (1994). Mechanisms of oxygen activation during plant stress. Proc. R. Soc. Edinb. Sect. B., Biol. Sci. 102, 131-154. doi: https:// doi.org/10.1017/S0269727000014068

Engels, C., Fuganti-Pagliarini, R., Marin, S. R., Marcelino-Guimarães, F. C., Oliveira, M. C., Kanamori, N., et al. (2013). Introduction of the rd29A:AtDREB2A CA gene into soybean (Glycine max L. Merril) and its molecular characterization in leaves and roots during dehydration. Genet. Mol. Biol. 36, 556–565. doi: 10.1590/S2415-47572013000400015

Fehr, W. R., Caviness, C. E., Burmood, D. T., and Pennington, J. S. (1971). Stage of development descriptions for soybeans, Glycine Max (L.) Merrill 1. Crop. Sci. 11, 929. doi: 10.2135/cropsci1971.0011183X001100060051x

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., et al. (2010). The Pfam protein families database. Nucleic. Acids Res. 38, D211-D222. doi: 10.1093/nar/gkt1223

Gaion, L. A., Baldo, A., Ferreira Júnior, D. da C., D'Amico-Damião, V., e Carvalho, R. F. (2017). Ethylene modulates tomato growth during drought stress. Holos Environ., Suppl., 56-68.

Gao, J., and Lan, T. (2016). Functional characterization of the late embryogenesis abundant (LEA) protein gene family from Pinus tabuliformis (Pinaceae) in Escherichia coli. Sci. Rep. 6, 19467. doi: 10.1038/srep19467

Gargallo-Garriga, A., Sardans, J., Pérez-Trujillo, M., Rivas-Ubach, A., Oravec, M., Vecerova, K., et al. (2014). Opposite metabolic responses of shoots and roots to drought. Sci. Rep. 4, 6829. doi: i10.1038/srep06829

Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., et al. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic. Acids. Res. 36, 3420-3435. doi: 10.1093/nar/gkn 176

Grunewald, W., Smet, I. de, Lewis, D. R., Löfke, C., Jansen, L., Goeminne, G., et al. (2012). Transcription factor WRKY23 assists auxin distribution patterns during Arabidopsis root development through local control on flavonol biosynthesis. Proc. Natl. Acad. Sci. U. S. A. 109, 1554–1559. doi: 10.1073/ pnas.1121134109

Ha, C. V., Watanabe, Y., Tran, U. T., Le, D. T., Tanaka, M., Nguyen, K. H., et al. (2015). Comparative analysis of root transcriptomes from two contrasting drought-responsive Williams 82 and DT2008 soybean cultivars under normal and dehydration conditions. Front. Plant. Sci. 6, 551. doi: 10.3389/fpls.2015. 00551

Haggag, W. M., El Habbasha, E. S., and Abouziena, H. (2015). Agriculture biotechnology for management of multiple biotic and abiotic environmental stress in crops. J. Chem. Pharm. Res. 7, 882-889.

Hernandez, D., Francois, P., Farinelli, L., Osteras, M., and Schrenzel, J. (2008). De novo bacterial genome sequencing: Millions of very moderate reads assembled on a desktop computer. Genome Res. 18, 802–809. doi: 10.1101/ gr.072033.107

671 Hoagland, D. R., and Arnon, D. I. (1950). The water-culture method for growing plants
672 without soil. Berkeley: California Agricultural Experiment Station.
673
674 Hrdlickova, R., Toloue, M., and Tian, B. (2017). RNA-Seq methods for transcriptome
675 analysis. Wiley Interdiscip. Rev. RNA 8, e1364. doi: 10. 1002/wrna.1364.
676
677 Hsieh, E. J., Cheng, M. C., and Lin, T. P. (2013). Functional characterization of an abiotic
678 stress-inducible transcription factor AtERF53 in Arabidopsis thaliana. Plant. Mol. Biol. 82,
679 223–237. doi: 10.1007/s11103-013-0054-z
680
681 Huang, X., Li, K., Xu, X., Yao, Z., Jin, C., and Zhang, S. (2015). Genome-wide analysis of
682 WRKY transcription factors in white pear (Pyrus bretschneideri) reveals evolution and
683 patterns under drought stress. BMC Genomics. 16, 1104 doi: 10.1186/s12864-015-2233-6
684
685 Hundertmark, M., Dimova, R., Lengefeld, J., Seckler, R., and Hincha, D. K. (2011). The
686 intrinsically disordered late embryogenesis abundant protein *LEA18* from Arabidopsis
687 thaliana modulates membrane stability through binding and folding. Biochim. Biophys. Acta.
688 1808, 446-453. doi: https://doi.org/10.1016/j. bbamem.2010.09.010
689
690 Jiang, M., and Zhang, J. (2002). Water stress-induced abscisic acid accumulation triggers the
691 increased generation of reactive oxygen species and up regulates the activities of antioxidant
692 enzymes in maize leaves. J. Exp. Bot. 53, 2401-2410. doi: 10.1093/jxb/erf090
693
694 Jiang, W., Wu, J., Zhang, Y., Yin, L., and Lu, J. (2015). Isolation of a WRKY30 gene from
695 Muscadiniarotundifolia (Michx)and validation of its function under biotic and abiotic stresses.
696 Protoplasma. 252, 1361-1374. doi: 10.1007/s00709- 015-0769-6.
697
698 Jiang, Y., and Deyholos, M. K. (2009). Functional characterization of Arabidopsis NaCl-
699 inducible WRKY25 and WRKY33 transcription factors in abiotic stresses. Plant Mol. Biol.
700 69, 91-105. doi: 10.1007/s11103-008-9408-3
701
702 Jiang, Y., Liang, G., and Yu, D. (2012). Activated expression of WRKY57 confers drought
703 tolerance in Arabidopsis. Mol. Plant 5, 1375–1388. doi: 10.1093/mp/sss080
704
705 Joshi, R., Singh, B., Bohra, A., and Chinnusamy, V. (2016). Salt stress signalling pathways:
706 specificity and crosstalk, in Managing Salinity Tolerance in Plants: Molecular and Genomic
707 Perspectives, eds. S. H. Wani, and M. A. Hossain (Boca Raton, FL: CRC Press), 51–78.
708
709 Kanehisa, M. and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes.
710 Nucleic. Acids. Res. 28, 27-30. doi: https://doi.org/10.1093/nar/28. 1.27
711
712 Khazaei, H., Street, K., Bari, A., Mickay, M., and Stoddard, F. L. (2013). The FIGS (Focused
713 Identification of Germplasm Strategy) approach identifies traits related to drought adaptation
714 in *vicia faba* genetic resources. PLoS One 8, e63107. doi:
715 https://doi.org/10.1371/journal.pone.0063107
716
717 Kido, E. A., Ferreira Neto, J. R., Silva, R. L., Belarmino, L. C., Bezerra Neto, J. P., Soares-
718 Cavalcanti, N. M., et al. (2013). Expression dynamics and genome distribution of
719 osmoprotectants in soybean: identifying important components to face abiotic stress. BMC
720 Bioinformatics. 14, S7. doi: 10.1186/1471-2105-14-S1-S7.

Kidokoro, S., Maruyama, K., Nakashima, K., Imura, Y., Narusaka, Y., Shinware, Z. K., et al. (2009) The phytochrome-interacting factor PIF7 negatively regulates DREB1 expression under circadian control in Arabidopsis. Plant Physiol. 151, 2046–2057.

Kidokoro, S., Watanabe, K., Ohori, T., Moriwaki, T., Maruyama, K,. Mizoi, J., et al. (2015). Soybean BRB1/CBF-type transcription factors function in heat and drought as well as cold stress-responsive gene expression. Plant J. 81, 505-518. doi: 10.1111/tpj.12746.

Kim, J. H., Choi, D., and Kende, H. (2003). The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in Arabidopsis . Plant. J. 36, 94-104. doi: 10.1046/j.1365-313X.2003.01862.x

Kim, J. S., Mizoi, J., Yoshida, T., Fujita, Y., Nakajima, J., Ohori, T., et al. (2011). An ABRE promoter sequence is involved in osmotic stress-responsive expression of the DREB2A gene, which encodes transcription factor regulating drought inducible genes in Arabidopsis. Plant Cell. Physiol. 52, 2136–2146. doi: 10.1093/pcp/pcr143.

Kim, J.-M., Sasaki, T., Ueda, M., Sako, K., and Seki, M. (2015). Chromatin changes in response to drought, salinity, heat, and cold stresses in plants. Front. Plant Sci. 6, 114. doi: 10.3389/fpls.2015.00114

Koski, L. B., Gray, M. W., Lang, B. F., and Burger, G. (2005). AutoFACT: an automatic functional annotation and classification tool. BMC Bioinformatics. 6, 151. doi: 10.1186/1471-2105-6-151

Kulik, A., Wawer, I., Krzywinska, E., Bucholc, M., and Dobrowolska, G. (2011). SnRK2 protein kinases-key regulators of plant response to abiotic stresses. OMICS 15, 859-872. doi: 10.1089/omi.2011.0091

Le Gall, H., Philippe, F., Domon, J-M., Gillet, F., Pelloux, J., and Rayon, C. (2015). Cell wall metabolism in response to abiotic stress. Plants 4, 112–166. doi: 10.3390/plants4010112.

Lee, S. J., Kang, J. Y., Park, H. J., Kim, M. D., Bae, M. S., Choi, H. I., et al. (2010). DREB2C interacts with ABF2, a bZIP protein regulating abscisic acid-responsive gene expression, and its overexpression affects abscisic acid sensitivity. Plant Physiol. 153, 716–727. doi: 10.1104/pp.110.154617

Li, J. B., Luan, Y. S., and Liu, Z. (2015). Overexpression of SpWRKY1 promotes resistance to Phytophthora nicotianae and tolerance to salt and drought stress in transgenic tobacco. Physiol. Plant 155, 248–266. doi: 10.1111/ppl. 12315

Li, R., Li, Y., Kristiansen, K., and Wang, J. (2008). SOAP: moderate oligonucleotide alignment program. Bioinformatics. 24, 713–714. doi: 10.1093/bioinformatics/btn025

Li, X., Duan, X., Jiang, H., Sun, Y., Tang, Y., Yuan, Z., et al. (2006). Genome-wide analysis of basic/helix-loop-helix transcription factor family in rice and arabidopsis. Plant Physiol. 141, 1167–1184. doi: 10.1104/pp.106.080580.

770 Luo, X., Bai, X., Sun, X., Zhu, D., Liu, B., Ji, W., et al. (2013). Expression of wild soybean
771 WRKY20 in Arabidopsis enhances drought tolerance and regulates ABA signalling. J. Exp.
772 Bot. 64, 2155–2169. doi: 10.1093/jxb/ert073
773
774 Marcolino-Gomes, J., Rodrigues, F. A., Fuganti-Pagliarini, R., Nakayama, T. J., Reis, R. R.,
775 Farias, J. R. B., et al. (2015). Transcriptome-wide identification of reference genes for
776 expression analysis of soybean responses to drought stress asevere the day. PLoS One 10,
777 e0139051. doi: https://doi.org/10.1371/ journal.pone.0139051
778
779 Martins, P. K., Jordão, B. Q., Yamanaka, N., Farias, J. R. B., Beneventi, M. A., Binneck, E.,
780 et al. (2008). Differential gene expression and mitotic cell analysis of the drought tolerant
781 soybean (Glycine max L. Merrill Fabales, Fabaceae) cultivar MG/BR46 (Conquista) under
782 two water deficit induction systems. Genet. Mol. Biol. 31, 512–521. doi:
783 http://dx.doi.org/10.1590/S1415-47572008000300 019
784
785 Molina, C., Rotter, B., Horres, R., Udupa, S. M., Besser, B., Bellarmino, L., et al. (2008).
786 SuperSAGE: the drought stress-responsive transcriptome of chickpea roots. BMC Genomics.
787 9, e553. doi: 10.1186/1471-2164-9-553
788 Morimoto, K., Mizoi, J., Qin, F., Kim, J. S., Sato, H., Osakabe, Y., et al. (2013). Stabilization
789 of Arabidopsis DREB2A is required but not sufficient for the induction of target genes under
790 conditions of stress. PLoS One 8, e80457. doi: 10.1371/journal.pone.0080457
791
792 Nair, A. S., Abraham, T. K., and Jaya, D. S. (2006). Morphological and physiological changes
793 in cowpea (Vigna unguiculata L.) subjected to water deficit. Indian J. Plant. Physiol. 11, 325–
794 328.
795
796 Nakano, T., Suzuki, K., Fujimura, T., and Shinshi, H. (2006). Genome-wide analysis of the
797 ERF gene family in Arabidopsis and rice. Plant Physiol. 140, 411–432. doi:
798 https://doi.org/10.1104/pp.105.073783
799
800 Nakashima, K., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2014). The transcriptional
801 regulatory network in the drought response and its crosstalk in abiotic stress responses
802 including drought, cold, and heat. Front. Plant Sci. 5, 170. doi: 10.3389/fpls.2014.00170.
803
804 Neves-Borges, A. C., Guimarães-Dias, F., Cruz, F., Mesquita, R. O., Nepomuceno, A. L.,
805 Romano, E., et al. (2012). Expression pattern of drought stress marker genes in soybean roots
806 under two water deficit systems. Genet. Mol. Biol. 35, 212–221. doi: 10.1590/S1415-
807 47572012000200003
808
809 Nogue, F., Mara, K., Collonnier, C., and Casacuberta, J. M. (2016). Genome engineering and
810 plant breeding: impact on trait discovery and development. Plant Cell. Rep. 35, 1475–1486.
811 doi: 10.1007/s00299-016-1993-z.
812 Ogawa, A., Kawashima, C., and Yamauchi, A. (2005). Sugar accumulation asevere the
813 seminal root axis, as affected by osmotic stress in maize: a possible physiological basis for
814 plastic lateral root development. Plant Prod. Sci. 8, 173–180. doi:
815 https://doi.org/10.1626/pps.8.173
816
817 Olvera-Carrillo, Y., Campos, F. M., Reyes, J. R., Garciarrubio, A., and Covarrubias, A. A.
818 (2010). Functional analysis of the group 4 late embryogenesis abundant proteins reveals their

819 relevance in the adaptive response during water deficit in arabidopsis. Plant Physiol. 154,
820 373-390. doi: 10.1104/pp.110.158964
821
822 Oya, T., Nepomuceno, A. L., Neumaier, N., Farias, J. R. B., Tobita, S., and ITO, O. (2004).
823 Drought tolerance characteristics of Brazilian soybean cultivars – evaluation and
824 characterization of drought tolerance of various Brazilian soybean cultivars in the field. Plant
825 Prod. Sci. 7, 129-137. doi: https://doi.org/ 10.1626/pps.7.129
826
827 Pantuwan, G., Fukai, S., Cooper, M., and Rajatasereekul, J. C. O. (2002). Yield response of
828 rice (Oryza sativa L.) genotypes to drought under rainfed lowland: 3. Plant factors
829 contributing to drought resistance. Field Crops Res. 73, 181–200. doi: 10.1016/S0378-
830 4290(01)00194-0
831
832 Paul, S., Gayen, D., Datta, S. K., and Datta, K. (2015). Dissecting root proteome of transgenic
833 rice cultivars unravels metabolic alterations and accumulation of novel stress responsive
834 proteins under drought stress. Plant Sci. 234, 133–143. doi: 10.1016/j.plantsci.2015.02.006.
835
836 Petrov, V., Hille, J., Mueller-Roeber, B., and Gechev, T. S. (2015). ROS-mediated abiotic
837 stress-induced programmed cell death in plants. Front. Plant Sci. 6, 69. doi:
838 10.3389/fpls.2015.00069
839
840 Pfaffl, M. W., Horgan, G. W., and Dempfle, L. (2002). Relative expression moderateware
841 tool (REST©) for group-wise comparison and statistical analysis of relative expression results
842 in real-time PCR. Nucleic Acids Res. 30, e36. doi: https://doi.org/10.1093/nar/30.9.e36
843
844 Phukan, U. J., Jeena, G. S., and Shukla, R. K. (2016). WRKY transcription factors: molecular
845 regulation and stress responses in plants. Front. Plant Sci. 7, 760. doi:
846 http://doi.org/10.3389/fpls.2016.00760
847
848 Qin, Y., Tian, Y., and Liu, X. (2015). A wheat salinity-induced WRKY transcription factor
849 TaWRKY93 confers multiple abiotic stress tolerance in Arabidopsis thaliana. Biochem.
850 Biophys. Res. Commun. 464, 428–433. doi: 10.1016/j.bbrc.2015.06.128
851
852 Rashid, M., Guangyuan, H., Guangxiao, Y., Hussain, J., and Xu, Y. (2012). AP2/ERF
853 transcription factor in rice: genome-wide canvas and syntenic relationships between monocots
854 and eudicots. Evol. Bioinform. 8, 321–355. doi: 10.4137/EBO.S9369
855
856 Rayko, E., Maumus, F., Maheswari, U., Jabbari, K., and Bowler, C. (2010). Transcription
857 factor families in ferred from genome sequences of photosynthetics tramenopiles. New
858 Phytol. 188, 52–66. doi: 10.1111/j.1469-8137.2010.03371.x
859
860 Riechmann, J. L., and Meyerowitz, E. M. (1998). The AP2/EREBP family of plant
861 transcription factors. Biol. Chem. 379, 633–646. doi: 10.1515/bchm. 1998.379.6.633
862 Rinerson, C. I., Scully, E. D., Palmer, N. A., Donze-Reiner, T., Rabara, R. C., Tripathi, P., et
863 al. (2015). The WRKY transcription factor family and senescence in switchgrass. BMC
864 Genomics. 16, 912. doi: 10.1186/s12864-015-2057-4
865
866 Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2010). edgeR: a bioconductor package
867 for differential expression analysis of digital gene expression data. Bioinformatics. 26, 139–
868 140. doi: 10.1093/bioinformatics/ btp616

Rodrigues, F. A., Fuganti-Pagliarini, R., Marcolino-Gomes, J., Nakayama, T. J., Molinari, H. B. C., Lobo, F. P., et al. (2015). Daytime soybean transcriptome fluctuations during water deficit stress. BMC Genomics. 16, 505.

Rodrigues, F. A., Marcolino-Gomes, J., Carvalho, J. F. C., Nascimento, L. C. do, Neumaier, N., Farias, J. R., et al. (2012). Subtractive libraries for prospecting differentially expressed genes in the soybean under water deficit. Genet. Mol. Biol. 35, 304–314. doi: 10.1590/S1415-47572012000200011

Rolla, A. A. de P., Carvalho, J. de F. C., Fuganti-Pagliarini, R., Engels, C., Rio, A. do, Marin, S. R., et al. (2014). Phenotyping soybean plants transformed with rd29A:AtDREB1A for drought tolerance in the greenhouse and field. Transgenic. Res. 23, 75–87. doi: 10.1007/S21248-013-9723-6

Roy, S. (2016). Function of MYB domain transcription factors in abiotic stress and epigenetic control of stress response in plant genome. Plant Signal. Behav. 11, e1117723. doi: 10.1080/15592324.2015.1117723

Rushton, D. L., Tripathi, P., Rabara, R. C., Lin, J., Ringler, P., Boken, A. K., et al. (2012). WRKY transcription factors: key components in abscisic acid signalling. Plant Biotechnol. J. 10, 2–11. doi: 10.1111/j.1467-7652.2011.00634.x

Sakuma, Y., Liu, Q., Dubouzet, J. G., Abe, H., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2002). DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. Biochem. Biophys. Res. Commun. 290, 998–1009. doi: https://doi.org/10.1006/bbrc.2001.6299

Salminen, A., Kauppinen, A., and Kaarniranta, K. (2015). 2-Oxoglutarate-dependent dioxygenases are sensors of energy metabolism, oxygen availability, and iron homeostasis: potential role in the regulation of aging process. Cell Mol. Life Sci. 72, 3897-3914. doi: 10.1007/s00018-015-1978-z

Santos, E. L. dos, Cattelan, A. J., Prete, C. E. C., Neumaier, N., Oliveira, M. C. N. de, Farias, J. R. B., et al. (2014). Níveis de disponibilidade hídrica sobre componentes de produção e rendimento de cultivares de soja. Global Sci. Technol. 7, 1-11. doi: 10.14688/1984-3801/gst.v7n3p1-11

Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. Nat. 463, 178–183. doi: 10.1038/nature08670

Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998). SMART, a Simple Modular ArchitectureResearch Tool: identification of signaling domains. Proc. Natl. Acad. Sci. U. S. A. 95, 5857–5864. doi: 10.1073/pnas.95.11.5857

Seeve, C. M., Cho, I.-J., Hearne, L. B., Srivastava, G. P., Joshi, T., Smith, D. O., et al. (2017). Water deficit-induced changes in transcription factor expression in maize seedlings. Plant Cell Environ. 40, 686–701. doi: 10.1111/pce.12891.

Shinozaki, K., and Yamaguchi-Shinozaki, K. (2007). Gene networks involved in drought stress response and tolerance. J. Exp. Bot. 58, 221–227. doi: https://doi.org/10.1093/jxb/erl164

Silva, E. C., Nogueira, R. J. M. C., Vale, F. H. A., Melo, N. F., and Araujo, F. P. (2009). Water relations and organic solutes production in four umbu tree (Spondias tuberosa) genotypes under intermittent drought. Braz. J. Plant Physiol. 21, 43-53. doi: 10.1590/S1677-04202009000100006

Singh, D., and Laxmi, A. (2015). Transcriptional regulation of drought response: a tortuous network of transcriptional factors. Front. Plant Sci. 6, 895. doi: 10.3389/fpls.2015.00895

Solano, R., Stepanova, A., Chao, Q., and Ecker, J. R. (1998). Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENEINSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. Genes Dev. 12, 3703–3714. doi: 10.1101/gad.12.23.3703

Song, J., Xing, Y., Munir, S., Yu, C., Song, L., Li, H., et al. (2016). An ATL78-Like RING-H2 finger protein confers abiotic stress tolerance through. Front. Plant Sci. 7, 1305. doi: 10.3389/fpls.2016.01305

Sun, J., Hu, W., Zhou, R., Wang, L., Wang, X., Wang, Q., et al. (2015). The Brachypodium distachyon BdWRKY36 gene confers tolerance to drought stress in transgenic tobacco plants. Plant Cell Rep. 34, 23–35. doi: 10.1007/s002990 14-1684-6

Takahashi, F., Sato-Nara, K., Kobayashi, K., Suzuki, M., and Suzuki, H. (2003). Sugar-induced adventitious roots in Arabidopsis seedlings. J. Plant Res. 116, 83-91. doi: 10.1007/s10265-002-0074-2

Tan, X., Yan, S., Tan, R., Zhang, Z., Wang, Z., and Chen, J. (2014). Characterization and expression of a D=GDSL-Like Lipase gene from Brassica napus in Nicotiana benthamiana. Protein. J. 33, 18-23. doi: 10.1007/s10930-013-9532-z

Tholkappian, P., Prakash, M., and Sundaram, M. D. (2001). Effect of AM-fungi on proline, nitrogen and pod number of soybean under moisture stress. Indian J. Plant. Physiol. 6, 98–99.

Trivedi, D. K., Ansari, M. W., and Tuteja, N. (2013). Multiple abiotic stress responsive rice cyclophilin: (OsCYP-25) mediates a wide range of cellular responses. Commun. Integr. Biol. 6: e25260. doi: http://dx.doi.org/10.4161/ cib.25260.

Urbanczyk-Wochniak, E., Luedemann, A., Kopka, J., Selbig, J., Roessner-Tunali, U., Willmitzer, L., et al. (2003). Parallel analysis of transcript and metabolite profiles: a new approach in systems biology. EMBO Rep. 4, 989–993. doi: 10.1038/sj.embor.embor944

Varshney, R. K., Bansal, K. C., Aggarwal, P. K., Datta, S. K., and Craufurd, P. Q. (2011). Agricultural biotechnology for crop improvement in a variable climate: hope or hype? Trends. Plant Sci. 16, 363–371. doi: 10.1016/j.tplants.2011.03. 004

Wang, S., Yang, S., Yin, Y., Guo, X., Wang, S., and Hao, D. (2009). An in silico strategy identified the target gene candidates regulated by dehydration responsive element binding

proteins (DREBs) in Arabidopsis genome. Plant Mol. Biol. 69, 167–178. doi: 10.1007/s11103-008-9414-5

Wani, S. H., and Gosal, S. S. (2011). Introduction of OsglyII gene into Oryza sativa for increasing salinity tolerance. Biol. Plant. 55, 536–540. doi: 10.1007/ s10535- 011-0082-y

Wani, S. H., Singh, N. B., Devi, T. R., Haribhushan, A., Jeberson, S. M., Malik, C. P., et al. (2013). Engineering abiotic stress tolerance in plants:extricating regulatory gene complex, in Conventional and Non-Conventional Interventions in Crop Improvement, eds. C. P. Malik, G. S. Sanghera, and S. H. Wani (New Delhi: CABI), 1–19.

Wei, W., Zhang, Y. X., Han, L., Guan, Z. Q., and Chai, T. Y. (2008). A novel WRKY transcriptional factor from Tlaspi caerulescens negatively regulates the osmotic stress tolerance of transgenic tobacco. Plant Cell Rep. 27, 795–803. doi: 10.1007/s00299-007-0499-0

Wen, Z., Yao, L., Wan, R., Li, Z., Liu, C., and Wang, X. (2015). Ectopic expression in Arabidopsis thaliana of an NB-ARC encoding putative disease resistance gene from wild chinese vitis pseudoreticulata enhances resistance to phytopathogenic fungi and bacteria. Front. Plant Sci. 6, 1087. doi: 10.3389/fpls. 2015.01087

Woodrow, P., Pontecorvo, G., Ciarmiello, L. F., Annunziota, M. G., Fuggi, A., and Carillo, P. (2012). Transcription factors and genes in abiotic stress, in Crop Stress Management: Perspective and Strategies, B. S. Venkateswarlu, A. K. Shanker, and M. C. Maheswari (Netherlands: Springer), 317-357.

Xie, Z., Zhang, Z. L., Zou, X. L., Huang, J., Ruas, P., Thompson, D., et al. (2005). Annotations and functional analyses of the rice WRKY gene superfamily reveal positive and negative regulators of abscisic acid signaling in aleurone cells. Plant Physiol. 137, 176–189. doi: 10.1104/pp.104.054312

Yamaguchi-Shinozaki, K., and Shinozaki, K. (2006). Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. Annu. Rev. Plant Biol. 57, 781–803. doi: 10.1146/annurev.arplant.57.032905. 105444

Yan, H., Jia, H., Chen, X., Hao, L., An, H., and Guo, X. (2014). The cotton WRKY transcription factor GhWRKY17 functions in drought and salt stress in transgenic Nicotiana benthamiana through ABA signaling and the modulation of reactive oxygen species production. Plant Cell Physiol. 55, 2060–2076. doi: 10.1093/pcp/pcu133

Yang, Z., Tian, L., Latoszek-Green, M., Brown, and D., Wu, K. (2005). Arabidopsis ERF4 is a transcriptional repressor capable of modulating ethylene and abscisic acid responses. Plant Mol. Biol. 58, 585–596. doi: 10.1007/s111 03-005-7294-5

Yoshimura, K., Masuda, A., Kuwano, M., Yokota, A., and Akashi, K. (2008). Programmed proteome response for drought avoidance/tolerance in the root of a C-3 xerophyte (wild watermelon) under water deficits. Plant Cell Physiol. 49, 226–241. doi: 10.1093/pcp/pcm180

Yu, F., Huaxia, Y., Lu, W., Wu, C., Cao, X., and Guo, X. (2012). GhWRKY15, a member of the WRKY transcription factor family identified from cotton (Gossypium hirsutum L.), is

involved in disease resistance and plant development. BMC Plant Biol. 12, 144. doi: 10.1186/1471-2229-12-144

Yu, Y., Wang, N., Hu, R., and Xiang, F. (2016). Genome-wide identifcation of soybean WRKY transcription factors in response to salt stress. Springerplus. 5, 920. doi: 10.1186/s40064-016-2647-x

Zhou, Q. Y., Tian, A. G., Zou, H. F., Xie, Z. M., Lei, G., Huang, J., et al. (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, 486–503. doi: 10.1111/j.14677652. 2008.00336.x

Zhu, J.-K. (2016). Abiotic stress signaling and responses in plants. Cell. 167, 313-324. doi: 10.1016/j.cell.2016.08.029

Zong, W., Tang, N., Yang, J., Peng, L., Ma, S., Xu, Y., et al. (2016). Feedback regulation of ABA signaling and biosynthesis by a bZIP transcription factor targets drought-resistance-related genes. Plant Physiol. 171, 2810–2825. doi: 10.1104/pp.16.00469

Table 1 - Differentially gene expression in leaves and roots of BR 16 and Embrapa 48 at moderate and severe stress condition.

Sample	Reads after adapter removal	Mapped reads [0..2 mis]	% of mapped reads [0..2 mis] (Mean)
Embrapa 48 leaf	12,480,047	6,139,313	49.2
Embrapa 48 root	12,241,551	6,137,155	50.2
BR 16 leaf	11,305,266	5,414,745	47.6
BR 16 root	11,150,778	5,555,411	49.7
Total	47,177,642	23,246,624	-

Table 2 - Embrapa 48 and BR 16 general RNA-seq analysis.

	Up-regulated genes number		**Down-regulated genes number**	
TRATAMENT	**Embrapa 48**	**BR 16**	**Embrapa 48**	**BR 16**
Leaf (moderate stress)	4693	2009	1238	2128
Leaf (severe stress)	4247	2311	2694	5167
Root (moderate stress)	2479	2274	3162	2165
Root (severe stress)	3031	2834	3659	3629
TOTAL	14450	9428	10753	13089
Embrapa 48 total genes		25203		
BR 16 total genes		22517		

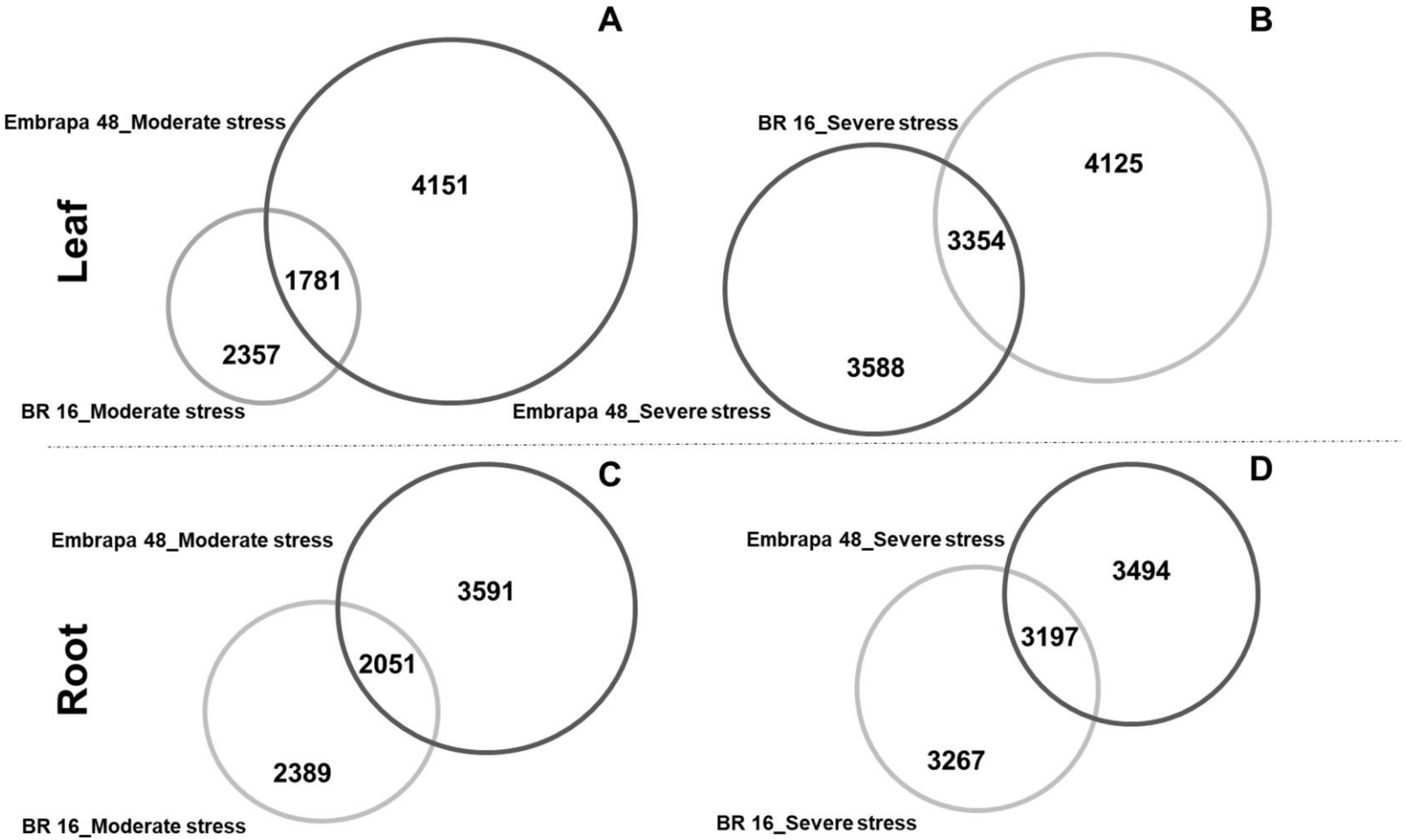


Figure 1 - Venn's diagram. Number of genes differentially expressed in both libraries, highlighting common genes between cultivars. (A) BR 16 and Embrapa 48 leaf in moderate stress library; (B) BR 16 and Embrapa 48 leaf in severe stress; (C) BR 16 and Embrapa 48 root in moderate stress library; (D) BR 16 and Embrapa 48 root in severe stress library. The size of the circle follows a proportion related to the number it represents. Dark gray circles represent the cultivar Embrapa 48 and light gray the BR 16.

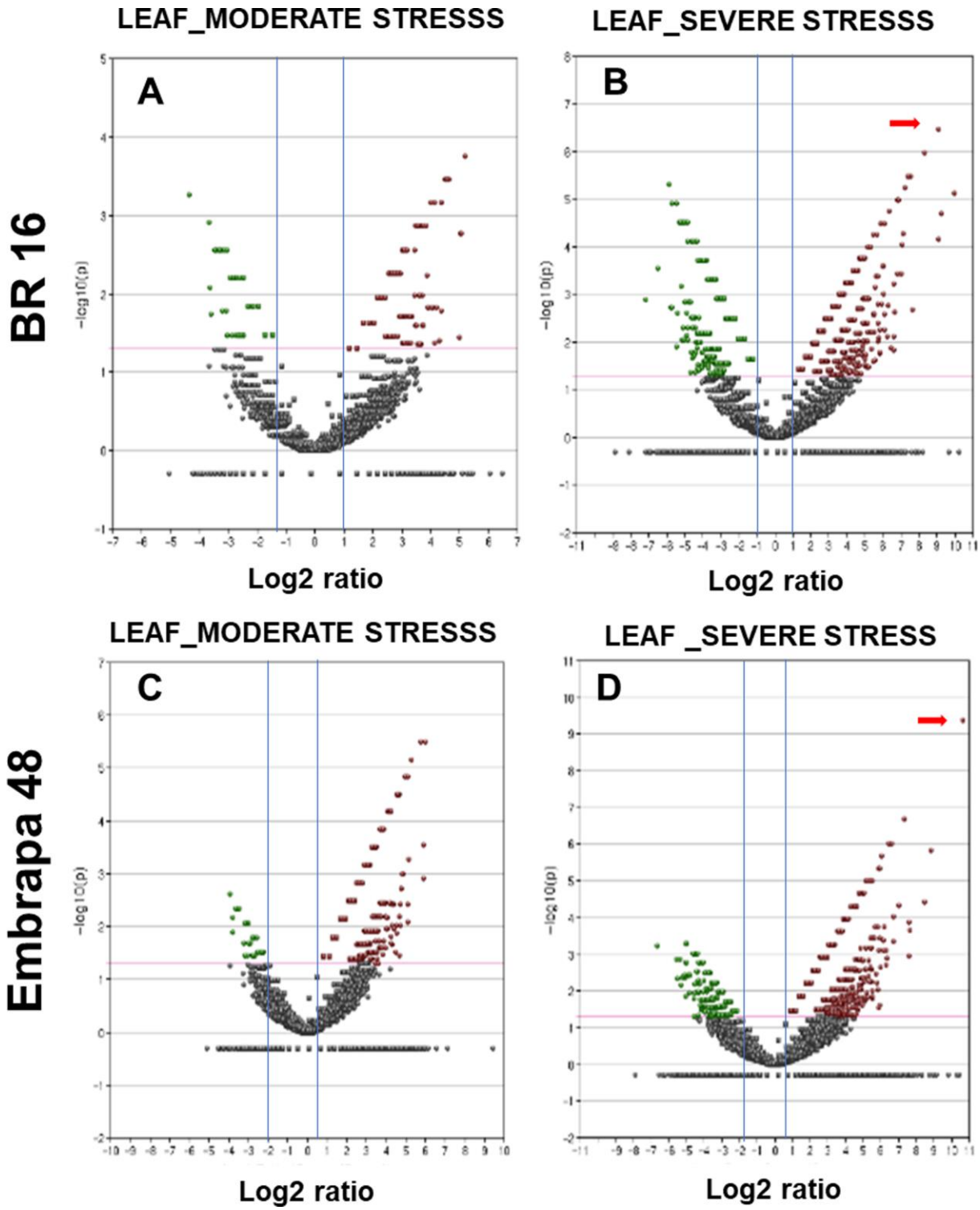


Figure 2 - Volcano plots. Expression data (fold-change) were plotted using log_2 scale (x-axis) and $-log_{10}$ transformation of the p-value (y-axis). Datasets were filtered to remove genes with low expression levels (blue line on the X-axis ≥1 and -1) and (red lines on the Y-axis), the red line is the threshold applied to delimitate a significance cut off ($p<0.001$). The red points were used to represent genes up regulated and green points to genes down regulated. The red arrows indicate the top highest expression in same genes. (A) leaf of BR 16 during moderate stress library; (B) leaf of BR 16 during stress severe stress library; (C) leaf of Embrapa 48 during moderate stress library; (D) leaf of Embrapa 48 during severe stress library.

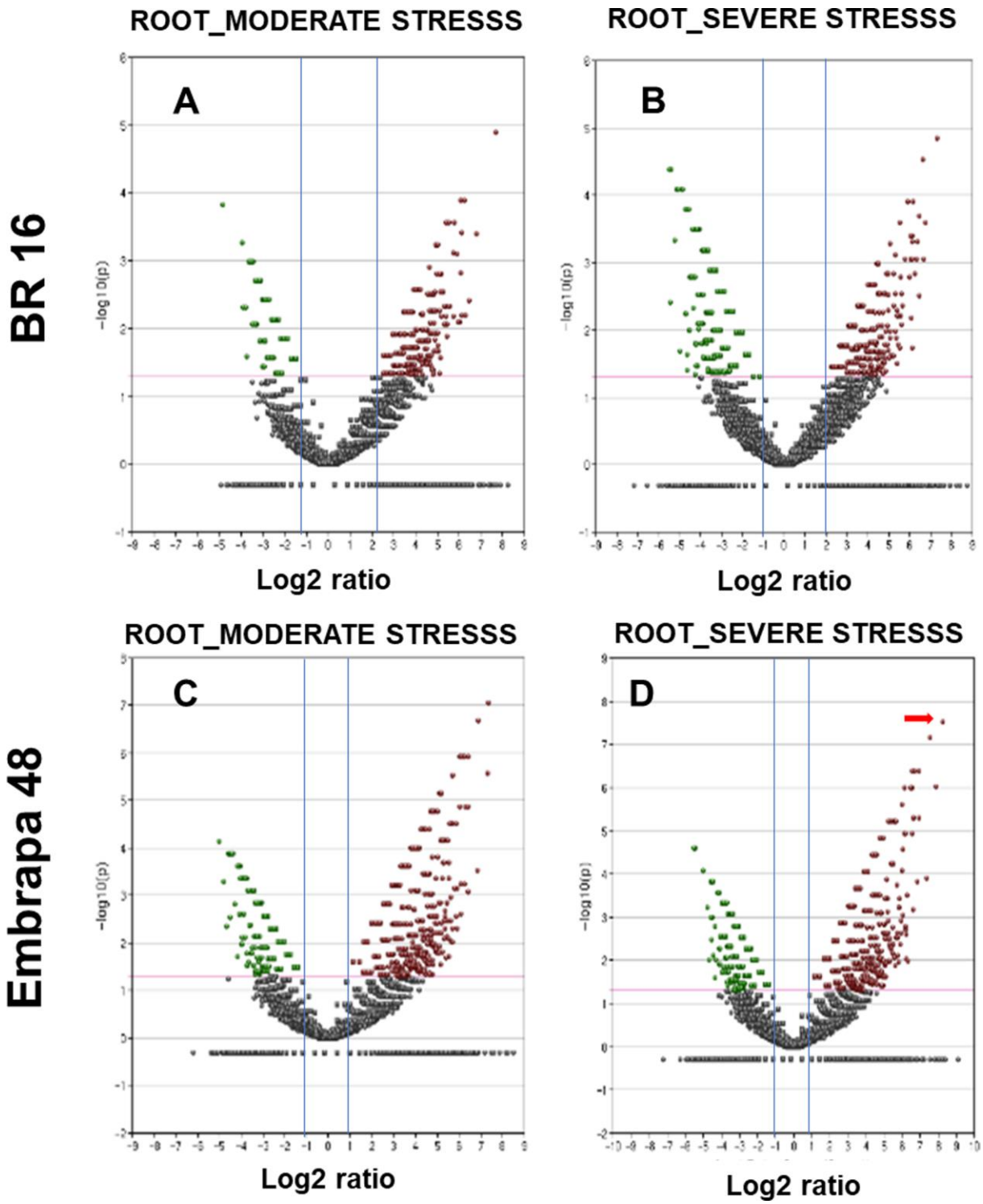


Figure 3 - Volcano plots. Expression data (fold-change) were plotted using $\log_2$ scale (x-axis) and $-\log_{10}$ transformation of the p-value (y-axis). Datasets were filtered to remove genes with low expression levels (blue line on the X-axis ≥1 and -1) and (red lines on the Y-axis), the red line is the threshold applied to delimitate a significance cut off ($p<0.001$). The red points were used to represent up-regulated genes and green points to down-regulated genes. The red arrows indicate the top highest expression in same genes. (A) roots of BR 16 during moderate stress library; (B) roots of BR 16 during severe stress library); (C) roots of Embrapa 48 during moderate stress library); (D) roots of Embrapa 48 during severe stress library.

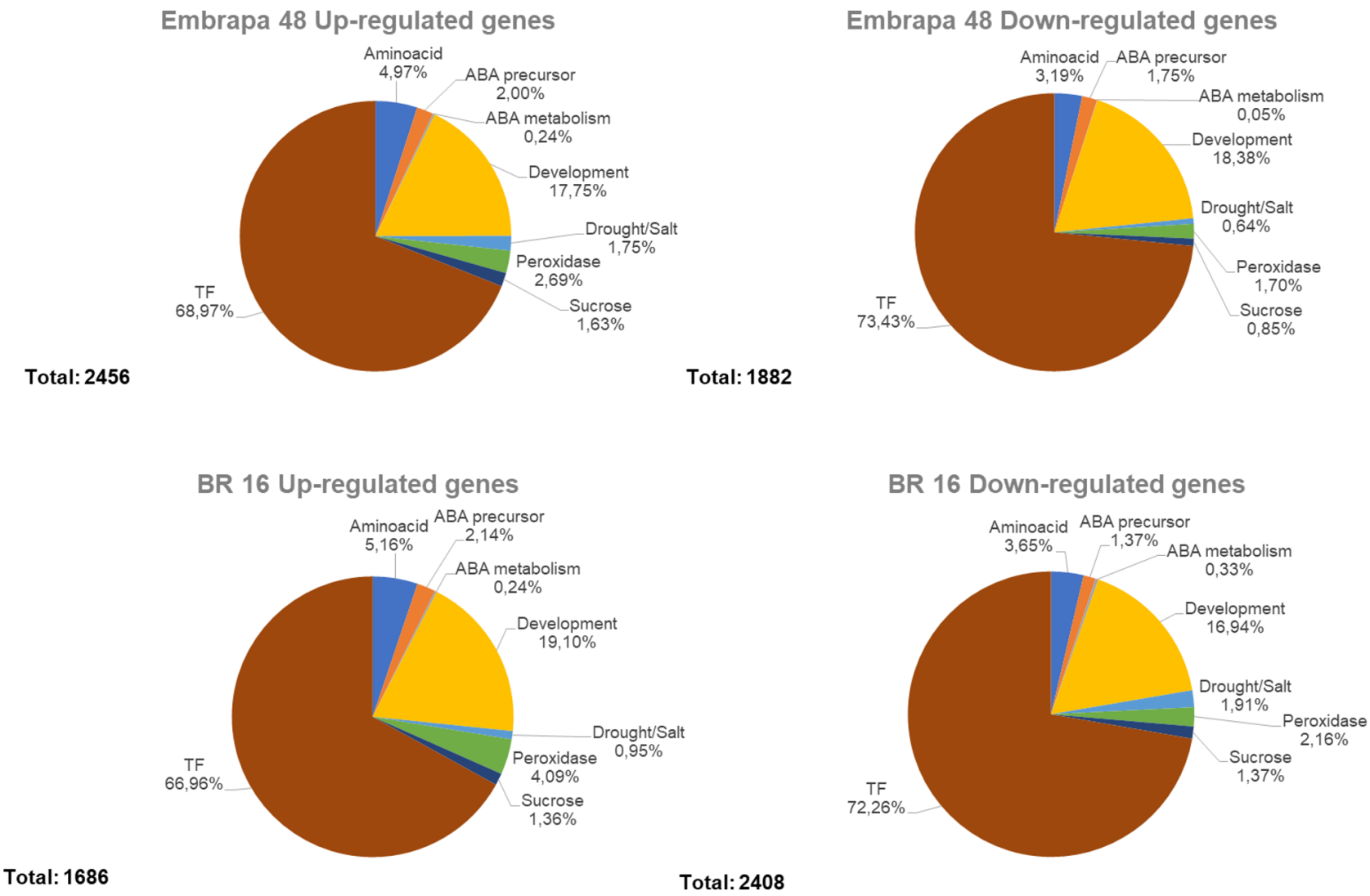


Figure 4 - Graphic representation of eight gene categories (enriched biological process from MapMan 3.6 ORC), analyzed in up and down-regulated gene groups to the cultivars BR 16 and Embrapa 48.

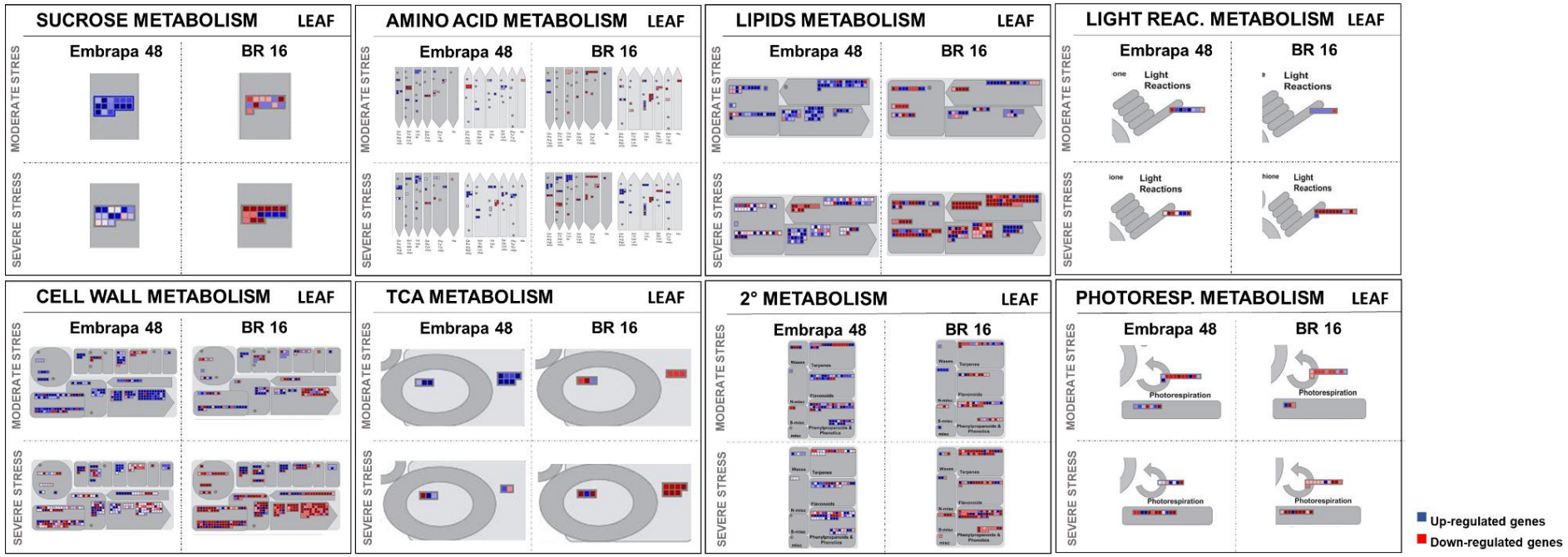


Figure 5 - Metabolism_overview_mapping (MapMan 3.6 ORC). BR 16 leaf moderate stress; BR 16 leaf severe stress; Embrapa 48 leaf moderate stress; Embrapa 48 leaf severe stress; Genes that were differentially expressed in response to water deficit were mapped to specific stress-related gene group. The color scale showed the $\log_2$ fold change: red = down-regulated and blue = up-regulated. Light reac. is the abbreviation of "light reaction" and photresp. is the abbreviation of "photorespiration".

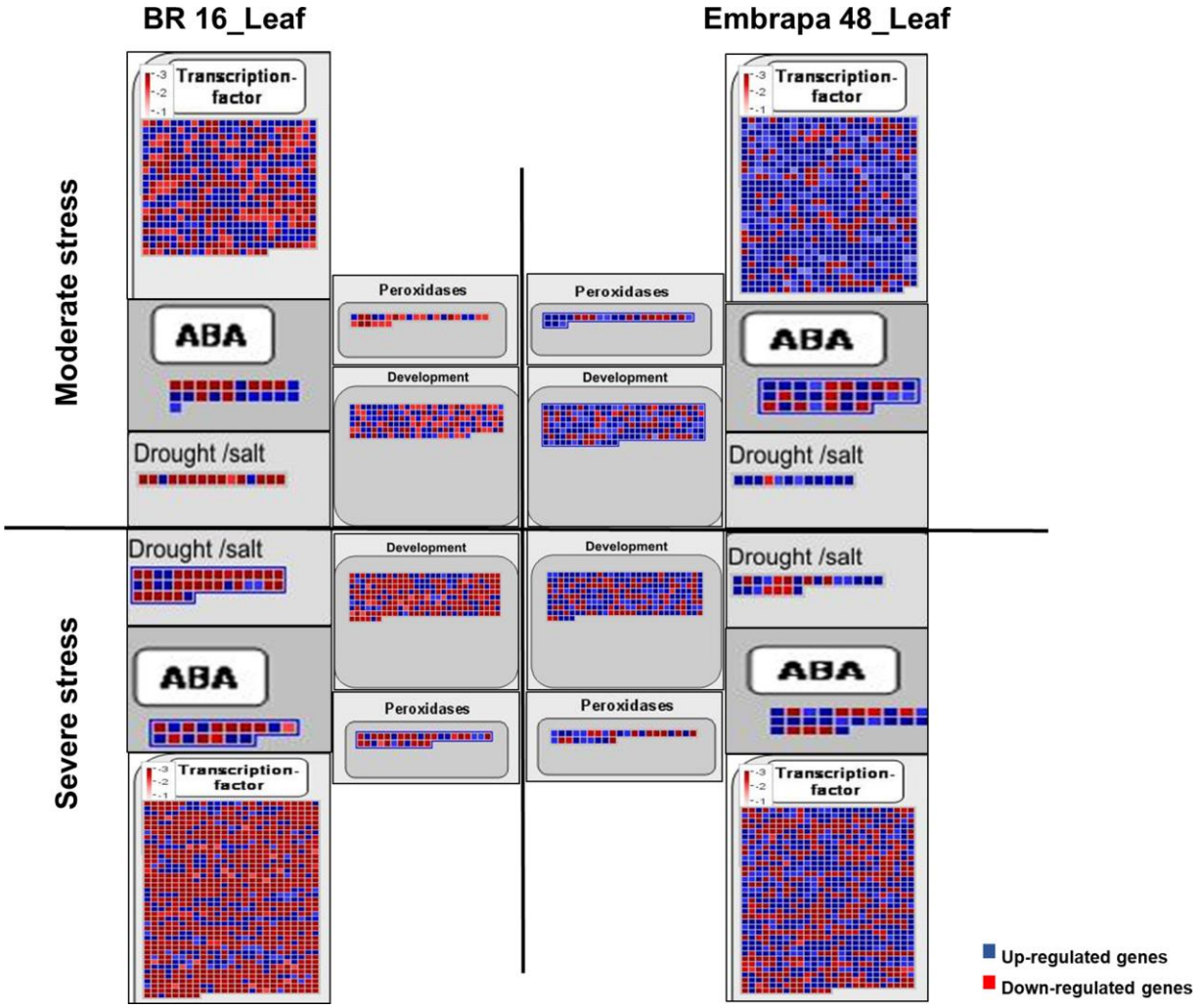


Figure 6 - Modified maps of MapMan 3.6 ORC. BR 16 leaf in moderate stress; BR 16 leaf in severe stress; Embrapa 48 leaf in moderate stress; Embrapa 48 leaf in severe stress. Functional roles triggered in soybean plants under water deficit stress. Genes that were differentially expressed in response to water deficit were mapped to specific stress-related pathways. The color scale show the $\log_2$ fold change: red = down-regulated and blue = up-regulated.

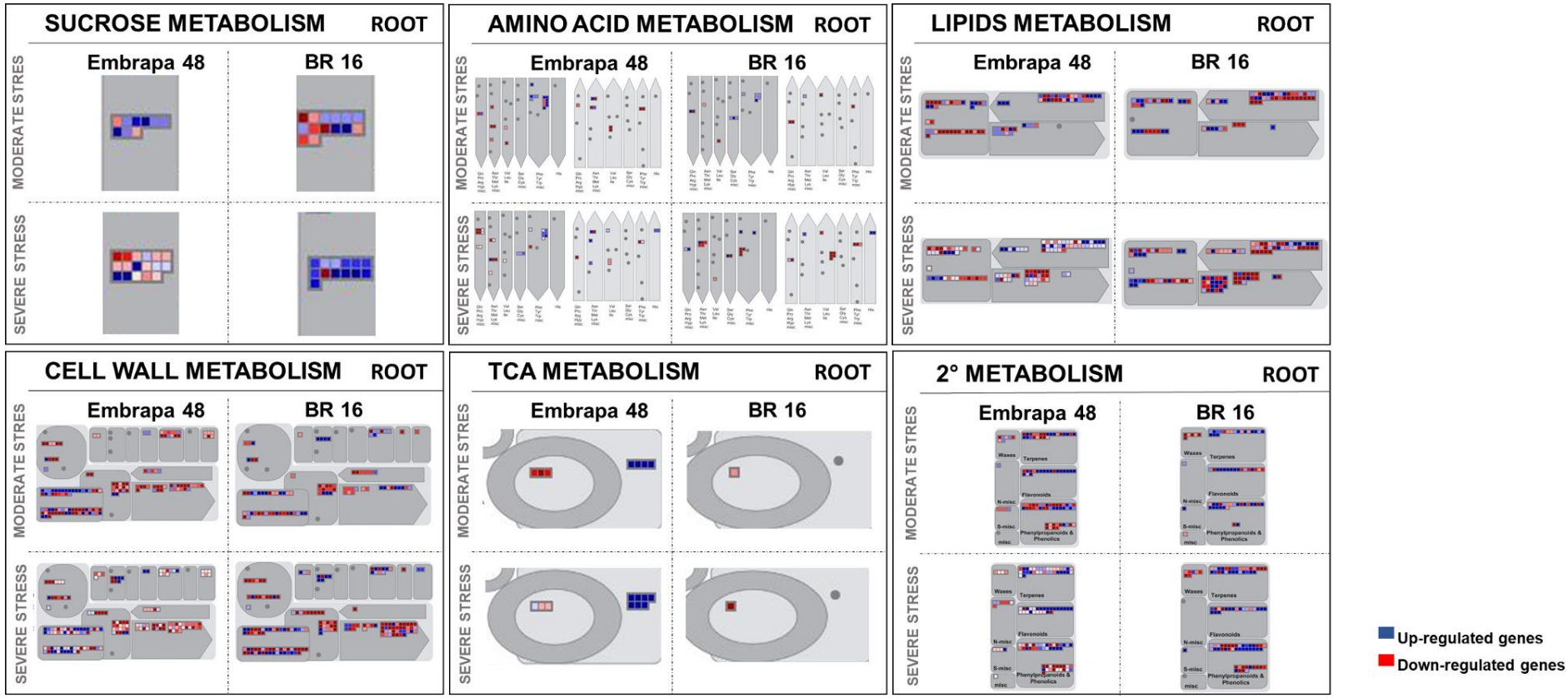


Figure 7 - Metabolism_overview_mapping (MapMan 3.6 ORC). BR 16 root in moderate stress; BR 16 root in severe stress; Embrapa 48 root in moderate stress; Embrapa 48 root in severe stress. Genes that were differentially expressed in response to water deficit were mapped to specific stress-related gene group. The color scale show the $\log_2$ fold change: red=down-regulated and blue=up-regulated.

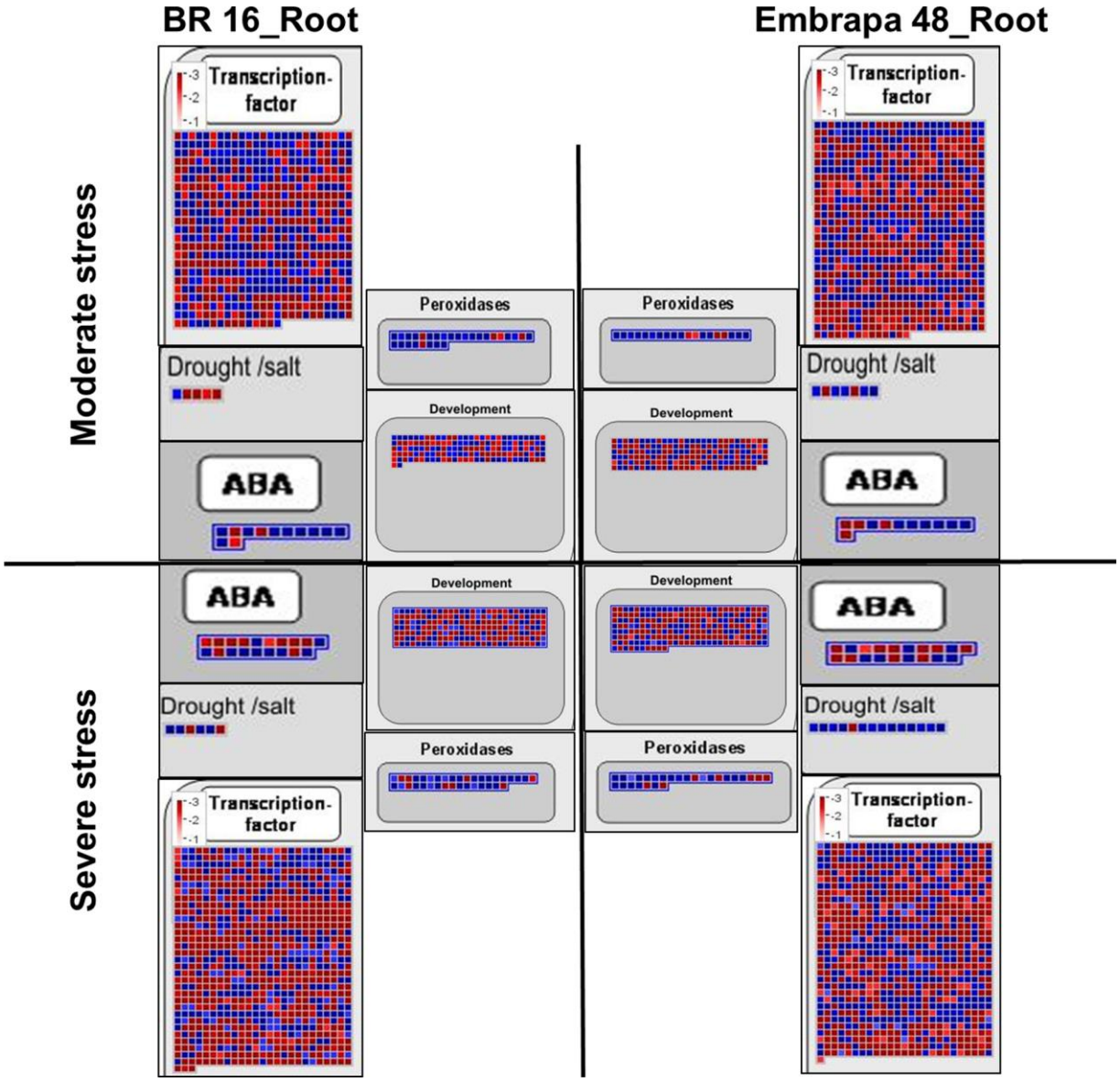


Figure 8 - Modified maps of MapMan 3.6 ORC. BR 16 root in moderate stress; BR 16 root in severe stress; Embrapa 48 root in moderate stress; Embrapa 48 root in severe stress. Functional roles triggered in soybean plants under water deficit stress. Genes that were differentially expressed in response to water deficit were mapped to specific stress-related pathways. The color scale show the $\log_2$ fold change: red = down-regulated and blue = up-regulated.

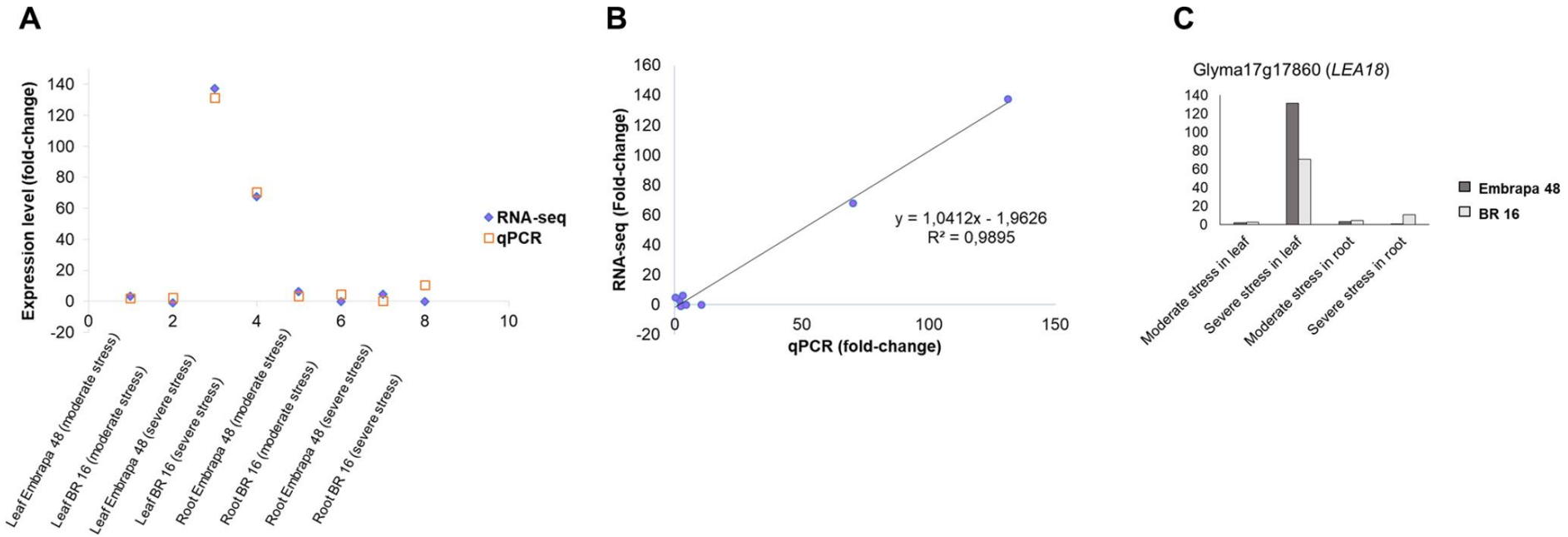
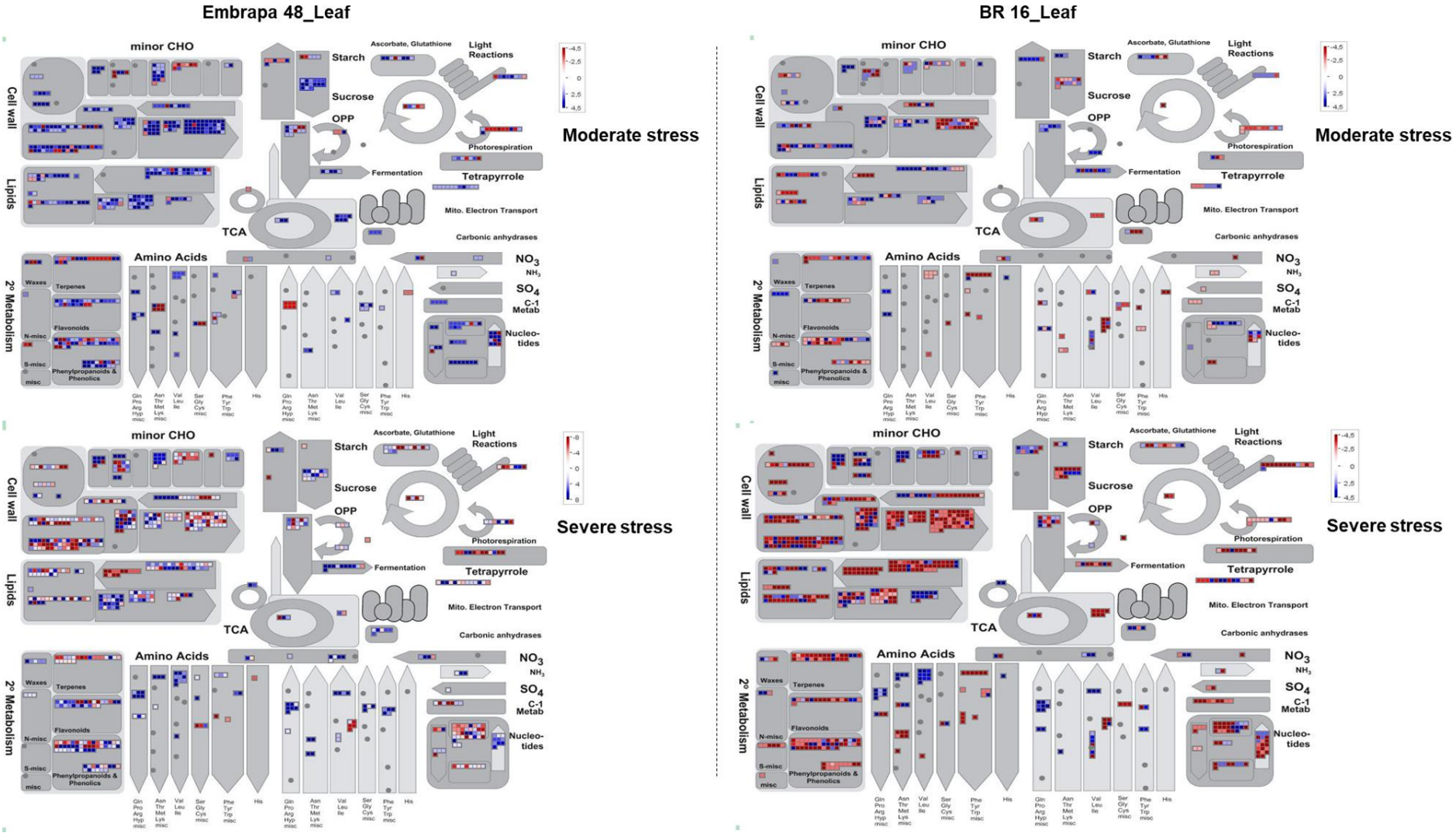
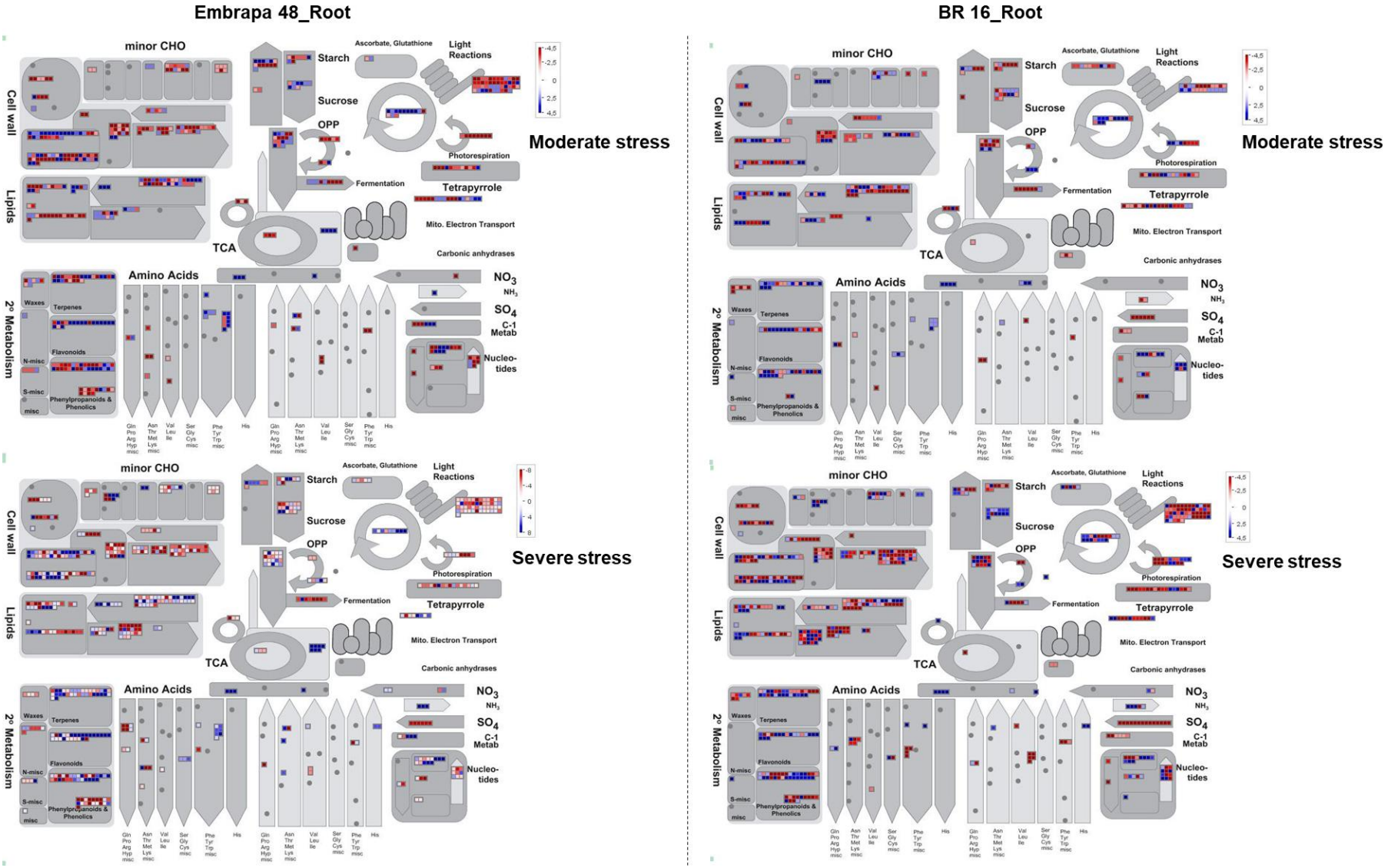


Figure 9 - RT-qPCR validation of differentially expressed genes in soybean under water deficit stress identified in RNA-seq analysis. (A) represents RNA-seq and RT-qPCR results to *LEA18* gene validated in Embrapa 48 leaf in moderate and severe library; Embrapa 48 root in moderate and severe library; BR 16 leaf in moderate and severe library; BR 16 root in moderate and severe library. (B) represents the correlation of the fold-change analyzed by between RNA-seq (y-axis) and qPCR (x-axis). (C) RT-qPCR to *LEA18* gene. Fold-change in (y-axis) and treatments to each gene (x-axis).



S1) Figure 1 Metabolism_overview_mapping of leaf (MapMan 3.6 ORC).



S2) Figure 1 Metabolism_overview_mapping of root (MapMan 3.6 ORC).

S3) Table 1 - ABA Regulatory Genes with differential expression in cultivars BR 16 and Embrapa 48 (the genes in purple represent the commons Glymas between BR 16 and Embrapa 48).

	Wm82.a1.v1	Wm82.a2.v1	PFAM_Description
BR 16[a]	Glyma08g42340	Glyma.08g310600	TB2/DP1, HVA22 family
	Glyma11g12740	Glyma.11g119700	TB2/DP1, HVA22 family
	Glyma12g04920	Glyma.12g044900	TB2/DP1, HVA22 family
	Glyma18g12530	Glyma.18g105400	TB2/DP1, HVA22 family
	Glyma19g20090	Glyma.19g067900	bZIP transcription factor
	Glyma19g30230	Glyma.19g122800	bZIP transcription factor
	Glyma20g23510	Glyma.20g102100	GRAM domain
Embrapa 48[b]	Glyma04g02730	Glyma.04g024900	TB2/DP1, HVA22 family
	Glyma01g05370	Glyma.01g045400	TB2/DP1, HVA22 family
	Glyma04g37440	Glyma.04g197600	TB2/DP1, HVA22 family
	Glyma06g20570	Glyma.06g192700	TB2/DP1, HVA22 family
	Glyma06g33130	Glyma.06g229400	GRAM domain
	Glyma09g29380	Glyma.09g164100	GRAM domain
	Glyma10g43290	Glyma.10g286800	GRAM domain
	Glyma10g43300	Glyma.10g286900	GRAM domain
	Glyma16g05520	Glyma.16g050700	GRAM domain
	Glyma17g10730	Glyma.17g098800	TB2/DP1, HVA22 family
	Glyma20g10600	Glyma.20g049200	bZIP transcription factor
	Glyma20g23510	Glyma.20g102100	GRAM domain
	Glyma0273s00200	no correspondence	NA
Common to both cultivar	Glyma03g37140	Glyma.03g213300	TspO/MBR family
	Glyma08g10190	Glyma.08g096200	Retinal pigment epithelial membrane protein
	Glyma08g41060	Glyma.08g229300	TB2/DP1, HVA22 family
	Glyma14g03090	Glyma.14g027000	TB2/DP1, HVA22 family
	Glyma15g40070	no correspondence	NA
	Glyma18g15480	Glyma.18g121600	TB2/DP1, HVA22 family

Glyma20g23500	Glyma.20g102600	GRAM domain	

[a] Gene expressed exclusively in BR16

[b] Gene expressed exclusively in Embrapa 48

S4) Table 1 - Genes selected for RNA-seq validation.

	Wm82.a1.v1	Wm82.a2.v1	Annotation	Embrapa 48 expression profile	BR 16 expression profile
Drought-responsive gene	Glyma17g17860	Glyma.17g164200	LEA 18 (Late embryogenesis abundant protein 18)	UP in leaf severe and moderate stress and UP in root moderate and severe stress time	UP in leaf severe stress and DOWN in leaf moderate stress

CAPÍTULO 3 – COMPARATIVE TRANSCRIPTIONAL PROFILE OF AP2/EREBP AND WRKY TRANSCRIPTION FACTORS IN SOYBEAN GENOTYPES WITH CONTRASTING WATER DEFICIT TOLERANCE IDENTITIES

Artigo preparado de acordo com as normas para submissão do periódico

Molecular Plant.

Comparative transcriptional dynamics of AP2/EREBP and WRKY families in leaf and root of soybean genotypes with contrasting water deficit tolerance identities

Transcriptional dynamics of AP2/EREBP and WRKY families

Rafaela Ribeiro Reis*[1], Juliana Marcolino-Gomes[2], Fabiana Aparecida Rodrigues[3], Silvana Rockenbach Marin[4]; Liliane Marcia Mertz-Henning[5], Alexandre Lima Nepomuceno[6]

[1] Londrina State University, Rodovia Celso Garcia Cid, Km 380. CEP 86051-900, Londrina, PR, Brazil.

[2] CNPq SEPN 507 BL B, - lt-2 térreo - Asa Norte - Brasília, DF - CEP 70740-522.

[3] Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso do Sul Rua Salime Tanure, s/n°. CEP 79400-000. Fone: (67) 3291-9600. Coxim - MS

[4] Embrapa Soja, Rodovia Carlos João Strass, Distrito de Warta, CEP 86001-970 PR, Brazil.

[5] Embrapa Soja, Rodovia Carlos João Strass, Distrito de Warta, CEP 86001-970 PR, Brazil.

[6] Embrapa Soja, Rodovia Carlos João Strass, Distrito de Warta, CEP 86001-970 PR, Brazil.

* Corresponding Author: rafaelarireis@gmail.com

ABSTRACT

Water deficit stress is considered the major threat to soybean production worldwide. Therefore, due to the genetic complexity that permeates drought tolerance traits, the identification of genes involved in water stress responses in plants is necessary. Here, new data about soybean plant responses under moderate and severe stress levels was assessed in two contrasting water deficit tolerance soybean cultivars, comparing leaves and roots. Genes from two important transcription factor families, AP2/EREBP and WRKY. Additionally, a difference in the expression pattern between both genotypes was detected, in which approximately 25,203 genes were modulated during water deficit stress in Embrapa 48 (tolerant), while for BR 16 (sensitive) it was 22,517. Evaluating only three categories (Transcription factor, AP2/EREBP and WRKY families), Embrapa 48 presented more than 609 up-regulated genes when compared to BR 16. The AP2/EREBP family exhibited greater activity in BR 16, reaching 10.72% among the up-regulated. Additionally, Embrapa 48 presented 3.99% down-regulated WRKY genes, while BR 16 showed 1.62%. The AP2/EREBP and WRKY families have distinct roles and expression profiles, often acting together and oppositely in the tolerance response. Finally, this study was identified and limited the top ten genes of the AP2/EREBP and WRKY gene families in the stress tolerance response under water deficit.

Keywords: RNA-seq, Transcription factors (TF), cross-talk, negative regulator.

1 INTRODUCTION

Soybean (*Glycine max* L. Merrill) crops have faced several losses in productivity due to water deficit stress. In stress conditions, plants develop molecular mechanisms of prevention and protection against damage caused by water restriction. Genetic products involving responses to drought can be classified into two groups: the functional and regulatory genes (De Carvalho, 2008; Molina et al., 2008). Different stress response strategies involving functional genes, which act directly against the damage caused by stress, are activated in an effort to counter the effects of adverse conditions, including morphological alterations in leaves, shoots and roots, variations in gene regulation, which triggers innumerable episodes of cascade expression induction, changes in stomatal conductance mediated by abscisic acid (ABA), osmoregulation from the accumulation of osmotically active molecules and the activity of antioxidant enzymes (Nakashima et al., 2014).

The precise control of the metabolic pathways and of this regulatory network of events that are triggered by stress are the key factors in determining the tolerance level of the plant. In this context, transcription factors (TFs) play an important role in the regulation of the process of signaling the perception of stress and transmitting it to the transduction pathway, activating defense genes. Regulatory genes are composed by several classes and families. Transcription factors (TF) are involved in the activation/inactivation of genes related to innumerable biological processes (Zhu, 2016). In *A. thaliana* nearly 6% of the proteome is dedicated to TF (Rayko et al., 2010).

The regulatory group primarily involve TFs (AP2/ERF, AREB, WRKY, MYC, NAC, bZIP, and MYB), protein kinases (MAPK- mitogen activated protein kinases, CDPK-calcium-dependent protein kinases) and others signals (Wani et al., 2013). The AP2/EREBP family includes a large group of plant-specific TFs and is characterized by the presence of highly conserved AP2/ERF (APETALA2/Ethylene Responsive Factor) binding domain, consisting of 58-60 amino acids. This TF interact directly with GCC-box and/or DRE/CRT (Dehydration-responsive element/C-repeat element) cis-acting elements at the promotor region (Yamaguchi-Shinozaki and Shinozaki, 2006). Riechmann and Meyerowitz (1998), in a study of genome-wide, identified a multitude of AP2/EREBP members of various plant species, such as 145 in *Arabidopsis thaliana (A. thaliana)*. AP2/EREBP superfamily is divided according to similarity of AP2/ERF domains in four major subfamilies: AP2 (Apetala 2), ERF (Ethylene response factor), DREB (Dehydration-responsive element-binding protein) and RAV (Related to ABI3/VP1) families.

The ERF and DREB/CBF (C-repeat binding factor) subfamilies have been largely studied due to their involvement in plant abiotic and biotic responses (Rashid et al., 2012; Nakashima et al., 2014). CBF/DREB associated transcription factors play an important role in genes regulation, responding to various stresses, such as cold, drought, high salinity and heat in different species (Kidokoro et al., 2015; Zong et al., 2016). According to Kidokoro et al. (2009), DRIP (*DREB*-Interacting Protein) and *PIF7* (Phytochrome-Interacting Factor 7) are proteins that negatively regulate *DREB* genes, and *DREB* 2A is negatively regulated by *GRF* (Growth-regulating factor 7).

WRKY also represents a transcription factor family with multiple roles in biotic/abiotic stress responses, as well as in developmental/physiological processes (Jiang et al., 2015). In a genome-wide study using SMART analysis, it was identified a total of 188

members containing WRKY (tryptophan, arginine, lysine and tyrosine) domains (Yu et al., 2016). This family is considered the most important TF family of plants and consists of 60 amino acid-severe four-stranded β-sheed WRKY DNA binding domain/s (DSD and Zing-finger) motifs. WRKY TFs interact with W-box (with core motif TTGACC/T) and clustered W-boxes present in the promoter of downstream genes. Recently studies of Chu et al. (2015) demonstrated that the overexpression of the *GhWRKY 41* gene enhanced salt tolerance and Li et al. (2015) observed that the overexpression of the *SpWRKY 1* gene enhanced drought tolerance, both in transgenic tobacco by studying *reactive oxygen species (ROS)* levels and stomatal conductance regulation.

WRKY has often been described as a booster of ROS production in cells. ROS, such as hydrogen peroxide, superoxide and hydroxyl radicals, have negative impact on the cell wall, inducing cell damage, oxidative stress and lipid peroxidation (Das and Roychoudhury, 2014). Thus, WRKY genes involved in this process can be recruited in extreme conditions, remaining inactive during the initial phases of stress. In soybean, three stress-induced genes were reported in Zhou et al. (2008): *GmWRKY13, GmWRKY21 and GmWRKY54*. *GmWRKY 21*-transgenic *A. thaliana* plants were tolerant to cold stress, whereas *GmWRKY 54* conferred salt and drought tolerance, possibly by *DREB 2A* and *STZ/Zat10* regulation. Additionally, the *A. thaliana* gene *AtWRKY 46* regulates stress hormones and developmental responses, facilitating growth of lateral roots in osmotic/salt stress through ABA signaling and auxin homeostasis control (Ding et al., 2015).

Over the years, several plants transcriptome of different tissues and species have aided in a better understanding of the gene expression regulation in response to adverse/stressful conditions. Although different technologies have been developed in soybean the molecular characterization for drought tolerance has been addressed only in few studies, leaving space to a vast unexplored gene possibility. The objective of this study was to gain insight into dynamics gene regulatory networks and cross-talk of AP2/EREBP and WRKY TF family in responses to water deficit stress in contrasting soybean genotypes Embrapa 48 (water deficit stress tolerant) and BR 16 (water deficit stress sensitive).

2 RESULTS

2.1 Comparative analysis of expression patterns of TF AP2/EREBP and WRKY families

A total of 47,177,642 reads were obtained after adapter removal, it was reduced to 23,246,624 reads (Table 1), when mapping considering only sequences with a maximum two mismatches in the first 32 bases. Furthermore, were obtained 55,787 mRNAs were obtained, whose, 51,322 were hit by at least 1 mRNA and 39,951 pass the low counts filter.

Two important TF families, AP2/EREBP and WRKY, were selected to study. Assuming the differentially expressed (activated and repressed) genes designated as a member of the AP2/EREBP family, in Embrapa 48 they represented, in number, 48.93% of the genes annotated as AP2/EREBP by the SoyBase database (reference genome Wm82.a1.v1), while in BR 16, it corresponded to 40.92%. For the class of WRKY genes, Embrapa 48 exhibited 28.46% of genes annotated for this family, with 19.23% represented in BR 16.

Among the up-regulated transcripts in the TF category, Embrapa 48 showed 85.86% members of TF class, 9.68% related to AP2/EREBP and 4.46% to the WRKY family, while BR 16 presented 84.07%, 10.72% and 5.21%, respectively (Figure 1). In relation to down-regulated genes classified in TF, BR 16 presented 3.36% more genes than Embrapa 48, reaching 4.63% of AP2/EREBP family and 1.62% of WRKY family. For Embrapa 48, the following results were obtained: 5.62% and 3.99 % of AP2/EREBP and WRKY family, respectively.

The expression of AP2/EREBP and WRKY familiy genes was suppressed or induced under water deficit stress for the genotypes, libraries and tissues (Figs. 2 and 2B). Furthermore, in general, the number of genes related to the AP2/EREBP family was higher than the WRKY family (Figs. 2A and 2B) for both cultivars.

Regarding the moderate and severe stress in leaves, Embrapa 48 presented more up-regulated genes (89 genes) in the AP2/EREBP family than BR 16 (33 genes) (Figure 2A). The difference between Embrapa 48 and BR16 was more intense in this gene family. For the WRKY classification, Embrapa 48 surpassed BR 16, exhibiting more negatively regulated genes (18 and 13 respectively). BR 16 presented 15 up-regulated genes in the moderate library, while Embrapa 48 showed 12 genes (Figure 2A).

Considering the moderate and severe stress in roots, the major contrasts between the cultivars was found for the AP2/EREBP family. Embrapa 48 exhibited 59 up-regulated genes more than BR16 and 16 down-regulated genes less than BR16. This discrepancy was not observed for the WRKY family, in which Embrapa 48 reached 45 up and 18 down-regulated genes and BR 16 reached 27 up and 13 down-regulated genes. In general, Embrapa 48 presented a more evident dynamic positive expression of both gene families than BR 16 (Figure 2A and B).

This can also be observed in the heat map data in Figure 3. In general, the dynamic expression of the AP2/EREBP family was more intense than WRKY, in which Embrapa 48 and BR 16 presented 248 and 213 differentially expressed genes, respectively. For WRKY, the cultivars were modulated to 114 and 86 genes, respectively (Figs. 3A and B). Considering AP2/EREBP, the gene activity was more intense in leaf than root for both gene cultivars (Figure 3A). Interestingly, in the AP2/EREBP category many genes did not occur in roots, which were evidenced by the black hole exposed by the heat map (Figure 3A). The dynamic of activation profile in leaf of AP2/EREBP and WRKY gene families varied significantly between cultivars. The tolerant cultivar demonstrated the highest number of differentially expression genes of both families when combining the two stress conditions (moderate and severe) (Figure 3A and B). Furthermore, for the AP2/EREBP family, Embrapa 48 presented a positively regulated differential gene expression than BR 16 in leaf in the two stress levels. However, considering the WRKY family, Embrapa 48 exhibited more differentially expression down-regulated genes in leaf and root when compared to BR 16 (Figure 3B).

2.2 Validation of genes expression

The RT-qPCR method to validate the RNA-seq data from the selection of the target genes considered a gene commonly described as responsive to water deficit stress, indicating the response of Late Embryogenesis Abundant protein 18 (*LEA18*) and some members of AP2/EREBP and WRKY families (Table 2 in S2). The RNA-seq values showed low variation compared with respective values from the RT-qPCR analyses (Figure 4B). The obtained expression pattern was widely compatible between both assays (R^2=0.9895) (Figure 4C). Additionally, gene from *LEA18* (Glyma17g17860) was highly up-regulated in severe stress in leaf, particularly in

Embrapa 48, reaching nearly twice the value of fc (131.3), while for BR 16 the fc value was 70.34 (Figure 4A).

Glyma08g01430 (*WRKY 75*) was highlighted in two treatments considering leaf in severe stress, in which Embrapa 48 obtained fc = 26.6, while for BR 16 fc = 9.94. Considering roots under severe, BR 16 exhibited fc = 2.52 and Embrapa 48 fc = 39.79. Glyma05g32040 (*AP2*) was highlighted in Embrapa 48 in leaf under severer stress (fc = 150.88), while BR 16 showed fc = 13.54. Glyma0041s00200 (*AP2*) showed positively differential regulated expression in BR 18 to all treatments. However, Embrapa 48 presented down-regulation for root in the two libraries (moderate and severe stress). For Glyma13g17250 (*ERF 018*), the differential expression was more intense in leaf in moderate stress, reaching fc = 32 and fc = 75.25 for Embrapa 48 and BR 16 respectively, while to severe stress, the cultivars presented fc = 8.8 and fc = 63.19, respectively. Considering Glyma17g14110 (*DREB 1E*) in leaf for both stress conditions, Embrapa 48 was more evident (fc = 482.16 in moderate and fc = 387 in severe stress) than BR 16 (fc = 82.74 in moderate and fc = 86 in severe stress). Finally, Glyma20g29410 (*DREB 1A*) was more intensively expressed in Embrapa 48, mainly in leaf severe stress, reaching fc = 26.18) and root in moderate stress (fc = 15.05) (Figure 4A).

2.3 Top genes in AP2/EREBP and WRKY families and analyze of promoters

To further understand the response to water deficit stress in soybean, the top 10 genes were selected within the AP2/EREBP and WRKY families (Table 2). These genes presented quite discrepant differential expression between Embrapa 48 and BR 16, achieving high fold-change in profile of up or down regulated, for tissues (leaf or root) and/or different libraries. The top 10 genes might be the key genes acting in the modulation against stress. Those genes, which are differentially expressed during stress, are good sources of stress-responsive promoters and cis-elements (Hernandez-Garcia and Finer, 2014). Among the evaluated genes, only it was not possible to obtain data from the Glyma16g05880 promoter region. The identification of what each promoter region exclusively contained was prioritized in an attempt to understand the levels of expression presented in this study. The gene Glyma06g04490, belonging to the AP2/EREBP family exhibited the highest number of cis-exclusive elements, reaching 17 classifications. In the WRKY family, the gene Glyma08g01430 exhibited the largest number of exclusive cis-elements, reaching 18

classifications (Table 1 in S3 and Table 2 in S3). Interestingly, for these two mentioned genes, the level of leaf expression under severe stress of Embrapa 48 was the highest among the top 10 (Table 2).

3 DISCUSSION

Signal transduction through membrane receptors transmit the information of the environmental condition that affects the plant, synchronizing an extensive and efficient gene cascade, altering the metabolic profile. Since plant development is largely affected by abiotic environmental stresses, such as water deficit stress, efforts have been dedicated to determining the characteristics and mechanisms involved in the water deficit stress tolerance response in plants (Neves-Borges et al., 2012). AP2/EREBP and WRKY are among the numerous gene families activated during the stress stages in the different tissues. These genes are described as important TF in drought tolerance and other stresses in several species. The increase of tolerance to such stresses (salinity, ionic stress, drought, and low temperature stresses) were obtained in a large number of transgenic plants of different species utilizing TF (DREB): *Oryza sativa* (Paul et al., 2015), *A. thaliana* (Chen et al., 2015) *Saccharum spp. Hybrid.Co 86032* (Augustine et al., 2015), *Glycine max* (Rolla et al., 2014), and TF (WRKY): *Nicotiana tabacum* (Sun et al., 2015), *Oryza sativa* (Cai et al., 2014).

Among these two large families, the AP2/EREBP family stands out, which, according to the expression standards presented here, have a greater number of genes activated in the tolerant cultivar (Embrapa 48) throughout the stress period, when compared to the sensitive cultivar (BR 16). This may be an indication of greater direct participation of the AP2/EREBP family in drought tolerance responses than the *WRKY* family. The expression of the *GmDREB 1* gene, member of AP2/EREBP family, under various abiotic stress conditions in soybean had been related. *DREB 1E*, for example, responded in heat (42 °C), cold (4 °C), NaCl (250 mM), and drought (four and seven days) of stress conditions (Kidokoro et al., 2015) (Figure 5). In the work, it was observed that *DREB 1E* (Glyma17g14110) was strongly active in leaf of Embrapa 48 under moderate and severe stress, participating in the water deficit stress tolerance process.

ERF 018 (Glyma13g17250), also a member of the AP2/EREBP family, have been shown to be involved in the response to both biotic and abiotic stresses (Nakano et al., 2006), reaching greater fold-change in leaf of BR 16 in moderate and severe

stress in relation the Embrapa 48 (Figure 5). However, Gaion et al. (2017) showed that lower sensitivity to ethylene enhances plant growth in both irrigated and stressed conditions. Some ERF that modulate ethylene-dependent transcription (positive or negative regulation) regulate the expression of ethylene inducible genes downstream of *EIN3* (Yang et al., 2005). Therefore, Embrapa 48 probably maintained low level of ethylene because it did not need to activate many ERF genes, such as ERF 018, maintaining a normal plant growth rate. When overexpressed, the *TcWRKY 53* (*Tlaspi caerulescens*) gene, two ethylene response factor (*ERF*) family genes, *NtERF5* and *NtEREBP-1*, presented low transcription levels in *TcWRKY 53* transgenic tobacco plants (Wei et al., 2008). Additionally, the expression of the LEA family, *NtLEA5* gene did not show alteration, suggesting that the increased osmotic stress tolerance occur due to an interaction with *ERF* (TF) (Figure 5).

For example, according to Morimoto et al. (2013) and Singh and Laxmi (2015), the *DREB 2A*, which is classified within the AP2/EREBP family, is regulated by two mechanisms under normal conditions: (I) *GRF7* (growth-Regulation Factor7) inhibits expression of *DREB 2A* binding to its moderate promoter position and (II) target degradation by 26S proteasome mediated proteolysis assisted by *DRIP1* (DREB2A-interacting protein1) and *DRIP2* protein. The TF *ERF 53* (Ethylene response factor 53) and two homologous, *RGLG2* (RING domain ligase 2) and *RGLG1* have been functioned similarly to *DREB 2A/DRIPs* regarding regulation processes. Thus, the activation of *DREB 2A* conditional to water deficit stress provides the organism with an efficient energy-saving tool (Hsieh et al., 2013) (Figure 5). As observed in this study, most of the differentially expressed genes to AP2/EREBP family occurred in leaf, and was not expressed in root, for both cultivars and the treatments. This indicates that genes reserved for root growth/depth induction are prioritized in this tissue and are prioritized in relation to those associated with AP2/EREBP family, probably expanding and diversifying the response, being one of the tolerance strategies carried out in plants (Ogawa et al., 2005). This response can also be attributed to the fact that AP2/EREBP genes are highly specific under organ/tissue regulation by drought (Jin et al., 2009; Wang et al., 2011).

On the other hand, the genes members of the WRKY family did not follow the same behavior of the AP2/EREBP family genes observed in roots (Figure 3A and B). WRKY TF plays important roles in physiological and cellular processes, including senescence (Besseau et al., 2012), root development (Grunewald et al., 2012), plant

growth (Yu et al., 2012), and others. For this reason, it is possible that the WRKY family is more active than AP2/EREBP in roots of all libraries and cultivars. Furthermore, it possibly does not act in tissue-specific processes (Figure 3B). The wide involvement of WRKY proteins in plant physiological processes makes the elucidation of their co-expression traits extremely complex (Rinerson et al. 2015). Dias et al. (2016) observed the expression pattern determined by RT-qPCR showing that, the *GmWRKY 6, GmWRKY 46, GmWRKY 56, GmWRKY 106* and *GmWRKY 149* genes are differentially expressed between a drought tolerant and a susceptible soybean genotype in stress conditions (Figure 5).

The interaction between many TF plays an important role in the mechanism of response, generating specificity in stress responses. Drought responsive gene promoters contain a DRE/CRT motif, in which ABA-independent DREB/CBF TF binds and act as a coupling element for ABRE in ABA-dependent gene expression (Singh and Laxini, 2015). It has been already shown that the DREB 1A/CBF3, DREB 2A, and DREB 1C proteins interact with AREB/ABF (Lee et al., 2010). Thus, there is a crosstalk between ABA-dependent and independent signaling and regulatory pathways (Figure 5). Under osmotic stress conditions, AREB/ABF TF and SnRK2s regulate the transcriptional activation of *DREB 2A* gene, suggesting a complex interaction between *DREB* and *AREB* regulatory regions at the transcript and protein level (Kim et al., 2011) (Figure 5).

The interpretation that the AP2/EREBP family might act in advanced/strongly front the stress than WRKY family and the possible distinct function in drought tolerance responses is supported by the fact that the activation of the WRKY family depends on ABF1, which belongs to the TF bZIP category. Additionally, Phukan et al. (2016) believe that in the future it will be necessary even more studies about TF dependent interaction between other TFs, such as MYC MYB, ERF and NAC, acting with WRKY to regulate various responses in order to elucidate the relations among them and the stress responses. ABA signal perception leads to induction of *WRKY 18* and *WRKY 40* and their product could bind to W-Box present in *WRKY 60* and thereby induce it (Figure 5). However, WRKY act in cross-talk with various gene class, such as AP2/EREBP, being often inactivated to trigger or even amplify the activity of other genes that are more important at that particular moment of stress. WRKY TF have also been reported as positive regulator of ABA-responsive genes like Dehydration Response Element Binding Protein 1a (*DREBP 1a), DREBP 2a MYB 2, ABF 4, ABI 4*

and Response to *ABA 18 (RAB18)* (Rushton et al., 2012). Additionally, for studies conducted in rice, WRKY genes have identified WRKY 25 and WRKY 33 as positive and negative regulators of ABA signaling (Xie et al., 2005).

Genetic studies confirmed that *WRKY 33* acts upstream of *NCED3/NCED5* to negative regulation in ABA biosynthesis (Li et al., 2015). Das and Roychoudhury (2014) showed the increased expression of *WRKY 71* and low induction was reported for *WRKY 24,* in ABA-treated, for the salt tolerant rice variety Pokkali. In *Glycine soja WRKY 20* was isolated (Figure 5), which regulates ABA signaling and enhanced drought tolerance. *GsWRKY 20* has also been reported with high expression in the shoot tips, flowering and the inflorescence meristems of wild soybean (Luo et al., 2013). *Glycine max* in transgenic lines overexpressing *GmWRKY 54* enhanced the drought and salt tolerance. High levels of *GmWRKY 54* are capable to modulate the expression of the TF gene, salt tolerance *Zinc finger* (*STZ/Zat10*) and *DREB 2A* (Figure 5). The overexpression *GmWRKY 13* transgenic plants decreased sensitivity to ABA, while these plants showed less tolerance to high salt and mannitol in relation to the wild types. Thus, it can be understood that *GmWRKY 13* is a negative regulator of abiotic stress response when overexpressed alone (Zhou et al., 2008). *GhWRKY 17* overexpressed in transgenic tobacco plants (*Nicotiana benthamiana*) increased drought and salt sensitivity, but lowered ABA sensitivity leading to low transcription of ABA-inducible genes like *AREB* (ABA-responsive element binding), *NCED* (9-cis-epoxycarotenoid dioxygenases), *DREB* (DEHYDRATION-RESPONSIVE ELEMENT-BINDING PROTEIN), *ERD* (Early Responsive to Dehydration), and *LEA*. It points the important influence of WRKY proteins in activating a proper response against dehydration and high salinity (Yan et al., 2014).

The signalization of cascades in salinity, drought, osmotic, and oxidative stresses is regulated by the *AtWRKY 8, AtWRKY 57* and *AtWRKY 28* genes (Babitha et al., 2013). In high ABA concentration, *AtWRKY 57* has enhanced its expression, reflecting on water deficit stress tolerance to the plant. In this same study, it was confirmed the binding of *WRKY 57* to the W-box of responsive to desiccation *29A (Rd29A)* and 9-cis-epoxycarotenoid dioxygenase 3 (*NCED 3*) promoters. Therefore, *WRKY 57* acts directly inducing genes that are responsive to stress (Jiang et al., 2012) (Figure 5).

Stress-induced promoters are useful tools for the regulation of genes linked to stress defense or those imminently lethal ones that are activated in response to stress related (Hernandez-Garcia and Finer, 2014). According to Hernandez-Garcia and Finer (2014), differentially expressed genes in stress condition are good sources of stress-responsive promoters and cis-regulatory elements (CREs). Promoters and *OsNCED 3* and *Wsi18* genes, induced in the synthesis and signaling of ABA, were strongly drought-inducible in high-salinity treatments in transgenic line of rice (Yi et al., 2011; Bang et al., 2013).

Among the top 10 selected genes categorized as belonging to the AP2/EREBP family, Glyma06g04490 is highlighted for presenting bZIP, MYB and WRKY CREs associated in promoter region, which is more important than others genes (Glyma20g34550, Glyma20g29410, Glyma19g31960, Glyma17g17010, Glyma15g19910, Glyma10g02080, Glyma02g42960, Glyma01g42500, Glyma03g26520). Among the top 10 genes and categorized as belonging to the WRKY family, Glyma13g44730, Glyma08g01430, Glyma03g05220 and Glyma06g06500 have the promoter region composed of at least two distinct cis-elements linked to (MYC and MYB), (MYB and ERF), (bZIP and WRKY) and (DREB and bZIP) respectively, deserving prominence among the top 10 genes. Gene families such as MYB, MYC, ERF, bZIP, DREB and WRKY transcription factors could also be good candidates for identifying stress-inducible promoters. Inducible promoters that mediate rapid responses under stress conditions can contribute as potential key-elements in the control genes that provide protection during the initial stages of stress (Hernandez-Garcia and Finer, 2014; Khan et al., 2017).

4. METHODS

4.1 Plant Materials and Experimental Design in hydroponic solution

4.1.1 Plant growth and drought stress treatment for Transcriptome analysis

Based on the method and experimental design proposed by Martins et al. (2008), it was performed an evaluated experiment with soybean transcriptomic data under water deficit conditions, using a hydroponic system. Seeds of the soybean cultivars BR 16 and Embrapa 48, classified for Oya et al. (2004) as sensitive and tolerant to water deficit, respectively, were germinated on germination paper during four days in a growth chamber at 25 ± 1 ºC and 100% relative humidity. Seedlings were placed in 36 L boxes containing 50% Hoagland's solution (without modifications) (Hoagland

and Arnon, 1950), which was continuously aerated and replaced on a weekly basis. These boxes were then transferred to a greenhouse under a natural photoperiod of approximately 12/12 h light/dark cycle, temperature of 30 ± 5 ºC and 60 ± 10 % relative humidity (RH), where the plants were allowed to grow until the V4 stage (Fehr et al., 1971). The experiment followed a randomized complete block design 3x5 factorial, with three replicates. The factors were two cultivars (BR 16 and Embrapa 48) and five times of dehydration (0, 25, 50, 125 and 150 min). Stress conditions were simulated by removing the plants out of the hydroponic solution and leaving them in boxes without nutrient solution for up to 150 min, under ambient-air exposure. For each time of water deficit, roots from 15 plants were collected, pooled and frozen in liquid nitrogen before storage at -80 ºC. The same procedure was performed for leaf samples.

4.1.2 Library construction and sequencing run

The total RNA was extracted from leaves and root samples from BR 16 and Embrapa 48 plants using the Trizol reagent (Invitrogen). Bulks of total RNA was made utilizing all samples of same tissue (25 and 50 min - moderate stress) and (125 and 150 min - severe stress) for leaf and root, generating four libraries from BR 16 cultivar and four libraries from Embrapa 48. Additionally, DNase treatment (Life Technologies Grand Island, NY, USA) and high-quality total RNA were used to analyze the transcripts. The RNA-seq libraries were built using the Nugen-OvationH kit according to the manufacturer's instructions (NuGEN Technologies Inc., San Carlos, CA, USA). The obtained libraries were subjected to sequencing by Illumina HiSeq2000 (Illumina, San Diego, CA, USA).

4.1.3 Mapping of reads and functional classification

Mapping of reads was performed with the soybean genome (Phytosome Glycine max v1.1) using the GeneSifter platform (http://www.geospiza.com/Products/ AnalysisEdition.shtml). To compare gene expression between different times and conditions, $\log^{2-}$ was used to transform the normalized reads per million mapped values (RPM). Then, an analysis was run and tested (t-test for two group comparisons). Libraries tagged with a barcode produced thousands of reads from each library. Contig sequences were submitted to the non-redundant protein database NCBI through Blast X (Altschul et al., 1997), to search for similarity to

406 known proteins. In addition, sequences were analyzed by the moderate ware
407 AutoFact (Koski et al., 2005), which is an automated annotation tool that assigns
408 biological information for a given sequence by comparing different databases. We
409 used the UniRef90 e UniRef100, KEGG (Kanehisa and Goto, 2000), Pfam (Finn et
410 al., 2010), Smart (Schultz et al., 1998) databases. In order to establish the GO (Gene
411 Ontology) terms (Ashburner et al., 2000), we employed the Blast2Go program to
412 classify the sequences according to the molecular function and biological process
413 described to the respective protein (Götz et al., 2008; Carbon et al., 2009).
414
415 **4.1.4 Analysis of differential gene expression**
416 Using the GeneSifter platform (http://www.geospiza.com/Products/
417 AnalysisEdition.shtml), we applied a pairwise comparison between the control and
418 water deficit treatment, using the two libraries synthesized from plants of the same
419 time period/treatment. In the pairwise analysis, we only used genes with more than
420 20 mapped reads to compare gene expression using the edgeR statistical test
421 (Robinson et al., 2010). The ratio of expression (fold-change) was acquired by
422 dividing values of gene expression under water deficit and control conditions. The
423 statistical test was combined with the multiple-hypothesis-testing correction method
424 of Benjamini and Hochberg (1995), which calculates the False Discovery Rate (FDR),
425 to qualify statistically significant, differentially expressed genes by avoiding inflation
426 of type-1 errors. Differential gene expression was considered significant at an
427 adjusted *p*-value ≤ 0.01, and down- and up-regulation was established in the range
428 of ≤ -2 to ≥ 2 fold-change (fc), respectively. We also applied a stringent statistical
429 significance cutoff (adjusted p-value ≤ 0.001) to improve confidence.
430
431 **4.1.5 Classification and characterization functional**
432 Analyses through MapMan 3.6 ORC1 (http://mapman.gabipd.org/web/guest/
433 mapman-version-3.6.0) allowed us to visualize differentially expressed genes, which
434 was calculated based on calibration with time 0 of stress (control) and were classified
435 by functional categories in several pathways. Three categories were mapped and
436 selected to analyses: transcript factor, the AP2/EREBP and WRKY families.
437

4.1.6 Validation of gene expression

4.1.6.1 Gene Selection

The gene selection for validation was performed through the search of data deposited in the GeneSifter - Analysis Edition (GSAE) System Requirements, an online analysis moderateware that assists in the rapid visualization of the amount of data generated by Next Gen Sequencing and Microarray Analysis. Differential expression gene was considered significant at an adjusted *p*-value ≤ 0.01, and down- and up-regulation was established in the range of ≤ -2 to ≥ 2 fold-change (fc), respectively. Seven genes were also explored by biological definitions and differential expression profiles: Glyma17g17860 (*LEA 18*), Glyma08g01430 (*WRKY 75*), Glyma05g32040 (*AP2*), Glyma0041s00200 (*AP2*), Glyma13g17250 (*ERF 018*), Glyma17g14110 (*DREB 1E*) and Glyma20g29410 (*DREB 1A*) (Table 1 in S2).

4.1.6.2 RT-qPCR analysis

Relative expression levels of the target genes were measured in root and leaf samples from Embrapa 48 and BR 16 plants. Bulk time point (0, 25-50, 125-150 min under water deficit) was analyzed, generating libraries (control, moderate and severe) with three biological replicates, each with three technical replicates. After DNAse treatment (Invitrogen/Life Technologies, Grand Island, NY, USA), high quality total RNA was used to synthesize cDNA strands (Superscript II First Strand Synthesis, Invitrogen/Life Technologies, Grand Island, NY, USA), and cDNA quality was verified using a standard PCR reaction with an actin primer that spanned an intronic region. Additionally, carrying out the amplification efficiency analysis, the genes were amplified by qPCR using a StepOne RT-qPCR Thermocycler (Applied Biosystems/Life Technologies, Grand Island, NY, USA) with the following cycling FAST parameters: 95 ºC for 20 seconds, 40 cycles at 95 ºC for 3 seconds, 60 ºC for 30 seconds and Melt Curve 95 ºC for 15 seconds and 60 ºC for 1 minute.

Data were collected during the extension phase, and dissociation curves were performed by heating each amplicon from 60 to 95 ºC and taking readings at one-degree intervals to verify the specificity of the primers. The Rest2009 moderateware package (Pfaffl et al., 2002) was used to evaluate the data because this program provided a more robust statistical analysis. The normalization of the real-time quantitative RT-qPCR was performed by taking the geometric average of the selected endogenous genes (FYVE and B-act, according with analyses of Marcolino-

Gomes et al. (2015), and the control plants (0 min under stress) were used to normalize the relative expression. Hypothesis testing was used to determine whether the differences between the control and treatment conditions were significant (Pfaffl et al., 2002).

4.1.6.3 Primer design and efficiency analysis

Primers for the target genes were designed based on the GeneModels of the selected genes using the program Primer Express 3.0 (Applied Biosystems/Life Technologies, Grand Island, NY, USA) (Table 2 in S2). Primer sequences were determined for the 3'end of each gene, and the amplicons spanned up to 150 base pairs (bp). Primer sequences were BLASTed against the soybean genome (Phytozome database v1.0, http://www.phytozome.net/search.php) to verify the specificity of each primer, and standard curves were produced from serial dilutions of a cDNA pool to estimate the efficiency of the PCR amplification reactions.

4.1.7 Promoter analysis

The drought gene regulatory sequences from the top 10 genes to AP2/EREBP and WRKY families selected and analyzed using promoter sequence 2 kb upstream of the translational start site. To map the domains regions, the Genomatix program was employed utilizing the tool Gene2Promoter, which provides access to promoter sequences of all genes annotated in the available genomes. Promoter regions were thoroughly annotated and validated according to highest scientific standards, including Genomatix proprietary technology (e.g. PromoterInspector, oligo-capping, comparative genomics) (http://www.genomatix.de/cgi-bin/c2p/c2p.pl?s=8a5d8b1c9bd7f6d975b5578c2d1de62a). In a more restrictive analysis, we selected only the cis-elements found at the upstream 200bp to the target gene. Additionally, the data were filtered by strictly identifying cis-elements unique to each evaluated gene. This filter was applied to the top 10 gene classes (AP2 / EREBP and WRKY)

4.2 Concluding remarks

The AP2/EREBP and WRKY families are responsive to water deficit stress, but they tend to be different, often opposite, in the tolerance response in some cases up and down-regulated or inducers and repressors of gene activation respectively. They can

operate in cross-talk between different metabolic pathways or in feedback control mechanism, as well as the WRKY 54, extending the response of FT as the DREB 2A. In addition, the top 10 genes might be potential study tools in the analysis of their promoters and regulatory mechanisms, since they were highlighted in a contrasting manner between the two evaluated cultivars, Embrapa 48 (tolerant) and BR 16 (sensitive).

REFERENCES

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**:3389-33402.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. **25**:25-29.

Augustine, S.M., Narayan, J.A., Syamaladevi, D.P., Appunu, C., Chakravarthi, M., Ravichandran, V., Tuteja, N., and Subramonian, N. (2015). Overexpression of EaDREB2 and pyramiding of EaDREB2 with the pea DNA helicase gene (PDH45) enhance drought and salinity tolerance in sugarcane (Saccharum spp. hybrid). Plant Cell Rep. **34**:247–263.

Babitha, K.C., Ramu, S.V., Pruthvi, V., Mahesh, P., Nataraja, K.N., and Udayakumar, M. (2013). Co-expression of AtbHLH17 and AtWRKY28 confers resistance to abiotic stress in Arabidopsis. Transgenic Res. **22**:327–341.

Bang, S.W., Park, S.H., Jeong, J.S., Kim, Y.S., Jung, H., Há, S.H., and Kim, J.K. (2013). Characterization of the stress-inducible OsNCED3 promoter in different transgenic rice organs and over three homozygous generations. Planta **237**:211–224.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Statist. Soc. B. **57**:289–300.

Besseau, S., Li, J., and Palva, E.T. (2012). WRKY54 and WRKY70 co-operate as negative regulators of leaf senescence in Arabidopsis thaliana. J. Exp. Bot. **63**:2667–2679.

Bucholska, J., Minkiewicz, P., Darewicz, M., and Iwaniak, A. (2018). Databases and associated bioinformatic tools in studies of food allergens, epitopes and haptens - a review. Pol. J. Food Nutr. Sci. **68**:103-113.

Cai, R., Zhao, Y., Wang, Y., Lin, Y., Peng, X., Li, Q., Chang, W., Jiang, H., Xiang, Y., and Cheng, B. (2014). Overexpression of a maize WRKY58 gene enhances drought and salt tolerance in transgenic rice. Plant Cell Tiss. Org. Cult. **119**:565–577.

Carbon, S., Ireland, A., Mungall, C.J., Shu, S., Marshall, B., and Lewis, S. (2009). AmiGO: online access to ontology and annotation data. Bioinformatics **25**:288-289.

Chen, H., Liu, L., Wang, L., Wang, S., and Cheng, X. (2015). VrDREB2A, a DREB-binding transcription factor from Vigna radiata, increased drought and high-salt tolerance in transgenic Arabidopsis thaliana. J. Plant Res. **129**:263-273.

Chu, X., Wang, C., Chen, X., Lu, W., Li, H., Wang, X., Hao, L., and Guo, X. (2015). The cotton WRKY gene GhWRKY41 positively regulates salt and drought stress tolerance in transgenic Nicotiana benthamiana. PLoS One **11**:e0157026.

Das, K., and Roychoudhury, A. (2014). Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. Front. Environ. Sci. **2**:1-13.

de Carvalho M.H.C. (2008). Drought stress and reactive oxygen species. Plant Signal. Behav. **3**:156-165.

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

Ding, Z.J., Yan, J.Y., Li, C.X., Li, G.X., Wu, Y.R., and Zheng, S.J. (2015). Transcription factor WRKY46 modulates the development of Arabidopsis lateral roots in osmotic/salt stress conditions via regulation of ABA signaling and auxin homeostasis. Plant J. 84:56–69.

Fehr, W.R., Caviness, C.E., Burmood, D.T., and Pennington, J.S. (1971). Stage of development descriptions for soybeans, Glycine Max (L.) Merrill. Crop Sci. **11**:929-931.

Finn, R.D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J.E., Gavin, O.L., Gunesekaran, P., Ceric, G., Forslund, K., et al. (2010). The Pfam protein families database. Nucleic Acids Res. **38**:D211-D222.

Gaion, L.A., Baldo, A., Ferreira Júnior, D. da C., D'Amico-Damião, V., and Carvalho, R.F. (2017). Ethylene modulates tomato growth during drought stress. Holos Environ. **supl**:56-68.

Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talón, M., Dopazo, J., and Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. **36:**3420-3435.

Grunewald, W., de Smet, I., Lewis, D.R., Löfke, C., Jansen, L., Goeminne, G., Bossche, R.V., Karime, M., de Rybel, B., Vanholme, B., et al. (2012). Transcription factor WRKY23 assists auxin distribution patterns during Arabidopsis root development through local control on flavonol biosynthesis. Proc. Natl. Acad. Sci. USA **109**:1554–1559.

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

Hoagland, D.R., Arnon, D.I. (1950). The water culture method for growing plants without soils. Berkeley: California Agricultural Experimental Station.

Hsieh, E.J., Cheng, M.C., and Lin, T.P. (2013). Functional characterization of an abiotic stress-inducible transcription factor AtERF53 in Arabidopsis thaliana. Plant Mol. Biol. **82**:223–237.

Jiang, W., Wu, J., Zhang, Y., Yin, L., and Lu, J. (2015). Isolation of a WRKY30 gene from Muscadiniarotundifolia (Michx)and validation of its function under biotic and abiotic stresses. Protoplasma **252**:1361–1374.

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

Jin, P., Huang, L.Y., Wang, D., Wu, H.M., Zhu, L.H., and Fu, B.Y. (2009). Expression profiling of rice AP2/EREBP genes responsive to abiotic stresses. Sci. Agric. Sin. **42**:3765-3773.

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. **28**:27-30.

Khan, M.R., Khana, I., Ibrarc, Z., Léona, J., and Naza, A.A. (2017). Drought-responsive genes expressed predominantly in root tissues are enriched with homotypic cis-regulatory clusters in promoters of major cereal crops. Crop J. **5**:195–206.

606 **Kidokoro, S., Maruyama, K., Nakashima, K., Imura, Y., Narusaka, Y., Shinwari,**
607 **Z.K., Osakabe, Y., Fujita, Y., Mizoi, J., Shinozaki, K., et al.** (2009). The
608 phytochrome-interacting factor PIF7 negatively regulates DREB1 expression
609 under circadian control in Arabidopsis. Plant Physiol. **151**:2046–2057.
610 **Kidokoro, S., Watanabe, K., Ohori, T., Moriwaki, T., Maruyama, K., Mizoi, J.,**
611 **Htwe, M.P.S., Fujita, Y., Sekita, S., Shinozaki, K., et al.** (2015). Soybean
612 DREB1/CBF-type transcription factors function in heat and drought as well as
613 cold stress-responsive gene expression. Plant J. **81**:505-518.
614 **Kidokoro, S., Watanabe, K., Ohori, T., Moriwaki, T., Maruyama, K., Mizoi, J.,**
615 **Htwe, M.P.S., Fujita, Y., Sekita, S., Shinozaki, K., et al.** (2015). Soybean
616 BRB1/CBF-type transcription factors function in heat and drought as well as cold
617 stress-responsive gene expression. Plant J. **81**:505-518.
618 **Kim, J.S., Mizoi, J., Yoshida, T., Fujita, Y., Nakajima, J., Ohori, T., Todaka, D.,**
619 **Nakashima, K., Hirayama, T., Shinozaki, K., et al.** (2011). An ABRE promoter
620 sequence is involved in osmotic stress-responsive expression of the DREB2A
621 gene, which encodes transcription factor regulating drought inducible genes in
622 Arabidopsis. Plant Cell Physiol. **52**:2136–2146.
623 **Koski, L.B., Gray, M.W., Lang, B.F., and Burger, G.** (2005). AutoFACT: an
624 automatic functional annotation and classification tool. BMC Bioinformatics
625 **6**:151.
626 **Lee, S.J., Kang, J.Y., Park, H.J., Kim, M.D., Bae, M.S., Choi, H.I., and Kim, S.Y.**
627 (2010). DREB2C interacts with ABF2, a bZIP protein regulating abscisic acid-
628 responsive gene expression, and its overexpression affects abscisic acid
629 sensitivity. Plant Physiol. **153**:716–727.
630 **Li, J.B., Luan, Y.S., and Liu, Z.** (2015). Overexpression of SpWRKY1 promotes
631 resistance to Phytophthora nicotianae and tolerance to salt and drought stress in
632 transgenic tobacco. Physiol. Plant. **155**:248–266.
633 **Luo, X., Bai, X., Sun, X., Zhu, D., Liu, B., Ji, W., Cai, H., Cao, L., Wu, J., Hu, M., et**
634 **al.** (2013). Expression of wild soybean WRKY20 in Arabidopsis enhances
635 drought tolerance and regulates ABA signalling. J. Exp. Bot. **64**:2155–2169.
636 **Marcolino-Gomes, J., Rodrigues, F.A., Fuganti-Pagliarini, R., Nakayama, T.J.,**
637 **Reis, R.R., Farias, J.R.B., Harmon, F.G., Molinari, H.B.C., Molinari, M.D.C.,**
638 **and Nepomuceno, A.** (2015). Transcriptome-wide identification of reference

genes for expression analysis of soybean responses to drought stress asevere the day. PLoS One **10**:e0139051.

Martins, P.K., Jordão, B.Q., Yamanaka, N., Farias, J.R.B., Beneventi, M.A., Binneck, E., Fuganti, R., Stolf, R., and Nepomuceno, A.L. (2008). Differential gene expression and mitotic cell analysis of the drought tolerant soybean (Glycine max L. Merrill Fabales, Fabaceae) cultivar MG/BR46 (Conquista) under two water deficit induction systems. Genet. Mol. Biol. **31**:512–521.

Molina, C., Rotter, B., Horres, R., Udupa, S.M., Besser, B., Bellarmino, L., Baum, M., Matsumura, H., Terauchi, R., Kahl, G., et al. (2008). SuperSAGE: the drought stress-responsive transcriptome of chickpea roots. BMC Genomics **9**:e553.

Morimoto, K., Mizoi, J., Qin, F., Kim, J.S., Sato, H., Osakabe, Y., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2013). Stabilization of Arabidopsis DREB2A is required but not sufficient for the induction of target genes under conditions of stress. PLoS One **8**:e80457.

Nakano, T., Suzuki, K., Fujimura, T., and Shinshi, H. (2006). Genome-wide analysis of the ERF gene family in Arabidopsis and rice. Plant Physiol. **140**:411–432.

Nakashima, K., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2014). The transcriptional regulatory network in the drought response and its crosstalk in abiotic stress responses including drought, cold, and heat. Front. Plant Sci. **5**:170.

Neves-Borges, A.C., Guimarães-Dias, F., Cruz, F., Mesquita, R.O., Nepomuceno, A.L., Romano, E., Loureiro, M.E., de Fátima, G.-de-S.M., and Alves-Ferreira, M. (2012). Expression pattern of drought stress marker genes in soybean roots under two water deficit systems. Genet. Mol. Biol. **35**:212–221.

Ogawa, A., Kawashima, C., and Yamauchi, A. (2005). Sugar accumulation asevere the seminal root axis, as affected by osmotic stress in maize: a possible physiological basis for plastic lateral root development. Plant Prod. Sci. **8**:173–180.

Oya, T., Nepomuceno, A.L., Neumaier, N., Farias, J.R.B., Tobita, S., and 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.

673 **Paul, S., Gayen, D., Datta, S. K., and Datta, K.** (2015). Dissecting root proteome of
674 transgenic rice cultivars unravels metabolic alterations and accumulation of
675 novel stress responsive proteins under drought stress. Plant Sci. **234**:133–143.
676 **Pfaffl, M.W., Horgan, G.W., and Dempfle, L.** (2002). Relative expression
677 moderateware tool (REST©) for group-wise comparison and statistical analysis
678 of relative expression results in real-time PCR. Nucleic Acids Res. **30**:e36.
679 **Phukan, U.J., Jeena, G.S., and Shukla, R.K.** (2016). WRKY Transcription factors:
680 molecular regulation and stress responses in plants. Front. Plant Sci. **7**:760.
681 **Rashid, M., Guangyuan, H., Guangxiao, Y., Hussain, J., and Xu, Y.** (2012).
682 AP2/ERF transcription factor in rice: genome-wide canvas and syntenic
683 relationships between monocots and eudicots. Evol. Bioinform. **8**:321-355.
684 **Rayko, E., Maumus, F., Maheswari, U., Jabbari, K., and Bowler, C.** (2010).
685 Transcription factor families in ferred from genome sequences of
686 photosynthetics tramenopiles. New Phytol. **188**:52–66.
687 **Riechmann, J. L., and Meyerowitz, E. M.** (1998). The AP2/EREBP family of plant
688 transcription factors. Biol. Chem. **379**:633–646.
689 **Rinerson, C.I., Scully, E.D., Palmer, N.A., Donze-Reiner, T., Rabara, R.C.,**
690 **Tripathi, P., Shen, Q.J., Sattler, S.E., Rohila, J.S., Sarath, G., et al.** (2015).
691 The WRKY transcription factor family and senescence in switchgrass. BMC
692 Genomics **6**:912.
693 **Robinson, M.D., McCarthy, D.J., and Smyth, G.K.** (2010). edgeR: a bioconductor
694 package for differential expression analysis of digital gene expression data.
695 Bioinformatics **26**:139–140.
696 **Rolla, A.A. de P., Carvalho, J. de F.C, Fuganti-Pagliarini, R., Engels, C., do Rio,**
697 **A., Marin, S.R.R, de Oliveira, M.C., Beneventi, M.A., Marcelino-Guimarães,**
698 **F.C., Farias, J.R., et al.** (2014). Phenotyping soybean plants transformed with
699 rd29A:AtDREB1A for drought tolerance in the greenhouse and field. Transgenic
700 Res. **23**:75–87.
701 **Rushton, D.L., Tripathi, P., Rabara, R.C., Lin, J., Ringler, P., Boken, A.K.,**
702 **Langum, T.J., Smidt, L., Boomsma, D.D., Emme, N.J., et al.** (2012). WRKY
703 transcription factors: key components in abscisic acid signalling. Plant
704 Biotechnol. J. **10**:2–11.

Schultz, J., Milpetz, F., Bork, P., and Ponting, C.P. (1998). SMART, a Simple Modular Architecture Research Tool: Identification of Signaling Domains. Proc Natl Acad Sci U S A. **95**:5857–5864.

Singh, D., and Laxmi, A. (2015). Transcriptional regulation of drought response: a tortuous network of transcriptional factors. Front. Plant Sci. **6**:895.

Sun, J., Hu, W., Zhou, R., Wang, L., Wang, X., Wang, Q., Feng, Z., Li, Y., Qiu, D., He, G., et al. (2015). The Brachypodium distachyon BdWRKY36 gene confers tolerance to drought stress in transgenic tobacco plants. Plant Cell Rep. **34**:23–35.

Wang, D., Pan, Y., Zhao, X., Zhu, L., Fu, B., and Li, Z. (2011). Genome-wide temporal-spatial gene expression profiling of drought responsiveness in rice. BMC Genomics **12**:149.

Wani, S.H., Singh, N.B., Devi, T.R., Haribhushan, A., and Jeberson, S.M. (2013). Engineering abiotic stress tolerance in plants: extricating regulatory gene complex, in Conventional and Non-Conventional Interventions in Crop Improvement, eds. C.P. Malik, G.S. Sanghera, and S.H. Wani (New Delhi: CABI),1–19.

Wei, W., Zhang, Y.X., Han, L., Guan, Z.Q., and Chai, T.Y. (2008). A novel WRKY transcriptional factor from Tlaspi caerulescens negatively regulates the osmotic stress tolerance of transgenic tobacco. Plant Cell Rep. **27**:795–803.

Xie, Z., Zhang, Z.L., Zou, X.L., Huang, J., Ruas, P., Thompson, D., and Shen, Q.J. (2005). Annotations and functional analyses of the rice WRKY gene superfamily reveal positive and negative regulators of abscisic acid signaling in aleurone cells. Plant Physiol. **137**:176-189.

Yamaguchi-Shinozaki, K., and Shinozaki, K. (2006). Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. Annu. Rev. Plant Biol. **57**:781–803.

Yan, H., Jia, H., Chen, X., Hao, L., An, H., and Guo, X. (2014). The cotton WRKY transcription factor GhWRKY17 functions in drought and salt stress in transgenic Nicotiana benthamiana through ABA signaling and the modulation of reactive oxygen species production. Plant and Cell Physiol. **55**:2060–2076.

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

Yi, N., Oh, S.J., Kim, Y.S., Jang, H.J., Park, S.H., Jeong, J.S., Song, S.I., Choi, Y.D., and Kim, J.K. (2011). Analysis of the Wsi18, a stress-inducible promoter that is active in the whole grain of transgenic rice. Transgenic Res. **20**:153–163.

Yu, F., Huaxia, Y., Lu, W., Wu, C., Cao, X., Guo, X. (2012). GhWRKY15, a member of the WRKY transcription factor family identified from cotton (Gossypium hirsutum L.), is involved in disease resistance and plant development. BMC Plant Biol. **12**:144.

Yu, Y., Wang, N., Hu, R. and Xiang, F. (2016). Genome-wide identifcation of soybean WRKY transcription factors in response to salt stress. Springerplus **5**:920.

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**:486-503.

Zhu, J.-K. (2016). Abiotic stress signaling and responses in plants. Cell **167**: 313-324.

Zong, W., Tang, N., Yang, J., Peng, L., Ma, S., Xu, Y., Li, G., and Xiong, L. (2016). Feedback regulation of ABA signaling and biosynthesis by a bZIP transcription factor targets drought-resistance-related genes. Plant Physiol. **171**:2810-2825.

Tables

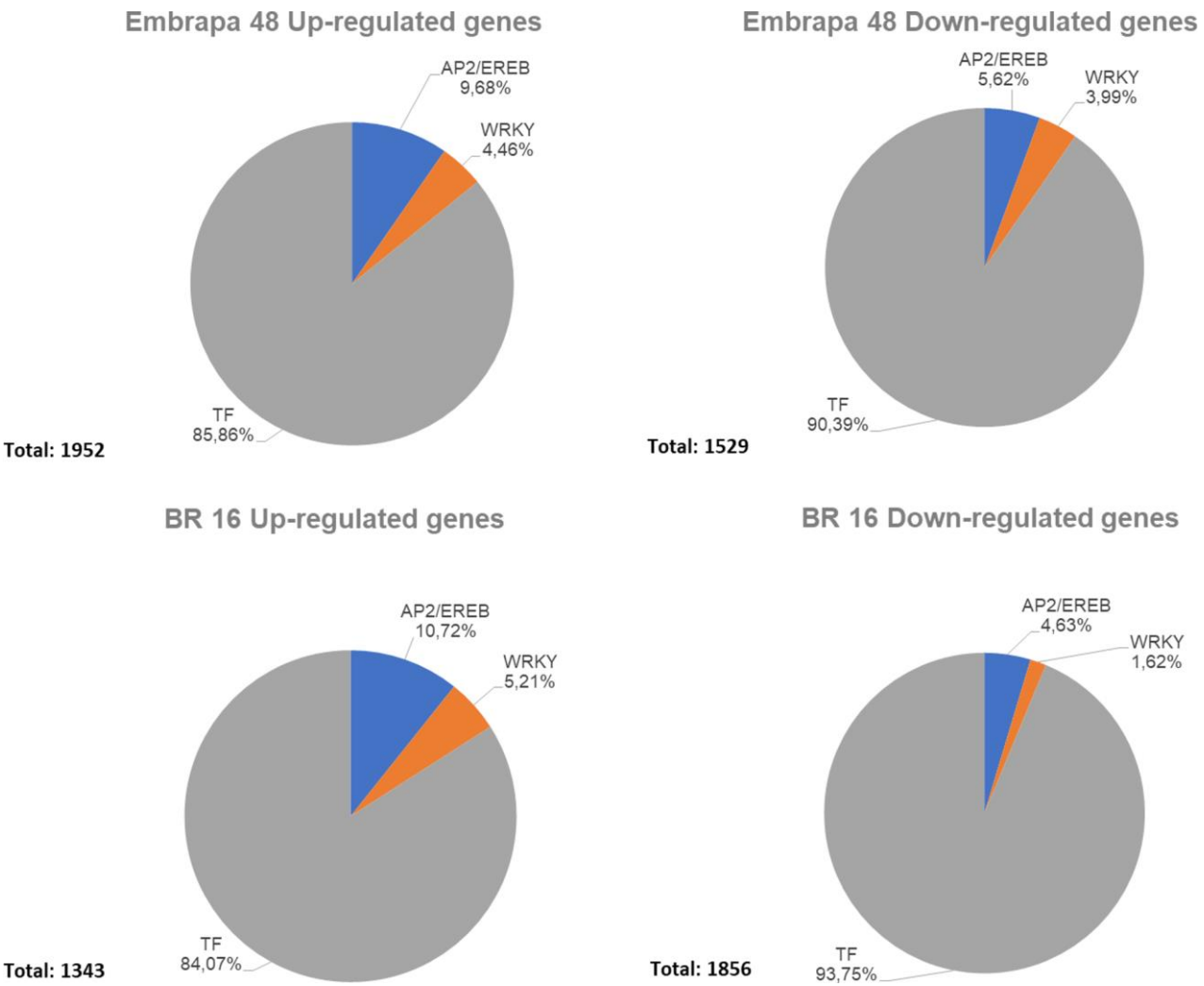
Table 1. Embrapa 48 and BR 16 general RNA-seq data.

Sample	Reads after adapter removal	Mapped reads [0..2 mis]	% of mapped reads [0..2 mis] (Mean)
Embrapa 48 leaf	12,480,047	6,139,313	49.2
Embrapa 48 root	12,241,551	6,137,155	50.2
BR 16 leaf	11,305,266	5,414,745	47.6
BR 16 root	11,150,778	5,555,411	49.7
Total	47,177,642	23,246,624	-

Table 2 - Top 10 genes.

Family	Wm82.a1.v1	Wm82.a2.v1	Genome localization	Embrapa 48 Leaf expression profile		BR 16 Leaf expression profile		Embrapa 48 Root expression profile		BR 16 Root expression profile	
				Short stress	Long stress	Short stress	Long stress	Short stress	Long stress	Short stress	Long stress
AP2/EREBP	Glyma01g42500	Glyma.01g216000.1	Chr01:54666766..54676069 reverse	-	-	14.38	12.23	110.96	110.73	28.29	27.24
	Glyma02g42960	Glyma.02g261700.1	Chr02:44767828..44770286 forward	-	-	-	-	-	10.05	-	-
	Glyma03g26520	Glyma.03g112700	Chr03:31930425..31931607	-	15.64	-	15.11	27.74	37.81	57.19	17.61
	Glyma06g04490	Glyma.06g042100.1	Chr06:3179136..3180497 reverse	4.24	48.02	-	-	-	-5.43	-	-
	Glyma10g02080	Glyma.10g016500.1	Chr10:1475606..1477360 reverse	2.12	11.17	-	2.16	6.6	5.4	-	-
	Glyma15g19910	Glyma.15g180000.1	Chr15:17375377..17376578 reverse	22.26	2.23	8.99	2.88	51.52	36.46	-	-
	Glyma20g29410	Glyma.20g155100	Chr02:39389515..39390591	12.72	8.93	5.39	7.91	6.6	6.3	19.47	28.61
	Glyma17g17010	Glyma.17g158300.1	Chr17:13595433..13600012 reverse	8.48	-	-	-	-25.74	-	-6.57	-14.54
	Glyma19g31960	Glyma.19g138000.1	Chr19:39934458..39938605 reverse	7.42	14.52	-3.34		-	-	-	2.75
	Glyma20g34550	Glyma.20g203500.1	Chr20:44033567..44034010 forward	-	8.1	-	24.46	6.6	16.2	8.52	18.71
WRKY	Glyma03g05220	Glyma.03g042700.1	Chr03:5405762..5409385 forward	-	-	-	-	75.16	66.57	28.95	-
	Glyma06g06500	Glyma.06g061800.1	Chr06:4648505..4651568 reverse	6.36	-	-	-	-2.27	-	3.04	-
	Glyma07g02630	Glyma.07g023300.1	Chr07:1772171..1774607 reverse	-	-	-	13.19	21.08	-	-	-
	Glyma08g01430	Glyma.08g011300.1	Chr08:901327..902577 forward	2.12	317.17	-	29.82	15.85	4.05	13.38	26.41
	Glyma08g02580	Glyma.08g021900.1	Chr08:1762687..1764910 reverse	-	-	-	-	36.19	22.33	-	-
	Glyma08g08720	Glyma.08g082400.1	Chr08:6229164..6231818 forward	7.42	15.64	2.25	-	-	-	-	-
	Glyma13g44730	Glyma.13g370100.1	Chr13:45564565..45566732 forward	-	13.09	-	-	-	-	109.91	108.77
	Glyma16g05880	Glyma.16g054400.1	Chr16:5336881..5339714 forward	-	-	-2.78	-	46.23	43.21	12.57	11.74
	Glyma18g06360	Glyma.18g056600.1	Chr18:4952759..4956347 forward	2.12	60.31	-	8.51	-	-	-	-
	Glyma19g26400	Glyma.19g094100.1	Chr19:33279312..33281763 reverse	-4.72	29.04	-	-	-	-	17.03	14.31

* The negative signal in front the number, represented down-regulation gene.



764

765 Figure 1. Graphic representation of eight gene categories (enriched biological process from MapMan 3.6 ORC) analyzed in up and
766 down-regulated gene groups to cultivars Embrapa 48 and BR 16.

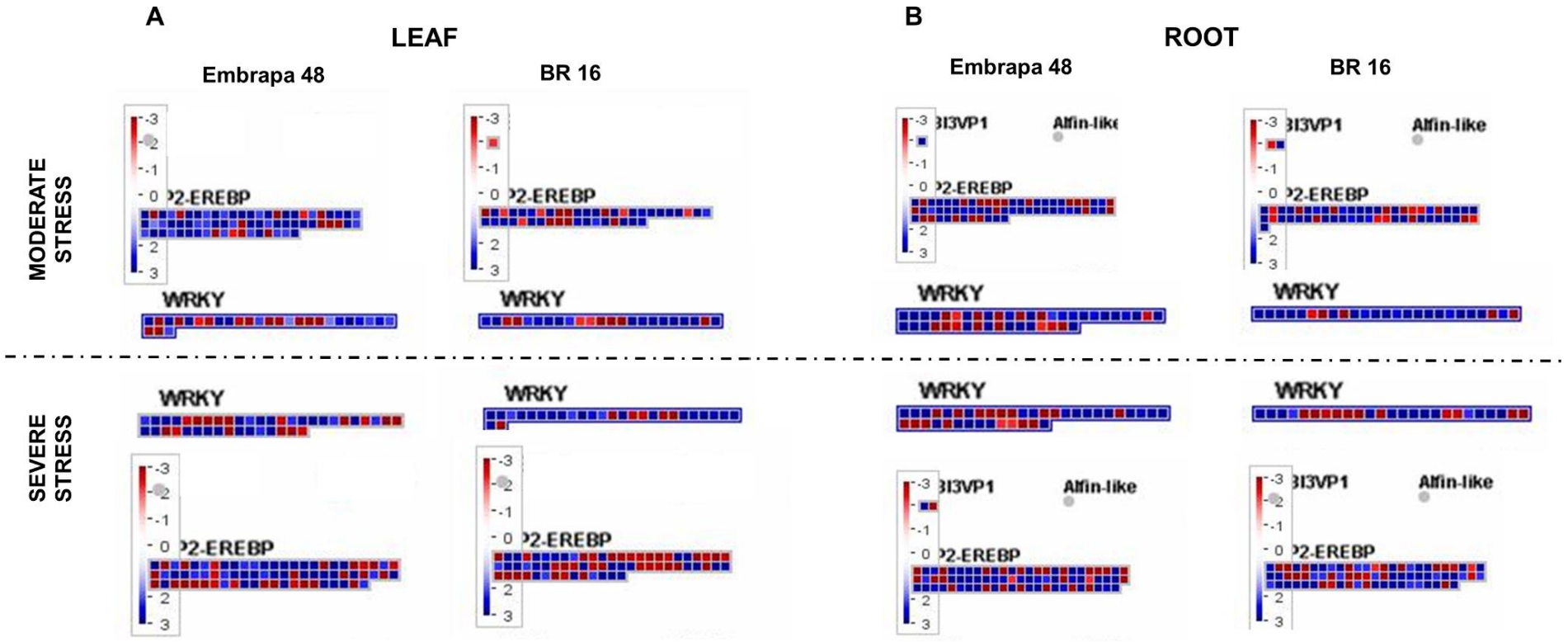


Figure 2. Modified maps of MapMan 3.6 ORC. Genes differentially expressed in response to water deficit were mapped to specific stress-related pathways. The color scale show the $\log_2$ fold change: red = down-regulated and blue = up-regulated.

(A) BR 16 leaf in moderate stress; BR 16 leaf in severe stress; Embrapa 48 leaf in moderate stress; Embrapa 48 leaf in severe stress. Functional roles triggered in soybean plants under water deficit stress.

(B) BR 16 root in moderate stress; BR 16 root in severe stress; Embrapa 48 root in moderate stress; Embrapa 48 root in severe stress.

774

775 Figure 3. Heat Map visualization of up and down-regulated genes characterized
776 within the category AP2 Family (A) and the WRKY Family (B) classified from
777 (MapMan 3.6 ORC) for thecultivars BR 16 and Embrapa 48, under different
778 treatments. Green cells represent down-regulated genes and red cells represent up-
779 regulated genes.

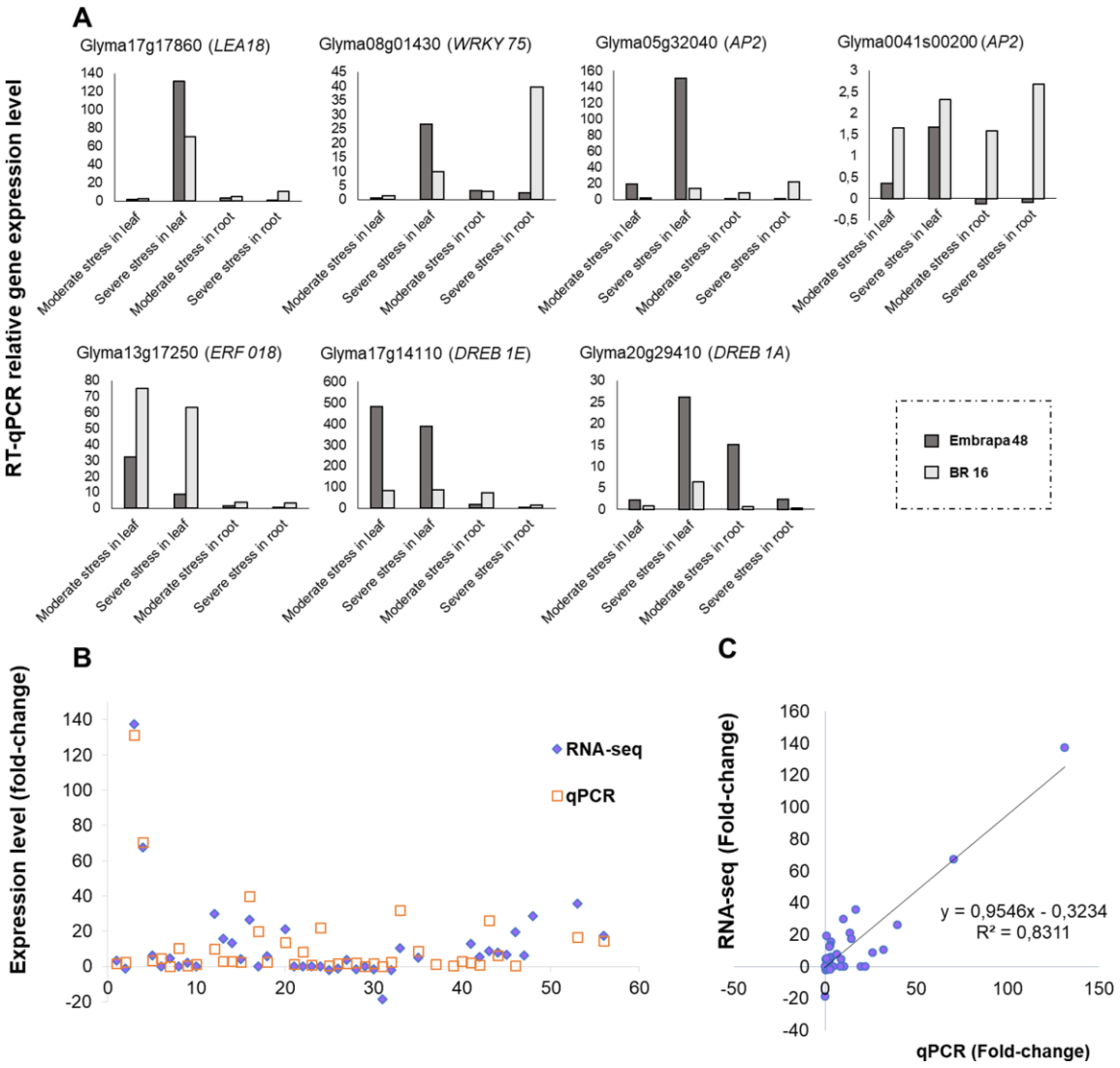


Figure 4. Validation of differentially expressed genes in soybean under water deficit stress.

(A)The RT-qPCR validation of differentially expressed genes in soybean under water deficit stress identified in RNA-seq analysis in Embrapa 48 leaf in moderate and severe library, Embrapa 48 root in a moderate and severe library, BR 16 leaf in a moderate and severe library and BR 16 root in a moderate and severe library. Fold-change in (y-axis) and treatments to each gene (x-axis).

(B) Represent the comparison between RNA-seq and qPCR results to genes validated in Embrapa 48 leaf in moderate and severe library; Embrapa 48 root in moderate and severe library; BR 16 leaf in moderate and severe library; BR 16 root in moderate and severe library. The relationship of the elements of the x-axis are best described in S1 (Figure 1).

(C) Correlation of the fold-change analyzed between RNA-seq (y-axis) and qPCR (x-axis).

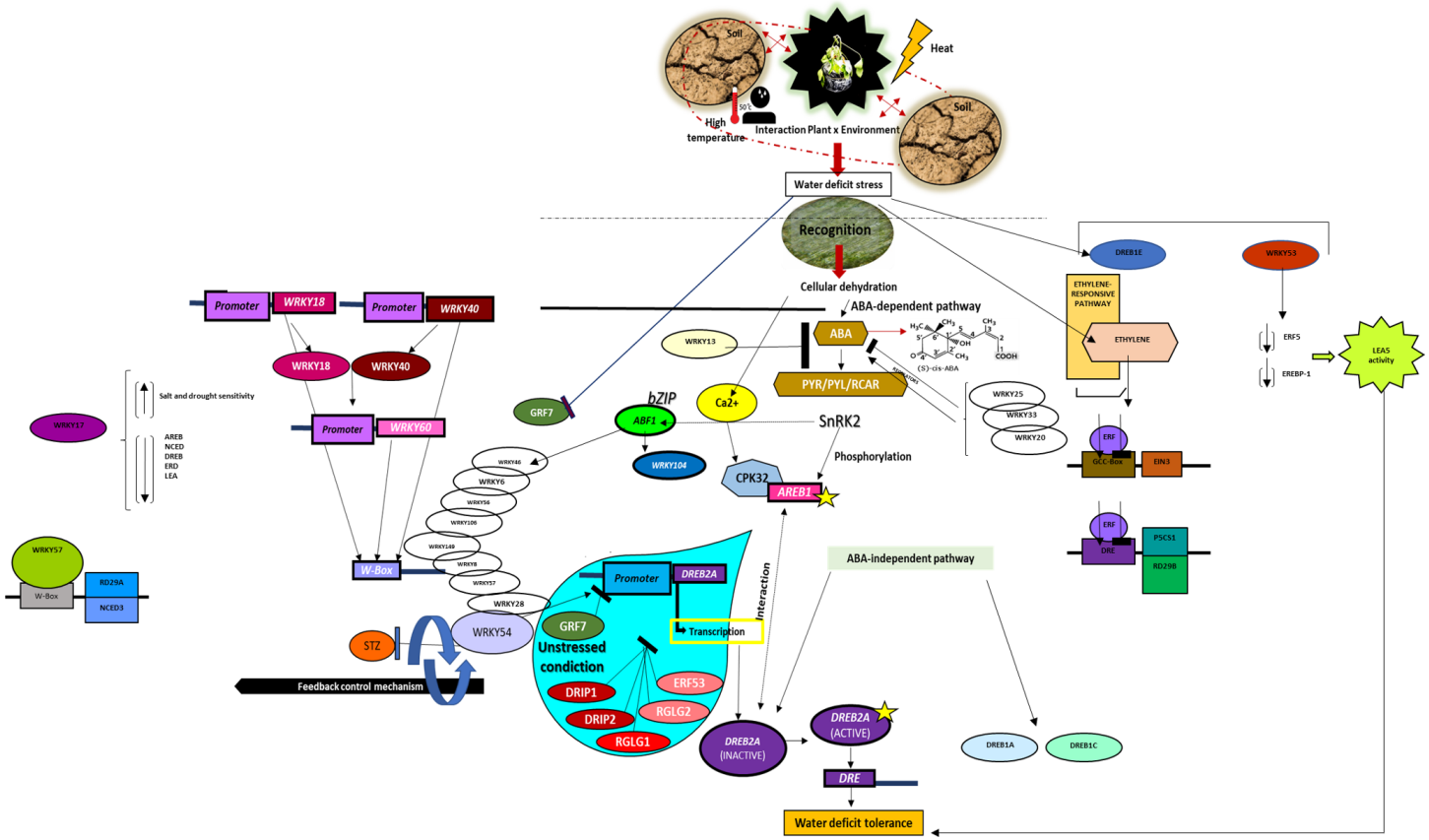
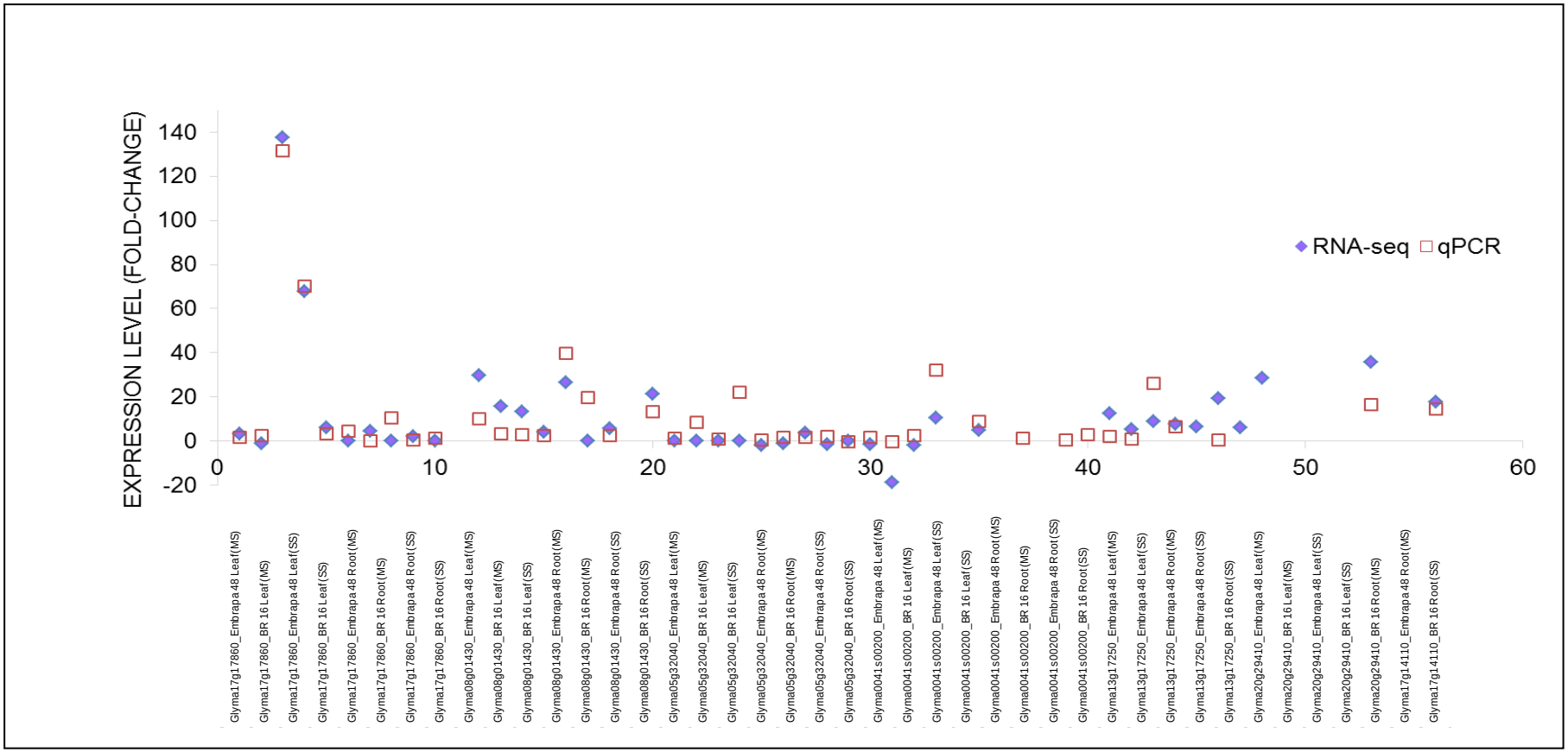


Figure 5. Cross-talk between WRKY and AP2/EREBP TF families in a ABA-mediated signaling. Transcriptional regulation activated in response to abiotic stress conditions. Transcriptional regulatory network involves interaction of Cis-regulatory elements (CREs) and TF. Phosphorylated DREB2A, AREB/ABF TFs, AREB1 and 2, ABF3 and 1 are active-genes (yellow star), resulting in induction of gene expression with DRE and ABRE-promoter region respectively. GRF, DRIP1, DRIP2, RGLG1, RGLG2 and ERF53, WRKY13, WRKY104 are negative-regulated. WRKY25, WRKY33, WRKY20, WRKY535, WRKY17 and ERF are regulators.



S1) Figure 1. Legend A graph axis X

S2) Table 1 and Table 2

Table 1 | Genes selected for RNA-seq validation.

Wm82.a1.v1	Wm82.a2.v1	Annotation	Embrapa 48 up or down expression	BR 16 up or down expression
Glyma17g17860	Glyma.17g164200	LEA 18 (Late embryogenesis abundant protein 18)	UP in leaf severe and moderate stress and UP in root moderate and severe stress	UP in leaf severe stress and DOWN in leaf moderate stress
Glyma13g17250	Glyma.13g112400	ERF 018 (AP2 domain)	UP in leaf and root	UP in leaf and root
Glyma20g29410	Glyma.20g155100	DREB 1A (AP2 domain)	UP in leaf and root	UP in leaf and root
Glyma17g14110	Glyma.17g131900	DREB 1E (AP2 domain)	UP in leaf and root	UP in leaf and root
Glyma0041s00200	Glyma.U037700	(AP2 domain)	UP in leaf and DOWN in root	DOWN in leaf and root
Glyma05g32040	Glyma.05g186700	(AP2 domain)	not occur	UP in leaf
Glyma08g01430	Glyma.08g011300	WRKY 75 (WRKY domain)	UP in leaf and root in severe and moderate stress	UP in leaf severe stress and root severe and moderate stress

Table 2 | Primers design for RNA-seq validation.

	Wm82.a1.v1	Wm82.a2.v1	Sequence
Primer Name	**Glyma17g17860-F**	Glyma.17g164200	5´AAAGGCACAGAGTGATGAAT3´
	Glyma17g17870-R		3´CTTGATGACCTTGTGTACCA5´
	Glyma08g01430-F	Glyma.08g11300	5´CCTGAAAGGTGGCAAAGAAA3´
	Glyma08g01430-R		3´CTTCACATTGCAACCTCGAT5´
	Glyma05g32040-F	Glyma.05g186700	5´GGAGGAAGAACCGAGGAGAA3´
	Glyma05g32040-R		3´GCTTTGTTGCCTCTGAAACG5´
	Glyma0041s00200-F	Glyma.U037700	5´GCTTGGACAGAAGCAACCTG3´
	Glyma0041s00200-R		3´TTCCTTTTTGCCCTTCCTT5´
	Glyma20g29410-F	Glyma.20g155100	5´AGGTCAGCCTGCCTCAACTT3´
	Glyma20g29410-R		3´GGCACCATCCCTTCTTCTTG5´
	Glyma13g17250-F	Glyma.13g112400	5´GAAAATGGGGCAAATGGGTA3´
	Glyma13g17250-R		3´GTCGGGGAAATTGAACTTGG5´
	Glyma17g14110-F	Glyma.17g131900	5´AGGATTTGGCTCGGAACGTA3´
	Glyma17g14110-R		3´CAATCAGCAACGGCATCAAT5'

827 S3) Table 1and Table 2

828 Table 1 | Promoter analysis of AP2/EREBP family

Gene symbol	Glyma	GeneID	Detailed Matrix Information	Start position	End position	Anchor position	Genomic start pos	Genomic end pos	Strand	Sequence
Ethylene-responsive transcription factor ERF096-like	Glyma.20g203500	100775830	Heat shock transcription factor B2A	16	38	27	44028311	44028333	+	aaagaaga**TTCT**tccggtagaac
			Heat shock transcription factor B2A	71	93	82	44028366	44028388	+	gaagaatc**TTCT**tccggatgaag
			Heat stress transcription factor B-2a (HSF6)	19	41	30	44028314	44028336	+	gaagattc**TTCC**ggtagaacaaa
			Secondary wall NAC binding elements	76	96	86	44028371	44028391	+	atcttcttccggat**GAAG**aaa
			NAC domain containing protein 2	73	101	87	44028368	44028396	+	agaatcttcttccggat**GAAG**aaattatt
			AGL15, Arabidopsis MADS-domain protein AGAMOUS-like 15	79	99	89	44028374	44028394	+	ttcttccggatgaa**GAAA**tta
			Sequence motif from the promoters of different sugar-responsive genes	92	110	101	44028387	44028405	+	ag**AAAT**tatttttttaat
			AC-type motifs, MYB46/MYB83-responsive elements	152	172	162	44028447	44028467	+	caacaat**GTTG**ctgggtgtat
			Transcription factor TGA2	178	196	187	44028473	44028491	+	atgtccgtg**ACTT**caaatg
AP2-like ethylene-responsive transcription factor At2g41710	Glyma.20g155100	100786651	KH and zinc finger CCCH domain-containing protein	7	17	12	39387798	39387788	+	aaaa**AAAG**aag
			KH and zinc finger CCCH domain-containing protein	21	31	26	39387784	39387774	+	ataa**AAAG**gat
			Dof zinc finger protein DOF4.2	16	38	27	39387789	39387767	+	agttaata**AAAA**ggatttttcca
			Dof zinc finger protein DOF4.2	114	136	125	39387691	39387669	+	agtaattg**AAAA**ggaaaaaacaa
			Avian C-type LTR TATA box	49	65	57	39387756	39387740	+	cagtata**TTAG**cttgag
			Arabidopsis 6b-interacting protein 1-like 1	57	75	66	39387748	39387730	+	tagcttgag**GTGA**attgtt
			Storekeeper (STK), plant specific DNA binding protein important for tuber-specific and sucrose-inducible gene expression	74	88	81	39387731	39387717	+	ttt**TAAA**cgattgta
			NAC WITH TRANSMEMBRANE MOTIF 1-LIKE 8 (NTL8/NTM1-like 8)	110	132	121	39387695	39387673	+	agcaagtaattgaaaa**GGAA**aaa
			Floral homeotic protein AGL15	113	133	123	39387692	39387672	+	aagtaattgaaaag**GAAA**aaa
			Paired amphipathic helix domain-containing protein	131	147	139	39387674	39387658	+	aaa**CAAC**tagtaaccac
			AP2/ERF and B3 domain-containing transcription factor RAV1	152	164	158	39387653	39387641	+	acc**AACA**aaaaga
			Suppressor of overexpression of CO 1 (AGL20)	151	171	161	39387654	39387634	+	gaccaacaaaaaga**GAAA**caa
			PBF (MPBF)	151	173	162	39387654	39387632	+	gaccaacaa**AAAG**agaaacaaga

Gene symbol	Glyma	GeneID	Detailed Matrix Information	Start position	End position	Anchor position	Genomic start pos	Genomic end pos	Strand	Sequence
			NAC domain containing protein 16	152	180	166	39387653	39387625	+	accaacaaaaagagaaa**CAAG**aaaaagcc
			Dof3 - single zinc finger transcription factor	166	188	177	39387639	39387617	+	aaacaagaa**AAAG**ccacttgggc
			Myb domain protein r1 (ATMYB44)	180	200	190	39387625	39387605	+	cacttgggca**GTTA**catatgc
			WRKY DNA-binding protein 50	38	54	46	39939244	39939228	+	ttctttt**GGAC**tttactt
			WUSCHEL-related homeobox 13	65	75	70	39939217	39939207	+	cat**CAAT**tagt
			Zinc finger of Arabidopsis thaliana 6 (Cold induced zinc finger protein 2)	79	89	84	39939203	39939193	+	caaa**CACT**ata
AP2-like ethylene-responsive transcription factor At2g41710	Glyma.19g138000	100786651	SQUA promoter binding proteins	110	126	118	39939172	39939156	+	ttttt**GTAC**aaaaactt
			Homeobox-leucine zipper protein HAT2	122	140	131	39939160	39939142	+	aacttaaa**TGAT**attaaaa
			Homeobox protein 34	136	154	145	39939146	39939128	+	taaaattt**TAAT**tattcac
			Flowering locus C	151	171	161	39939131	39939111	+	tcacacaaaaacaa**GAAA**tat
			Cis-element involved in SA (salicylic acid) induction of secretion-related genes via NPR1	173	187	180	39939109	39939095	+	aga**GAAG**aattaaca
			Homeobox-leucine zipper protein REVOLUTA (REV, IFL1)	189	207	198	39939093	39939075	+	ttcctat**ATGA**tttaatct
			GT1-Box binding factors with a trihelix DNA-binding domain	38	56	47	13601129	13601111	+	ttttattaa**GTTA**attagt
			Homeobox protein 25	41	59	50	13601126	13601108	+	tattaagt**TAAT**tagttaa
			Myb domain protein 57	41	61	51	13601126	13601106	+	tattaagttaa**TTAG**ttaaca
AP2-like ethylene-responsive transcription factor ANT	Glyma.17g158300	100792451	Myb domain protein 27	45	65	55	13601122	13601102	+	aagttaat**TAGT**taacaaaat
			Plant specific floral meristem identity gene LEAFY (LFY)	118	130	124	13601049	13601037	+	g**GCCA**gtggttga
			Myb domain protein 31 (AtMYB31)	116	136	126	13601051	13601031	+	ctggccagtg**GTTG**aaacttg
			Trihelix transcription factor GT-1	145	163	154	13601022	13601004	+	ctgttattg**GTTA**gaccta
			GT2-box and GT3-box motifs	158	176	167	13601009	13600991	+	gacctagtg**GTAA**attgtg
			CA-rich element	13	31	22	17377482	17377464	+	gccttca**AACC**acttatat
Ethylene-responsive transcription factor ERF017	Glyma.15g180000	100813688	ULT1 interacting factor 1 (AT4G37180)	58	74	66	17377437	17377421	+	atgtccaag**ATTC**aaat
			Myb domain protein 62 (BW62B)	91	111	101	17377404	17377384	+	aacatattcaat**TAGG**agggc
			Mammalian C-type LTR TATA box	109	125	117	17377386	17377370	+	ggctt**TAAA**aaagataa

Gene symbol	Glyma	GeneID	Detailed Matrix Information	Start position	End position	Anchor position	Genomic start pos	Genomic end pos	Strand	Sequence
			Dof zinc finger protein DOF4.7 (AT4G38000)	109	131	120	17377386	17377364	+	ggctttaaa**AAAG**ataattattc
			Class I GATA factors	116	132	124	17377379	17377363	+	aaaaa**GATA**attattca
			M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)	141	161	151	17377354	17377334	+	catcc**AACA**ggcaaagttgca
			Myb family transcription factor (G2-like family)	159	175	167	17377336	17377320	+	gcaaataaa**ATTC**ccca
			WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box	178	194	186	17377317	17377301	+	agatt**TTGA**caatgtac
			Lentivirus LTR TATA box	64	80	72	1478017	1478001	+	caa**TATA**atacggagta
			Lentivirus LTR TATA box	158	174	166	1477923	1477907	+	ata**TATA**ataaaattga
Ethylene-responsive transcription factor RAP2-2-like	Glyma.10g016500	100780763	Squamosa promoter-binding-like protein 14	66	82	74	1478015	1477999	+	atataa**TACG**gagtatt
			DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT) sequences of sporamin and beta-amylase genes	101	113	107	1477980	1477968	+	aa**AACT**atttata
			Calmodulin binding WRKY transcription factor 11	166	182	174	1477915	1477899	+	taaaa**TTGA**caatatga
			Lateral organ boundaries	2	22	12	3181355	3181335	+	gtgg**CGGC**gcaaggcgagtgc
			Promoter elements involved in MgProto (Mg-protoporphyrin IX) and light-mediated induction	12	42	27	3181345	3181315	+	aagg**CGAG**tgcctgccactcgtgctcaactg
			Brassinazole-resistant 1 protein	25	43	34	3181332	3181314	+	gccact**CGTG**ctcaactgc
			ABA (abscisic acid) inducible transcriptional activator	71	87	79	3181286	3181270	+	gtttg**CCAC**gtggcggt
			Tobacco bZip transcription activator (TAF-1)	71	89	80	3181286	3181268	+	gtttgcc**ACGT**ggcggtgt
			BBES1/BZR1-like protein 3 (AT4G18890)	73	91	82	3181284	3181266	+	ttgcca**CGTG**gcggtgtct
			Cis-acting element conserved in various PAL and 4CL promoters	99	119	109	3181258	3181238	+	ttttttacaggtg**GGTG**taa
Dehydration-responsive element binding protein 2	Glyma.06g042100	732579	Myb domain protein 55	103	123	113	3181254	3181234	+	tttacagg**TGGG**tgtaacggg
			Myb-related protein 3R-1 (PC-MYB1)	113	133	123	3181244	3181224	+	ggtgt**AACG**ggattggcgcgt
			E2F transcription factor 3	120	142	131	3181237	3181215	+	cgggatt**GGCG**cgtgggtcgtca
			Coupling element 3 (CE3), non-ACGT ABRE	123	141	132	3181234	3181216	+	gattgg**CGCG**tgggtcgtc
			Rice iron-related transcription factor 2	123	141	132	3181234	3181216	+	gattgg**CGCG**tgggtcgtc
			Arabidopsis thaliana signal-responsive gene1, Ca2+/ calmodulin binding protein homolog to NtER1 (tobacco early ethylene-responsive gene)	126	142	134	3181231	3181215	+	tgg**CGCG**tgggtcgtca
			WRKY transcription factor 8	164	180	172	3181193	3181177	+	ttttg**TTGC**cttttcta

Gene symbol	Glyma	GeneID	Detailed Matrix Information	Start position	End position	Anchor position	Genomic start pos	Genomic end pos	Strand	Sequence
Dehydration-responsive element-binding protein 2C-like	Glyma.02g261700	100798277	Soybean embryo factor 4	181	191	186	3181176	3181166	+	aa**TTTT**tatca
			Root hair-specific element with a 3-nucleotid spacer between left part (LP) and right part (RP)	175	199	187	3181182	3181158	+	tttctaaatttttat**CACG**ttgtga
			DIVARICATA 1 (myb related R-R-type factor AT5G58900)	180	198	189	3181177	3181159	+	aaattttt**TATC**acgttgtg
			bZIP protein G-Box binding factor 1	8	26	17	44767736	44767754	+	aagcgc**CACG**tcaccacat
			Phytochrome interacting factor3-like 5	8	26	17	44767736	44767754	+	aagcgc**CACG**tcaccacat
			Myb family transcription factor REVEILLE 1	48	64	56	44767776	44767792	+	atcagaaa**TATC**tagag
			Heat stress transcription factor A-6b	49	71	60	44767777	44767799	+	tcagaaatA**TCT**agagaatgcta
			SHI related sequence 1 (STYLISH 1) (secondary DNA binding preference)	147	159	153	44767875	44767887	+	acact**AGCC**agaa
			Myb-like protein of Petunia hybrida	158	178	168	44767886	44767906	+	aaataaaacc**GATA**agccgaa
			Shoot-apical-meristem arrest 2 (AT5G08750)	169	181	175	44767897	44767909	+	ataa**GCCG**aaaaa
			LOB domain-containing protein 19	171	191	181	44767899	44767919	+	aag**CCGA**aaaacattgagttc
Dehydration-responsive element-binding protein 1E	Glyma.01g216000	100789097	Arabidopsis thaliana class A heat shock factor 1a	98	120	109	54671112	54671090	+	actaaat**ATTC**ttaaaatacttg
			GAAA motif involved in pollen specific transcriptional activation	121	137	129	54671089	54671073	+	tccga**GAAA**ctctttga
			Homeobox 51, Late Meristem Identity 1	2	20	11	54671208	54671190	+	tttaaaaa**TAAT**ggttttt
			Homeobox-leucine zipper protein ATHB-7	157	175	166	54671053	54671035	+	agtaaaattg**ATTG**aacac
			Myb domain protein 56	81	101	91	54671129	54671109	+	tatattaaac**GTTA**tgtacta
			NAC domain containing protein 5	112	134	123	54671098	54671076	+	aaat**ACTT**gtccgagaaactctt
			Salt tolerance zinc finger (ZAT10)	161	183	172	54671049	54671027	+	aaattgattgaacact**CACT**gtt
			Squamosa promoter-binding-like protein 9	91	107	99	54671119	54671103	+	gttat**GTAC**taaatatt
			Zea mays MYB-related protein 1 (transfer cell specific)	62	80	71	54671148	54671130	+	atatata**TATA**taatataa
			Zinc finger and SCAN domain containing 4	171	185	178	54671039	54671025	+	a**ACAC**tcactgttaa
Uncharacterized LOC100527194	Glyma.03g112700	100527194	Sunflower homeodomain leucine-zipper protein Hahb-4	14	32	23	31932050	31932032	+	gtgaagtatt**ATTA**acata
			Growth-regulating factor 9	34	44	39	31932030	31932020	+	tttc**TGAC**ata
			Homeobox-leucine zipper protein ATHB-6	66	84	75	31931998	31931980	+	caaaa**TAAT**aattaaaata

Gene symbol	Glyma	GeneID	Detailed Matrix Information	Start position	End position	Anchor position	Genomic start pos	Genomic end pos	Strand	Sequence
			Homeobox-leucine zipper protein ATHB-53	137	155	146	31931927	31931909	+	ttaat**CTAT**aatttttta
			Indeterminate(ID)-domain 5 protein (RAVEN)	191	209	200	31931873	31931855	+	agctt**TTGT**ttttttat

829

830

831 Table 2 | Promoter analysis of WRKY family

Gene symbol	Glyma	GeneID	Detailed Matrix Information	Start position	End position	Anchor position	Genomic start pos	Genomic end pos	Chromosome	Strand	Sequence
WRKY53	Glyma.19g094100	732586	Sunflower homeodomain leucine-zipper protein Hahb-4	126	144	135	33282000	33281982	19	+	tatatatatc**ATTA**nnnnn
Probable WRKY transcription factor 26	Glyma.18g056600	100127397	GA-regulated myb gene from barley	13	33	23	4951780	4951800	18	+	atatgtattc**GTTA**acgtttt
			OCS-like elements	13	33	23	4951780	4951800	18	+	atatgtattcgtta**ACGT**ttt
			GAZ-like 3 (AT5G22990)	25	37	31	4951792	4951804	18	+	ta**ACGT**tttcctt
			Squamosa promoter-binding-like protein 9	80	96	88	4951847	4951863	18	+	tatgt**GTAC**tttaaatt
			Target of early activation tagged 1 (RAP2.7) (tertiary DNA binding preference)	85	95	90	4951852	4951862	18	+	gtac**TTTA**aat
			Class I GATA factors	130	146	138	4951897	4951913	18	+	gaata**GATA**aaatttaa
			Botrytis-susceptible1 (MYB108)	143	163	153	4951910	4951930	18	+	ttaaggatcaa**TTAG**gaaatt
			WUSCHEL-related homeobox 13	148	158	153	4951915	4951925	18	+	gat**CAAT**tagg
			Myb domain protein 59	147	167	157	4951914	4951934	18	+	ggatcaat**TAGG**aaatttggg
			Homeodomain glabrous 9	167	183	175	4951934	4951950	18	+	gcatat**TAAA**tatttta
Probable WRKY transcription factor 40	Glyma.13g370100	100798375	Homeobox-leucine zipper protein ATHB-24	32	50	41	45563654	45563672	13	+	actacttttt**ATTA**ctaat
			KANADI 4 (Aberrant Testa Shape) (secondary DNA binding preference)	80	92	86	45563702	45563714	13	+	ttgatg**ATTC**ata
			Homeobox 51, Late Meristem Identity 1	102	120	111	45563724	45563742	13	+	tattttaa**TAAT**gatttta
			Transcription factor NAC2	110	138	124	45563732	45563760	13	+	taatgattttatatgca**CATG**aaataagg
			Anther-specific myb gene from tobacco	131	151	141	45563753	45563773	13	+	aaataaggtt**GTTA**gactgag
			ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	150	168	159	45563772	45563790	13	+	aggat**ACAT**gtgtttttag
			ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	178	196	187	45563800	45563818	13	+	acaca**ACAA**atgtctattc
Probable WRKY transcription factor 28	Glyma.08g082400	100816891	WRKY DNA-binding protein 45	3	19	11	6227329	6227345	8	+	tatgt**TTAA**ctttttac
			WRKY DNA-binding protein 70	5	21	13	6227331	6227347	8	+	tgttt**AACT**ttttacac
			Reproductive Meristem 1 (secondary DNA binding preference)	14	24	19	6227340	6227350	8	+	tttt**ACAC**gtt
			Homeobox-leucine zipper protein HAT2	93	111	102	6227419	6227437	8	+	atagtaaa**TGAT**caaagta

Gene symbol	Glyma	GeneID	Detailed Matrix Information	Start position	End position	Anchor position	Genomic start pos	Genomic end pos	Chromosome	Strand	Sequence
Probable WRKY transcription factor 41	Glyma.08g021900	100127375	GCN4, conserved in cereal seed storage protein gene promoters, similar to yeast GCN4 and vertebrate AP-1	1	17	9	1765392	1765376	8	+	tgatt**TGAG**tcattttt
			Circadian clock associated 1 (secondary DNA binding preference)	21	37	29	1765372	1765356	8	+	ttatta**GATA**ttgatta
			PBF (MPBF)	90	112	101	1765303	1765281	8	+	gagcactca**AAAG**ggttaacaca
			Upstream sequence elements in the promoters of U-snRNA genes of higher plants	146	162	154	1765247	1765231	8	+	aatctc**CCAC**ttcctcc
WRKY25 protein	Glyma.08g011300	100127378	TESMIN/TSO1-like CXC 2	6	36	21	899849	899879	8	+	aggtaattaaaata**CAAA**ttattaaaactaa
			Zinc finger of Arabidopsis thaliana 6 (Cold induced zinc finger protein 2)	27	37	32	899870	899880	8	+	ttaa**AACT**aaa
			Myb-domain transcription factor werewolf	35	55	45	899878	899898	8	+	aaaatatact**GTTC**tgtattc
			ABORTED MICROSPORES	49	67	58	899892	899910	8	+	tgtatt**CATC**tgatcttga
			GATA transcription factor 19 (HANABA TARANU LIKE 2, HANL2)	54	68	61	899897	899911	8	+	tcatct**GATC**ttgat
			ABA response elements	76	92	84	899919	899935	8	+	gtattgt**ACGT**gaaaat
			NACL-inducible gene 1	78	90	84	899921	899933	8	+	att**GTAC**gtgaaa
			Squamosa promoter binding protein-like 14	76	92	84	899919	899935	8	+	gtattg**TACG**tgaaaat
			Two-component response regulator ARR14 (secondary DNA binding preference)	102	112	107	899945	899955	8	+	gaa**GATT**ttca
			Arabidopsis NAC domain containing protein 19	101	127	114	899944	899970	8	+	ggaagattttcacgt**TACG**tcacacag
			Myb domain protein 98	104	124	114	899947	899967	8	+	agattttcac**GTTA**cgtcaca
			Phytochrome interacting factor3-like 5	105	123	114	899948	899966	8	+	gattttt**CACG**ttacgtcac
			Rice transcription activator-1 (RITA), basic leucin zipper protein, highly expressed during seed development	111	127	119	899954	899970	8	+	cacgtt**ACGT**cacacag
			AGL15, Arabidopsis MADS-domain protein AGAMOUS-like 15	113	133	123	899956	899976	8	+	cgt**TACG**tcacacagaaaatt
			P1BS, PHR1 binding sequences	123	141	132	899966	899984	8	+	cacaga**AAAT**tctactatt
			Heat shock transcription factor B2A	124	146	135	899967	899989	8	+	acagaaaa**TTCT**actattattta
			AP2/ERF and B3 domain-containing transcription factor RAV1	157	169	163	900000	900012	8	+	gtc**AACA**aaataa
			Homeobox protein 25	181	199	190	900024	900042	8	+	gctacagt**TAAT**tgttttt
Probable WRKY transcription factor 40	Glyma.07g023300	100807966	Myb family transcription factor REVEILLE 1	40	56	48	1775459	1775443	7	+	agtaaaaa**CATC**ttgaa

Gene symbol	Glyma	GeneID	Detailed Matrix Information	Start position	End position	Anchor position	Genomic start pos	Genomic end pos	Chromosome	Strand	Sequence
			Heat stress transcription factor A-6b	41	63	52	1775458	1775436	7	+	gtaaaaac**ATCT**tgaaaaaaaaa
			Trihelix transcription factor GT-2	100	118	109	1775399	1775381	7	+	tagatgttg**GTCA**aaattc
			DREB and EAR motif protein 5 (RAP2.9)	11	31	21	4652431	4652411	6	+	atgaaaca**CCGA**taaaaccac
			Mammalian C-type LTR TATA box	18	34	26	4652424	4652408	6	+	accga**TAAA**accacgaa
			NAC domain containing protein 13	19	47	33	4652423	4652395	6	+	ccgataaaaccacgaat**CAAG**caaattgc
			NAC secondary wall thickening promoting factor 1	26	46	36	4652416	4652396	6	+	aaccacgaatcaa**GCAA**attg
			Nodulin consensus sequence 2	42	56	49	4652400	4652386	6	+	aattgc**CTCA**taata
			Drosophila initiator motifs	72	82	77	4652370	4652360	6	+	tt**TCAT**tcggc
Protein Polar localization during asymmetric division and redistribution-like	Glyma.06g061800	102667044	CCHC-type zinc knuckle protein	92	108	100	4652350	4652334	6	+	tttcg**TTCA**cttgtttt
			Wheat NAC-domain DNA binding factor (DNA binding site II)	140	166	153	4652302	4652276	6	+	gtttgtaatggctat**GACG**tcatgaaa
			NAM-like protein CUP-SHAPED COTYLEDON 1	143	171	157	4652299	4652271	6	+	tgtaatggctatgacgt**CATG**aaaacacg
			bZIP transcription factor Elongated Hypocotyl 5	149	167	158	4652293	4652275	6	+	ggctatg**ACGT**catgaaaa
			Root hair-specific element with a 2-nucleotid spacer between left part (LP) and right part (RP)	153	177	165	4652289	4652265	6	+	atgacgtcatgaaaa**CACG**tattgt
			NAC domain containing protein 3	153	179	166	4652289	4652263	6	+	atgacgtcatgaaaa**CACG**tattgtaa
			bZIP protein G-Box binding factor 1	162	180	171	4652280	4652262	6	+	tgaaaa**CACG**tattgtaaa
			RAP2.2, involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	11	21	16	5404783	5404793	3	+	aa**ATCT**atata
			Calmodulin-binding transcription activator 1 (AtSR2)	22	38	30	5404794	5404810	3	+	aga**CGCG**aaatatttac
			Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	43	61	52	5404815	5404833	3	+	catgctc**TTAC**gttgattt
Probable WRKY transcription factor 33	Glyma.03g042700	102666898	AS1/AS2 repressor complex binding motif II	52	60	56	5404824	5404832	3	+	acg**TTGA**tt
			WRKY transcription factor 7	50	66	58	5404822	5404838	3	+	ttacg**TTGA**tttttcaa
			Target of early activation tagged (EAT) 2 (secondary DNA binding preference)	73	83	78	5404845	5404855	3	+	ca**ACCT**aggca
			NAC with transmembrane mtif 1-like 8 (NTL8/NTM1-like 8)	121	143	132	5404893	5404915	3	+	atagattatcttct**GAAG**taaca
			NAC domain containing protein 5	125	147	136	5404897	5404919	3	+	atta**TCTT**ctgaagtaacattct

Gene symbol	Glyma	GeneID	Detailed Matrix Information	Start position	End position	Anchor position	Genomic start pos	Genomic end pos	Chromosome	Strand	Sequence
			Heat stress transcription factor B-3	136	158	147	5404908	5404930	3	+	aagtaaca**TTCT**atatttcttta
			Opaque-2 regulatory protein	169	185	177	5404941	5404957	3	+	ttgggtata**TCAT**cata
			PHAVOLUTA (Homeobox-leucine zipper protein ATHB-9)	173	191	182	5404945	5404963	3	+	gtatatca**TCAT**aataaat

832