

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

**INVESTIGAÇÃO DO MECANISMO DE TOLERÂNCIA À SECA EM
PLANTAS TRANSGÊNICAS DE SOJA QUE EXPRESSAM UMA
OSMOTINA DE *SOLANUM NIGRUM***

Lariane Frâncio

Porto Alegre, Julho de 2019

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

**INVESTIGAÇÃO DO MECANISMO DE TOLERÂNCIA À SECA EM
PLANTAS TRANSGÊNICAS DE SOJA QUE EXPRESSAM UMA
OSMOTINA DE *SOLANUM NIGRUM***

Lariane Frâncio

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

Orientadora: Profa. Dra. Maria Helena Bodanese Zanettini

Coorientadora: Dra. Luisa Abruzzi de Oliveira Busatto

Porto Alegre, Julho de 2019

INSTITUIÇÃO E FONTE FINANCIADORA

Este trabalho foi desenvolvido nos laboratórios (i) Genética Molecular Vegetal e (ii) Cultura de Tecidos e Transformação Genética de Plantas do Departamento de Genética da Universidade Federal do Rio Grande do Sul.

Recebeu apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e Chamada INCT – MCTI/CNPq/CAPES/FAPs nº 16/2014, Ativos 369 Biotecnológicos Aplicados a Seca e Pragas em Culturas Relevantes para o Agronegócio (INCT Biotec 370 Seca-Pragas) [88887.136360/2017-00 - 465480/2014-4].

“Eu tentei 99 vezes e falhei, mas na centésima tentativa eu consegui, nunca desista de seus objetivos mesmo que esses pareçam impossíveis, a próxima tentativa pode ser a vitoriosa.”

Albert Einstein

AGRADECIMENTOS

Agradeço em primeiro plano a Deus, autor da minha fé, pela companhia.

Aos meus amados e queridos pais, Cláudio e Odete, pela vida, pelos meus primeiros passos, pela minha formação, pelo inabalável laço de amor que nos une, pela incansável luta e incentivo nos momentos difíceis e nas alegrias e principalmente, por acreditarem em meus ideais.

À minha querida irmã Larissa, que muito me incentivou e por todo o seu amor.

Ao meu amado namorado Aléxis Kin, pelo constante apoio, compreensão, dedicação, companheirismo, carinho, amor e incentivo para lutar pelos meus sonhos.

À Universidade Federal do Rio Grande do Sul, pela excelente oportunidade. Em especial ao secretário do Programa de Genética e Biologia Molecular, Elmo por toda ajuda e paciência ao longo do meu mestrado.

À minha orientadora, Profa. Dra. Maria Helena Bodanese-Zanettini, pela oportunidade, confiança, apoio, orientação, disponibilidade e dedicação para realização deste trabalho. Agradeço também pelas horas dispendidas, pelos conselhos, incentivos e pelas palavras de conforto e carinho, que certamente levarei para minha vida.

À minha coorientadora Dra. Luisa A. de Oliveira Busatto pelos ensinamentos e conhecimentos passados, pela orientação, pela amizade, paciência, empenho e disposição constante para o êxito de minhas atividades.

Ao Dr. Ricardo Weber pela ajuda, ensinamentos, risadas, positivismo e principalmente pelos conselhos em momentos difíceis ou de dúvidas.

Aos colegas e amigos do Laboratório de Genética Molecular Vegetal em especial à colega Giulia pela ajuda no desenvolvimento deste trabalho, mas principalmente pela amizade, conselhos e reflexões positivas para a vida. À Letícia Gal pelo apoio técnico e amizade.

Aos demais familiares e amigos que estiveram ao meu lado durante esta jornada.

E a todos que de alguma maneira contribuíram para a realização deste trabalho.

Muito obrigada!

RESUMO

Estresses ambientais são amplamente responsáveis por limitar o rendimento da soja. Para mitigar os impactos gerados pela deficiência hídrica, ferramentas de biologia molecular estão sendo utilizadas para o desenvolvimento de plantas geneticamente modificadas mais tolerantes à seca. Osmotinas ou proteínas semelhantes a osmotinas (OLPs) são moduladas sob estresses abióticos e bióticos. Estudos prévios do nosso grupo de pesquisa mostraram que dois eventos independentes (B1 e B3) de plantas transgênicas de soja expressando uma osmotina (SnOLP) de *Solanum nigrum* tiveram um incremento na tolerância à seca. No entanto, o mecanismo de tolerância ainda não foi elucidado. O presente trabalho tem como objetivo investigar o mecanismo de tolerância à seca das plantas de soja transgênica. As duas linhagens transgênicas foram cultivadas em recipientes plásticos (1L) contendo substrato, em sala de crescimento com temperatura e umidade controladas. Parte das plantas, no estágio vegetativo (V6), foram submetidas ao estresse hídrico pela supressão da irrigação durante sete dias. Para as demais plantas (amostras controle), o suprimento de água foi mantido. Variáveis fisiológicas (conteúdo de água na folha e fluorescência da clorofila) foram monitoradas para avaliar a condição das plantas sob estresse. Os resultados confirmaram que as plantas transgênicas apresentaram melhor desempenho quando comparadas às plantas não transgênicas. Dois conjuntos (*pools*) de folhas de quatro plantas por tratamento foram coletados após o período de estresse hídrico. O RNA total extraído das folhas foi utilizado para a construção de bibliotecas que foram enviadas para sequenciamento. No total, doze bibliotecas foram produzidas, e o método de sequenciamento usado foi o *pair-end*. As sequências resultantes do RNA-seq foram agrupadas, validadas e mapeadas no genoma da soja, disponível no banco de dados Phytozome. Os dados obtidos foram normalizados pelo DESeq2 e analisados por duas abordagens diferentes: (i) bottom-up (DESeq2) e (ii) top-down (The Transcriptogramer). Na primeira, parte-se de uma lista de genes diferencialmente expressos (DEGs) para a identificação das rotas metabólicas. Na segunda, a partir de categorias de ontologia gênica (GOs) diferencialmente expressas são identificados os DEGs. Utilizando o software DESeq2, 115 DEGs foram detectados quando plantas transgênicas foram comparadas com plantas não transgênicas (NT). Na comparação das plantas transgênicas B1 com plantas NT na condição de seca, foram detectados 87 DEGs, sendo 43 genes induzidos e 44 genes reprimidos. Da mesma forma, comparando as plantas B3 com plantas NT sob seca, foram detectados 98 DEGs, dos quais 54 genes induzidos e 44 genes reprimidos. Trinta e seis (59%) do total de 61 genes induzidos, e 34 (63%) dos 54 genes reprimidos são compartilhados pelos eventos B1 e B3. A comparação entre cada evento transgênico na situação irrigado e seca e das plantas não transgênicas na situação irrigado e seca, também foi realizada. Um total de 2044 e 1505 DEGs foram identificados em plantas transgênicas B1 e B3, respectivamente ($\log_2\text{FoldChange} \geq 2$ e $\text{padj} \leq 0,001$). Em relação ao evento B1, 769 genes induzidos e 1275 genes reprimidos. A mesma análise para B3 revelou 541 genes induzidos e 964 genes reprimidos. A exclusão dos DEGs em comum com plantas não

transgênicas resultou em 395 (46.5%) genes induzidos e 234 (13.6%) genes reprimidos compartilhados pelos eventos B1 e B3. Desse total, 261 e 58 foram induzidos nos eventos B1 e B3, respectivamente; 251 e 91 foram reprimidos em B1 e B3, respectivamente. No método Transcriptogramer, os perfis de expressão foram projetados em uma lista de ordenamento e categorias de ontologia gênica. Os resultados permitiram a identificação de sete DEGs, em dez GOs diferencialmente expressas. Foi realizada, também, a análise de DEGs nas outras 6829 GOs (não diferencialmente expressas) e identificados 121 DEGs. As GOs e os DEGs estão envolvidos principalmente em processos biológicos relacionados à regulação do ciclo celular e manutenção do ambiente celular. Replicação, reparo e metilação do DNA; biossíntese do ribossomo e síntese metabólica secundária também foram moduladas. A validação dos DEGs será realizada pela técnica RT-qPCR.

ABSTRACT

Environmental stresses are largely responsible for limiting soybean yield. In order to mitigate the impacts generated by water deficit, molecular biology tools are being used to develop genetically modified plants more tolerant to drought. Osmotins or osmotins-like proteins (OLPs) are modulated under abiotic and biotic stresses. Previous studies of our research group showed that two independent events (B1 and B3) of soybean transgenic plants expressing a *Solanum nigrum* osmotin (SnOLP) had an increment in drought tolerance. However, the tolerance mechanism has not been elucidated yet. The present study aims to investigate the drought tolerance mechanism of the transgenic soybean plants. The two homozygous transgenic lines were cultivated in plastic containers (1L) containing substrate in a growth room with controlled temperature and humidity. Part of the plants, in vegetative stage (V6), was submitted to water deficit by the irrigation suppression during seven days. For the other plants (control samples) the water supply was maintained. Physiological variables (relative leaf water content and chlorophyll fluorescence) were monitored in order to evaluate the plant condition under stress. Results confirmed that the transgenic plants presented better performance when compared to the non-transgenic (NT) plants. Two pools of leaves from four plants per treatment were collected after the period of water deficit. The total RNA extracted from the leaves was used for the construction of libraries which were sent for sequencing. In total, twelve libraries were produced, and the sequencing method used was the pair-end. The resulting sequences of the RNA-seq were clustered, validated and mapped into the soybean genome available at Phytozome database. The data obtained were normalized by DESeq2 and analyzed using two different approaches: (i) bottom-up (DESeq2) and (ii) top-down (The Transcriptogramer). In the first, a list of differentially expressed genes (DEGs) is produced and then the metabolic pathways are identified. In the second, from differentially expressed gene ontology categories (GOs) the DEGs are identified. Using DESeq2 115 DEGs were detected when transgenic plants were compared to NT plants. In the comparison of B1 versus NT plants in the drought condition, 87 DEGs were detected, in which 43 up and 44 downregulated. In the same way, comparing B3 versus NT plants under drought, 98 DEGs were detected, in which 54 up and 44 downregulated. Thirty six (59%) out of the 61 upregulated and 34 (63%) out of the 54 downregulated genes are shared by B1 and B3 events. The comparison between each transgenic event (irrigated and dry situation) and non-transgenic plants (irrigated and dry situation) was also performed. A total of 2044 and 1505 DEGs were identified in B1 and B3 events respectively ($\log_2\text{FoldChange} \geq 2$ and $\text{padj} \leq 0.001$). Regarding B1 event, 769 upregulated and 1275 downregulated. In the same way, for B3 event were detected 541 upregulated and 964 downregulated. Exclusion of DEGs in common with non-transgenic plants resulted in 395 (46.5%) upregulated and 234 (13.6%) downregulated shared by B1 and B3 events. Of this total, 261 and 58 were upregulated in B1 and B3 events, respectively; 251 and 91 were downregulated in B1 and B3, respectively. In the Transcriptogramer method, the expression profiles were projected

on a list of ordering and GOs. The results allowed selecting seven DEGs within ten differentially expressed GOs. Analyzes of DEGs in the order 6829 soybean GOs (not differentially expressed) were performed and 121 DEGs were identified. The GOs and DEGs identified are mainly involved in biological processes related to cell cycle regulation and maintenance of internal cell environment. DNA replication, repair and methylation; ribosome biosynthesis and secondary metabolic synthesis were also modulated. Validation of the DEGs is being carried out by the RT-qPCR technique.

LISTA DE ABREVIATURAS E SIGLAS

ABA - Ácido abscísico (do inglês, *Abscisic Acid*)

bp - Pares de bases do DNA (do inglês, *base pairs*)

cDNA - Ácido desoxirribonucleico complementar (do inglês, *Complementary DNA*)

CTAB - Brometo de Cetiltrimetilamônio (do inglês, *Cetyltrimethylammonium Bromide*)

DEGs – Genes diferencialmente expressos (do inglês, *Differential Expressed Genes*)

DNA - Ácido Desoxirribonucleico (do inglês, *Deoxyribonucleic Acid*)

DNase - Desoxirribunoclease (do inglês, *Deoxyribonuclease*)

dNTP - Desoxirribonucleotídeos trifosfatados (do inglês, *Deoxyribonucleotide*)

GFP - Proteína fluorescente verde (do inglês, *Green Fluorescent Protein*)

GO – Ontologia gênica (do inglês, *Gene Ontology*)

MS - Meio de cultura Murashige e Skoog

NaCl - Cloreto de sódio

NT - Não transgênica

OLP – Proteína semelhante à Osmotina (do inglês, *Osmotin-like proteins*)

pH - Potencial Hidrogeniônico

PCR - Reação em cadeia da DNA polimerase (do inglês, *Polymerase Chain Reaction*)

pI - Ponto isoelétrico

RNA - Ácido Ribonucleico (do inglês, *Ribonucleic Acid*)

RNase - Ribonuclease (do inglês, *Ribonuclease*)

RT-qPCR - Reação em Cadeia da DNA Polimerase de forma quantitativa precedida de Transcrição Reversa (do inglês, *Reverse Transcription quantitative Polymerase Chain Reaction*)

SUMÁRIO

1 INTRODUÇÃO	11
1.2 A importância da soja	11
1.2 Estresse hídrico na soja.....	12
1.3 Osmotinas	14
1.4 Ferramentas para análise do transcriptoma	16
2 OBJETIVOS	18
2.1 Objetivo Geral	18
2.2 Objetivos Específicos	18
3. ARTIGO CIENTÍFICO.....	19
5 CONSIDERAÇÕES FINAIS.....	59
6 APÊNDICE: RESULTADOS PRELIMINARES.....	60
REFERÊNCIAS	65

1 INTRODUÇÃO

1.2 A importância da soja

Dentre as leguminosas, a soja [*Glycine max* (L.) Merrill] é uma das culturas mais importantes em nível econômico mundial. Rica em proteínas e óleos vegetais, é um alimento básico para humanos e outros animais, (Yin et al., 2014; Chen et al., 2018; Nawaz et al., 2017; Xu et al., 2017; Zhang; Song, 2017), além de fornecer matéria prima para a produção de biocombustíveis (Zhang; Song, 2017).

A cultura da soja tem elevada significância na composição das exportações de países como Estados Unidos da América, Brasil e Argentina. O plantio da soja, em larga escala, no Brasil, teve início em 1960 e, atualmente, é o principal responsável pelo crescente volume de exportações do agronegócio e o consequente avanço da economia nacional. O Brasil é hoje o segundo maior produtor desta oleaginosa, com uma produção de 113.923 milhões de toneladas, competindo com os Estados Unidos, sendo este o maior produtor mundial, com uma produção de 117.208 milhões de toneladas (CONAB, 2017). Estima-se que a safra brasileira 2019/2020 seja de 123 milhões de toneladas, colocando o Brasil como maior produtor do mundo. Porém, este valor apenas será alcançado se não ocorrerem problemas climáticos no período (CONAB, 2019).

Nos últimos 20 anos, o crescimento médio anual da produção de soja no Brasil foi de 3.5 milhões de toneladas, o que representa um incremento de 13,4% ao ano. A produção aumentou de 26 milhões de toneladas, na safra 1996/1997, para 95 milhões de toneladas, na safra 2015/2016. A área plantada cresceu um milhão de hectares por ano e o aumento da produtividade foi de aproximadamente 34 kg/hectare/ano. Esse crescimento ocorreu, em grande parte, pelas ações de Pesquisa, Desenvolvimento e Inovação (PD&I); assim como pela capacidade do produtor brasileiro de absorver as tecnologias disponíveis (Balbinot Junior et al., 2017).

Assim como outras leguminosas, a soja é sensível a uma variedade de estresses abióticos (Oh; Komatsu, 2015). As variações de muitos fatores ambientais desempenham um papel importante no controle da produção, qualidade e composição química da semente de soja. Entre muitos desses fatores ambientais, o estresse hídrico, é um dos mais importantes, causando reduções significativas no rendimento e qualidade da soja (Wijewardana; Reddy; Bellaloui, 2019). Para se adaptar a ambientes adversos,

as plantas desenvolveram uma série de respostas de defesa que envolve a modulação da expressão de muitos genes (Chen et al., 2019).

1.2 Estresse hídrico na soja

Entre os estresses abióticos, a seca é o principal estresse ambiental que limita a sobrevivência das plantas e a produção global dos alimentos (Tripathi et al., 2016; Fàbregas; Fernie, 2019). Devido à mudança climática global, o impacto do estresse por deficiência hídrica está recebendo mais atenção e atualmente é um desafio-chave na agricultura. Estima-se que as reduções de rendimento relacionadas à seca para as principais culturas no mundo atingirão mais de 50% até 2050 (Li et al., 2009). O impacto negativo do estresse hídrico na produtividade agrícola está presente em todos os estádios de desenvolvimento, desde a germinação até a maturação das sementes. Neste cenário será desafiador atender às demandas alimentares da população global (Sharma et al., 2018).

Estudos estão sendo realizados para investigar os mecanismos moleculares utilizados pelas plantas para combater a seca, com o objetivo de obter culturas tolerantes (Fàbregas; Fernie, 2019).

De acordo com a revisão de Manavalan et al. (2009), os mecanismos utilizados pelas plantas para superar o estresse hídrico, podem ser classificados em três grupos: (i) fuga da seca; (ii) prevenção da seca; e (iii) tolerância à seca. A fuga da seca ocorre quando a planta completa seu ciclo de vida durante o período de abastecimento de água suficiente, antes do início da seca. O sistema de plantio de soja precoce (*Early Soybean Planting System*), utilizado no sul dos EUA, é um exemplo desta estratégia. A segunda estratégia envolve mecanismos que ajudam a planta a manter o alto nível de água durante os períodos de estresse, seja pela eficiente absorção de água pelas raízes ou pela redução da evapotranspiração das partes aéreas. Já na terceira estratégia a planta mantém o turgor e continua o metabolismo, mesmo com baixo potencial hídrico, por tolerância protoplasmática ou síntese de osmoprotetores, osmólitos ou solutos compatíveis.

As plantas respondem à seca reprogramando suas vias transcricionais, proteômicas e metabólicas, a fim de proteger as células dos danos causados pelo estresse. Ferramentas de genômica, transcriptômica e proteômica têm sido amplamente

utilizadas para estudar as vias de resposta ao estresse em muitas espécies de plantas. Assim, atualmente, sabe-se que muitas enzimas metabólicas, enzimas de desintoxicação, chaperonas, proteínas quinases e fatores de transcrição desempenham papéis essenciais nas respostas das plantas ao estresse (revisado em Fàbregas; Fernie, 2019).

Quando as plantas se desenvolvem sob deficiência hídrica, passam por uma transição de fase saturada (a utilização da água não depende do conteúdo de água no solo), para uma segunda fase (o uso da água está diretamente relacionado à disponibilidade de água no solo). Essa transição está associada a uma redução na condutância estomática. A condutância estomática é uma variável chave que influencia a troca de gases e de vapor de água nas folhas (Manavalan et al., 2009). Liu et al. (2003) demonstraram que o estresse hídrico diminui a taxa relativa de expansão foliar, a condutância estomática e o turgor foliar, enquanto ocorre um aumento do conteúdo de ácido abscísico (ABA) na folha e no xilema. O controle da condutância estomática é um traço fisiológico promissor para o desenvolvimento da tolerância à seca em soja. Já o ajuste osmótico (OA, do inglês *Osmotic Adjustment*), definido como o acúmulo de solutos, ocorre nos tecidos vegetais em resposta a um aumento da deficiência hídrica.

O ajuste osmótico mantém a condutância estomática e fotossíntese em menores potenciais hídricos, atrasa a senescência e morte foliar, melhora o crescimento das raízes e aumenta a captação de água do solo, à medida que a disponibilidade de água diminui (Manavalan et al., 2009). James et al. (2008) demonstraram correlação significativa entre o conteúdo relativo de água (RWC, do inglês *Relative Water Content*) e OA em sete genótipos de soja, mostrando que os genótipos que apresentaram maior potencial osmótico, apresentaram também menor declínio de RWC, sugerindo assim a capacidade desses em manter o turgor por mais longo período.

Manavalan et al. (2009), revisaram as principais características fisiológicas relacionadas aos mecanismos de tolerância à seca em soja. Em resposta ao estresse, as plantas desencadeiam vários mecanismos de defesa, como a síntese de proteínas que irão participar em vias metabólicas relacionadas ao aumento da tolerância ao estresse.

As raízes são o primeiro órgão a perceber as tensões provocadas pela seca. Durante o estresse hídrico, a produção de ABA nas raízes e seu transporte para as folhas são considerados um sinal chave, levando ao fechamento estomático e ao crescimento reduzido. O pH, as citocininas, o precursor de etileno, o malato e outros fatores estão

envolvidos na comunicação raiz-parte aérea durante o estresse hídrico. As plantas podem adaptar-se à seca desenvolvendo uma raiz principal mais longa, para ajudá-las a alcançar as camadas mais profundas do solo, onde a água está disponível (Oh; Komatsu, 2015). Além disso, um sistema extensivo de raízes fibrosas pode ser útil para aumentar a captação da umidade superficial do subsolo. O aumento do peso seco, comprimento total, volume e número de raízes laterais são indicadores de maior tolerância à seca em soja (Manavalan et al., 2009).

O desenvolvimento das plantas de soja ocorre em dois estádios: vegetativo (V) e reprodutivo (R). A etapa vegetativa compreende os estádios VE (emergência de cotilédones) a Vn (numerados de acordo com a quantidade de folhas trifolioladas totalmente expandidas). O estágio R inclui: (i) R1 e R2 (florescimento); (ii) R3 e R4 (formação de vagens) e (iii) R5-R8 (maturação e enchimento de sementes). Em soja, quando a seca foi imposta durante os últimos estádios vegetativos e/ou iniciais do reprodutivo (R1-R2), um grande aumento no crescimento radicular foi observado. O crescimento da raiz foi menos afetado quando a seca foi imposta no estágio R4 e cessou no R5. Em outro estudo, foi demonstrado que as plantas de soja submetidas à situação de seca antes da floração tiveram maiores rendimentos de sementes do que as plantas submetidas ao estresse após a floração, isso porque já haviam desenvolvido um robusto sistema radicular antes do florescimento (revisado em Fàbregas; Fernie, 2019).

Avanços no melhoramento genético, sequenciamento do genoma/transcriptoma e caracterização funcional de genes da soja, fornecem oportunidades para entender os padrões globais da expressão gênica e sua associação com o desenvolvimento de fenótipos específicos, bem como para o estabelecimento de estratégias de melhoramento via engenharia genética.

1.3 Osmotinas

Osmotinas (*osmotine-like proteins*, OLPs) são proteínas produzidas pelas plantas e moduladas durante diferentes estresses abióticos e bióticos. Por apresentarem sequência e estrutura muito similar às taumatinas, proteínas doces encontradas nos frutos do arbusto africano *Thaumatococcus daniellii*, foram classificadas dentro da família das PR-5, também conhecida por *Thaumatococcus-like proteins* (TLPs) (Stintzi et al., 1991).

As osmotinas foram identificadas em diferentes espécies, incluindo *Cucumis sativus* L., *Arabidopsis thaliana*, *Arabidopsis lyrata* L., *Carica papaya* L., *Glycine max* L., *Medicago truncatula*, *Vitis vinifera* L., *Populus trichocarpa*, *Sorghum bicolor* L., *Oryza sativa* L. e *Zea mays* L. Propriedades como atividade antifúngica e tolerância ao estresse osmótico tem tornado as osmotinas alvos potenciais para o desenvolvimento de culturas transgênicas (revisado em Hakim et al., 2018).

Em plantas, foram descobertas inicialmente em células de tabaco, que se adaptaram a ambientes com baixo potencial osmótico. Esta função está ligada à sua ação como osmólito compatível (Abdin et al., 2011). As osmotinas desempenham, também um papel importante na proteção das membranas plasmáticas sob baixo potencial de água (Viktorova et al., 2012).

As osmotinas não atuam como fator de transcrição por não conterem domínios de ligação ao DNA (Abdin et al., 2011) e não possuem motivos alergênicos (Liu et al., 2010). Trata-se de proteínas secretoras que apresentam três domínios com dobramento semelhante ao de outras taumatinas, como: a zeamatina (Min et al., 2004) e a proteína PR-5d do tabaco (Koiwa et al., 1999). Entre os domínios I e II há a formação de uma fenda ácida. Alterações químicas nos resíduos ao redor da fenda ácida explicam o processo particular de atividade antifúngica.

A expressão de osmotinas em plantas transgênicas, como tomate, tabaco, algodão, soja, sésamo, cenoura, amora e arroz, tem levado à tolerância das mesmas à seca. Tem sido relatado que a sua expressão em plantas transgênicas tem causado aumento na expansão foliar, no teor de clorofila e no teor relativo de água. Esses resultados indicam a capacidade da osmotina de proteger a clorofila e a maquinaria fotossintética sob condições de seca (revisado por Hakim et al., 2018).

Apesar de numerosos estudos terem sido realizados para determinar o papel da osmotina na tolerância ao estresse hídrico, seu mecanismo de ação permanece desconhecido.

Nosso grupo de pesquisa está investigando este mecanismo em plantas de soja transgênicas que superexpressa uma osmotina (*SnOLP*) de *Solanum nigrum* L. var. *americanum*. O estudo, está sendo desenvolvido com plantas transgênicas de soja de dois eventos independentes (B1 e B3) que expressam o gene *SnOLP* previamente obtidas por Weber et al. (2014). O gene *SnOLP* isolado do genoma de *Solanum nigrum* L. var. *americanum*, mostra uma alta homologia com o que codifica a osmotina do

tabaco (Campos et al., 2002). Já foi demonstrado que a superexpressão da osmotina de tabaco em diferentes espécies de plantas conferem tolerância a estresses abióticos, especialmente salinidade e seca (Barthakur et al., 2001; Noori e Sokhansanj, 2008; Parkhi et al., 2009; Goel et al., 2010; Das et al., 2011; Subramanyam et al., 2011; Subramanyam et al., 2012). Os resultados obtidos por Weber et al (2014) mostraram que a expressão de *SnOLP* em soja transgênica melhorou as respostas fisiológicas e a produtividade das plantas quando submetidas à deficiência hídrica, destacando o potencial desse gene para aplicações biotecnológicas.

1.4 Ferramentas para análise do transcriptoma

Na era pós-genômica, a utilização de abordagens sistemáticas ou ômicas, aumentaram a compreensão de complexas redes de regulação molecular associadas a mecanismos de adaptação e tolerância a estresses tanto abióticos quanto bióticos. Entre as ômicas, a transcriptômica é utilizada para análise do padrão de expressão de RNAs codificantes e não-codificantes. A metabolômica é uma ferramenta poderosa para analisar o elevado número de metabólitos. A proteômica, por sua vez, analisa perfis de expressão e modificações de proteínas, colaborando para o entendimento das redes regulatórias (Cramer et al., 2011).

As tecnologias de sequenciamento de RNA (RNA-Seq) estão se tornando cada vez mais prevalentes. Uma dessas técnicas é o RNA-seq, que além de ter alto rendimento possui vantagens tais como: não limitar a detecção de transcritos, baixo sinal de fundo, quantificação de alta precisão dos níveis de expressão e requerimento relativamente baixo de amostras (Yin et al., 2014). Essa técnica tem sido amplamente utilizada em plantas. Em soja, várias análises do transcriptoma em resposta à seca já foram registradas (Chen et al., 2013; Belamkar et al., 2014; Shin et al., 2015; Rodrigues et al., 2015; Arraes et al., 2015; Tripathi et al., 2016).

As análises de sequenciamento de RNA frequentemente fornecem um alto número de genes diferencialmente expressos entre as amostras. Portanto, as atuais ferramentas de análise de bioinformática são valiosas para a exploração de resultados e identificação de genes individuais de interesse biotecnológico (Conesa et al., 2016).

Para análises de dados de transcriptome duas estratégias tem sido usadas: (i) *top-down* (por exemplo, DESeq2) e (ii) *bottom-up* (por exemplo, The Transcriptogramer). Na primeira, parte-se de uma lista de genes diferencialmente expressos (DEGs, do inglês *Differentially Expressed Genes*) para a identificação das rotas metabólicas. Na segunda, a partir de categorias de ontologia gênica diferencialmente expressas são identificados os DEGs.

O DESeq2 é um pacote disponível na plataforma online Bioconductor (<http://www.bioconductor.org/packages/release/bioc/html/DESeq2.html>) para identificar os DEGs, com base na contagem de transcritos (*reads*). Este método possibilita a normalização e a análise estatística dos dados.

O Transcriptogramer é uma ferramenta de bioinformática que permite a análise da expressão gênica em escala de genoma completo, que utiliza como base um método de ordenamento de interações entre proteínas de um organismo. Assim, o transcriptograma apresenta e analisa dados de transcrição em uma escala genômica ampla e facilitando a interpretação biológica. A grande utilidade do transcriptograma é a facilidade de identificação de regiões do genoma com maior alteração transcricional e a possibilidade de relacionar as regiões alteradas com os processos biológicos (Rybarczyk-Filho et al., 2011;)

2 OBJETIVOS

2.1 OBJETIVO GERAL

O presente estudo tem como objetivo investigar o mecanismo de tolerância à seca de plantas transgênicas de soja, que superexpressam o gene *SnOLP* codificador de uma osmotina de *Solanum nigrum*.

2.2 OBJETIVOS ESPECÍFICOS

- a) Avaliar variáveis fisiológicas nas plantas transgênicas em condições controle e sob deficiência hídrica.
- b) Comparar a atividade transcricional de plantas transgênicas e não-transgênicas de soja e identificar os genes diferencialmente expressos em condições controle e sob deficiência hídrica;

3. ARTIGO CIENTÍFICO

Investigation of drought tolerance mechanism in soybean transgenic plants expressing a *Solanum nigrum* osmotin (SnOLP)

Lariane Frâncio, Luisa A. de Oliveira-Busatto, Giulia R. Faillace, Frank Guzman, Ricardo L. M. Weber, Ciliana Rechenmacher, Débora Favero, Christian Bredemeier, Rita M. C. de Almeida, Maria Helena Zanettini.

O Capítulo 1 será apresentado em forma de um artigo científico em preparação que posteriormente a sua conclusão será submetido a um periódico especializado. As figuras e tabelas principais estão distribuídas ao longo do texto para facilitar a leitura e interpretação. As figuras e tabelas suplementares encontram-se no final do artigo após as referências bibliográficas.

**Investigation of drought tolerance mechanism in soybean transgenic plants
expressing a *Solanum nigrum* osmotin (SnOLP)**

**Lariane Frâncio ¹, Luisa A. de Oliveira-Busatto ¹, Giulia R. Faillace ¹, Frank
Guzman ², Ricardo L. M. Weber ¹, Ciliana Rechenmacher ¹, Débora Favero ³,
Christian Bredemeier ³, Rita M. C. de Almeida ⁴, Maria Helena Zanettini ^{1*}.**

¹Programa de Pós-Graduação em Genética e Biologia Molecular, Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul (UFRGS), Brasil.

²Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul (UFRGS), Brasil.

³Departamento de Plantas de Lavoura, Universidade Federal do Rio Grande do Sul (UFRGS), Brasil.

⁴Instituto de Física, Universidade Federal do Rio Grande do Sul (UFRGS), Brasil.

*Corresponding author: Universidade Federal do Rio Grande do Sul, Instituto de Biociências, Departamento de Genética, Programa de Pós-Graduação em Genética e Biologia Molecular, (PPGBM), Av. Bento Gonçalves, 9500 – Prédio 43312, sala 212D, CEP:91501-970, Postal Code 15053, Porto Alegre, RS, Brazil.

E-mail address: maria.zanettini@ufrgs.br

1. Introduction

Soybean [*Glycine max* (L.) Merrill] is one of the most important crops in the world economy, being a source of protein and vegetable oil, staple food for humans and other animals, besides supplying raw material for the production of biofuels. Soybean cultivation has high significance in the composition of exports from countries such as the United States of America, Brazil and Argentina (CONAB, 2017). Despite this positive scenario, soybean yield in most of the producing areas is affected by numerous biotic and abiotic stresses (Zhang et al., 2107).

Among the abiotic stresses, drought is the main environmental stress threatening plant survival and crop productivity worldwide. Due to global climate change, the impact of water-deficit stress in plants is receiving more attention and at present is a key challenge in agriculture (Fàbregas; Fernie, 2019). New strategies for the improvement of drought responses in crop species are therefore valuable.

Plants respond to drought by reprogramming their transcriptional, proteomic, and metabolic pathways in order to protect cells against environmental stress conditions. Substantial effort has been directed to elucidate the molecular mechanisms used by plants to cope with drought and to translate this knowledge into suitable biotechnological tools with the aim to obtain crops that are more drought tolerant. (Fàbregas; Fernie, 2019). Among the available tools, the development of transgenic crops by the overexpression of genes that confer drought tolerance has been proven to be successful and has a significant impact on agricultural production.

Osmotins (osmotin-like proteins, OLPs) are proteins produced by plants during different abiotic and biotic stresses (Hakim et al., 2018). They are members of the pathogenesis-related proteins (PR-5) family and have been identified in different plants species (Faillace et al., 2019). Properties such as antifungal activity and osmotic stress tolerance have made osmotins potential targets for the development of transgenic crops. The overexpression of an osmotin encoding-gene in a plant species may confer tolerance to various biotic and abiotic stresses. For example, the overexpression of an osmotin gene in tobacco (Barthakur et al., 2001) and soybean (Weber et al., 2014) plants enhanced tolerance to drought stress.

Numerous studies have been carried out to determine the physiological role of osmotins in stress tolerance, but their mechanism of action remains unknown (Hakim et

al., 2018). Using high through-put sequencing techniques such as RNA-Seq, it is possible to investigate the expression of genes and identify potential candidates for functional characterization studies (Le et al., 2012).

Our research group obtained two independent events of soybean transgenic plants expressing an osmotin gene (*SnOLP*) from *Solanum nigrum* L., which shows high homology with a tobacco osmotin gene (Weber et al., 2014). In the present study, we have used two strategies in the analysis of RNA-seq data: DESeq2 (top-down) (Love et al., 2014) and The Transcriptogramer (bottom-up) (Rybarczyk-Filho et al., 2011). These two methods are complementary since the first presents a list of differentially expressed genes and from this it is possible to identify differentially expressed categories of gene ontology. In contrast, the second method provides a list of differentially expressed gene ontology categories and within these it is possible to identify differentially expressed genes. The main goal of this study is to contribute to the elucidation of the osmotin role in the response to water deficit.

2. Material and Methods

2.1 Plant Material

Two T8 homozygous transgenic lines (B1 and B3) expressing the *SnOLP* gene, previously obtained from independent transformation events (Weber et al., 2014) and non-transgenic plants (NT) were analyzed.

Plants were grown in 1-L plastic pots filled with a substrate/soil mixture. The experiment was conducted in a growth room at $28\text{ }^{\circ}\text{C} \pm 2\text{ }^{\circ}\text{C}$ air temperature, 16 h / 8 h day/night photoperiod, $60\% \pm 10\%$ air relative humidity and $200\text{ }\mu\text{mol m}^{-2}\text{ s}^{-1}$ photosynthetically active radiation. Four seeds were sown per pot, leaving two plants by thinning after plant emergence.

2.2 PCR analysis

Genomic DNA was extracted from transgenic and non-transgenic soybean plants using CTAB method procedure described by Doyle and Doyle (1987). PCR analyses were performed with specific primers for *SnOLP* gene (Table 1). The PCR reaction

mixture consisted of 250 ng of DNA template, 0.25 mM of dNTPs, 1.5 mM MgCl₂, 1X Taq Buffer, 2 units of Taq® DNA Polymerase (Invitrogen), 0.25 µM of each primer. Reactions were hot-started (3 min at 94 °C) and subjected to 25 cycles as follows: 1 min at 94 °C; 1 min at 50 °C and 2 min at 72 °C with a final extension of 72 °C for 5 min. PCR-amplified products were analyzed in 1,2% agarose gel, stained with gel red and visualized under UV light.

2.3 Soybean drought stress assay

Treatments consisted of two water regimes (control/irrigated and drought stress/non-irrigated) imposed 26 days after sowing at V6 growth stage (six nodes on the main stem with fully developed trifoliolate leaves). From plant emergence up to V6 growth stage, pots were weighted daily and irrigated with water (if necessary) to maintain substrate moisture at approximately 90% field capacity. After this growth stage, plants of transgenic lines and NT were separated into two groups: one group (control) continued to receive regular irrigation to maintain 90% field capacity, and the other group was subjected to drought stress induced by watering suspension for seven days. The experiment was carried out in a randomized block design with four biological replicates. Each pot is considered as an experimental unit.

2.4 Physiological analysis

In order to characterize drought stress and physiological responses, relative leaf water content (RWC) and chlorophyll fluorescence (maximum quantum yield of photosystem II) were determined at the uppermost fully expanded leaves.

RWC was evaluated seven days after irrigation suspension. This parameter estimates the current water content relative to the maximum water content the leaves can hold at full turgidity and is given by the relation: $RWC(\%) = [(fresh\ weight - dry\ weight) / (turgid\ weight - dry\ weight)] \times 100$. Briefly, the fresh weight of three leaves per plant was determined. After that, leaves were left floating on distilled water, in Petri dishes, for 24 hours and then the turgid weight was recorded. Finally, leaves were dried at 65°C for 48 hours and the dry weight was determined (Virginia et al., 2012).

The maximum quantum yield of photosystem II (ϕ PSII) was determined under dark-adapted (30 min prior to the measurement) and natural light conditions (steady-state light-adapted fluorescence) using a portable pulse modulation fluorometer (Model OS1-FL, Opti-Sciences, Hudson, USA). ϕ PSII is given by the relation F_v/F_m (dark-adapted) or F_v'/F_m' (light-adapted), where F_v and F_v' denote the variable fluorescence and F_m and F_m' the maximal fluorescence from dark and light-adapted leaves, respectively (Baker, 2008). Measurements were performed daily at the same time (11:00am) between four and eight days after irrigation suspension on the adaxial surface of the uppermost fully expanded leaf (in four leaves per replicate).

2.5 Statistical Analysis

RWC and chlorophyll fluorescence data of NT, B1 and B3 lines grown under the two water regimes were compared by the F-test. When significant ($p < 0.05$), means were compared by Tukey's test using the software SPSS Statistics version 17.0 (SPSS Inc., Chicago, USA).

2.6 RNA Isolation and Library Preparation

Two pools of leaves from four plants were collected for each treatment and frozen in liquid nitrogen for gene expression analysis and sequencing. Samples were stored at -80°C until the analyses were performed. Total RNA extraction was performed with Trizol buffer (Thermo Fisher Scientific) according to the manufacturer-recommended protocol. The quality of samples was verified using spectrophotometric measurements performed on NanoDrop equipment. Before library preparation, RNA was cleaned with a RNA Clean & ConcentratorTM – 25 kit (Zymo Research). The RNA-Seq libraries were prepared using a SureSelect Strand-Specific RNA Library Prep for Illumina Multiplexed Sequencing Kit (Agilent Technologies) according to the manufacturer's instructions. In total, 12 RNA-Seq libraries were constructed (each genotype and treatment in duplicate).

2.7 Differential Expression Analysis

The presence of adapters and the quality of reads produced by the sequencing were determined for each library using FastQC software (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Based on this data, the Trim Galore! (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) was used to eliminate the reads sequences with a quality below 30, as well as, the sequences of the universal Illumina adapters. Then, the cleaned reads were anchored with TopHat2 (Kim et al., 2013) to the soybean genome (Phytozome - *Glycine max* Wm82.a2.v1) using the "fr-firststrand" option. This option was followed considering the methodology used in libraries preparation (Strand-specific RNAseq) and was confirmed by the `infer_experiment.py` script of the RSeQC program (Wang et al., 2012).

The counting tables of reads mapped on each gene were generated with the featureCounts module of the Subread software (Liao et al., 2013) from the anchor bam extension files generated by TopHat2. The criteria used to create the counting tables were: (1) fragments (reads pairs) were counted instead of individual reads, (2) pairs of reads that were anchoring on different chromosomes or anchoring on the same chromosomes but on different strand were not considered, (3) reads that anchor in multiple places of the genome were not considered.

Considering the RSeQC results, the option "reversely stranded" was selected. In order to perform the statistical analysis and to identify the differentially expressed genes, the counting tables were analyzed using the DESeq2 package 1.12.3 (Love et al., 2014) of the Bioconductor using a FDR (False Discovery Rate) of 0.001 and a fold change of 2. For each comparison of treatments, all genes with differential expression lower than 0.05 were separated into two groups: one group formed by up-regulated genes and another group formed by down-regulated genes.

The protein sequences of up-regulated genes were obtained from Ensembl Plants and compared to the UniRef Enriched KEGG Orthology (UEKO) database (Guedes et al., 2011) using local blastX (Altschul et al., 1997). The blastX results were processed in the MySQL software (Oracle) and the KO codes obtained were viewed on the iPATH3 web server (Darzi et al., 2018).

The Transcriptogramer method (Rybarczyk-Filho et al., 2011) was also used to analyze the transcriptome data. This method begins by producing a list of proteins of a

given organism that are listed as begin associated to another protein, as listed in STRING database (<http://string.embl.de/>) with a minimum 800 string-score. Transcriptograms are gene expression profiles projected on the ordered list. The soybean ordered list presents 20317 genes (Oliveira-Busatto et al., submitted). This list is ordered in such a way that genes whose products participate in the same metabolic pathways are localized in nearby positions.

The input data for Transcriptogramer was the DeSeq2 normalized read counts. For the term enrichment analysis, an ordering file, a dictionary (converting the Glyma codes from version 1.1 to 2.0 of the soybean genome) and the file list containing the categories of Gene Ontology (GO) were used (Oliveira-Busatto et al., submitted). The window average used in this study was $r = 125$.

2.8 Gene expression validation

The total RNA previously purified was treated with DNase I (Promega) and cDNA synthesis was performed using 1 μ g of DNA-free RNA, M-MLV Reverse Transcriptase SystemTM (Invitrogen) and oligo (dT) primers. To evaluate relative gene expression, the first-strand cDNA reaction product was diluted at 1:50.

Primers were designed using Primer3 software (<http://frodo.wi.mit.edu/>) (Koressaar; Remm, 2007; Untergasser et al., 2012) according to the gene sequences (Table 1). RT-qPCR was performed on an Applied Biosystems 7500 Real-Time PCR System (Thermo Fisher Scientific). PCR-cycling conditions were implemented as follows: 5 min at 94°C, followed by 40 repetitions of 10 s at 94 °C, 15 s at 60 °C and 15 s at 72 °C. A melting curve analysis was performed at the end of the PCR run, over the range of 55-99 °C, increasing the temperature stepwise by 0.1 °C/s. Each 20 μ L reaction comprised 10 μ L cDNA (1:50 dilution), 1x PCR buffer (Invitrogen), 50 mM MgCl₂, 5 mM of each dNTP, 10 μ M of each primer, 2.5 μ L of SYBR-Green solution (1:100,000, Molecular Probes Inc., Eugene, OR), and 0.06 U Platinum Taq DNA Polymerase (Termo Fisher Scientific). Technical quadruple reactions were performed for each sample. Reactions without cDNA were used as negative controls.

Transcript levels of the 15 differentially expressed genes will be evaluated. The F-box protein and a Metalloprotease were used for gene expression normalization. The

expression data analyses will be performed using the 2- $\Delta\Delta$ Ct method (Livak and Schmittgen, 2001).

Table 1 - Primer set designed for PCR and RT-qPCR.

Gene	Sequence 5' -3'	Sense	Amplicon (pb)
<i>SnOLP</i> – PPS1	CGCGGATCCATGGGCTACTTGAG ATCT	Forward	744
<i>SnOLP</i> – PCPT	CCCAAGCTTTTAC TTGGCCACTTCATC	Reverse	
<i>Glyma02g32058</i>	GTGGGATGCTTGCAATGCTG	Forward	167
	AGTTGCTTTGAGTCCTTGAGTT	Reverse	
<i>Glyma15g19921</i>	GAGGCCTGTGCTCCTTCTTC	Forward	181
	ACATACGGGCAGTGAAGAACA	Reverse	
<i>Glyma14g07070</i>	CTCCGTGGATGAGTTCACCC	Forward	171
	TGCCCATCCAATGCTCCAAT	Reverse	
<i>Glyma.12G111600</i>	GGTGTTCCCTTGGACAGCTT	Forward	170
	GTTCTCCACTCCATGCAGCT	Reverse	
<i>Glyma19g40430</i>	TGCAGCAAACCGTGAATGTG	Forward	186
	TTTGTGAAGGCTCCCCATCC	Reverse	
<i>Glyma.19G115400</i>	ATTCCTCCTTGCTGAACCGG	Forward	184
	CGCTTATTCCTGGCCCTCAA	Reverse	
<i>Glyma11g35750</i>	TTTGACCGTGCTGAGATTGC	Forward	181
	TGAATCCCTGGAAACGGTGAG	Reverse	
<i>Glyma14g36750</i>	TGCAAGATAACAGGGCTGACTC	Forward	200
	CGTCTACGCTTAAAACCACAGC	Reverse	
<i>Glyma14g14290</i>	TGGCGGCATTGAAGATGAAG	Forward	248
	TGCATTGATCTCGCTATCGC	Reverse	
<i>Glyma11g19870</i>	ACTGATGGATTGGCAGATGC	Forward	234
	TTCAGCCTTCCCCACAAAAG	Reverse	
<i>Glyma08g11200</i>	TTCCGTTGCTCACTCCAATG	Forward	195
	AGGCCTTGTTGATCACTTGC	Reverse	
<i>Glyma06g16110</i>	TTTGTTGCCCTCCTGGAATG	Forward	171
	AGGGTTGGGAATTGGATTGC	Reverse	
<i>Glyma03g39460</i>	TCATCACGAGCCTTTGGTTG	Forward	178
	TCCGAGTGCATCCACAAATC	Reverse	
	AACATCACAAGCACGAGCAC	Reverse	

Table 1 - continuation.

<i>Glyma15g06160</i>	ACCCTGATAGGCACTTGGTTG	Forward	195
	AAACCCACTGAGCATCCATG	Reverse	
<i>Glyma10g42630</i>	ACAGACCCCAACCAACATTG	Forward	244
	AACATCACAAGCACGAGCAC	Reverse	
<i>F-Box</i>	GATAGGGAAATGTTGCAGGT	Forward	93
	CTAATGGCAATTGCAGCTCTC	Reverse	
<i>Metalloprotein</i>	ATGAATGACGGTTCATGTA	Forward	114
	GGCATTAAAGGCAGCTCACTCT	Reverse	
<i>CYP2</i>	CGGGACCAGTGTGCTTCTCA	Forward	154
	CCCCTCCACTACAAAGGCTCG	Reverse	

3. Results and discussion

3.1 Transgenic status confirmation

PCR analyzes for *SnOLP* gene detection were performed with DNA extracted from leaves of B1 e B3 plants. Non-transgenic plants (NT) were used as negative control. The 744 bp fragment was detected in PCR products from transgenic plants, whereas no amplification was observed in NT plants (Figure 1).

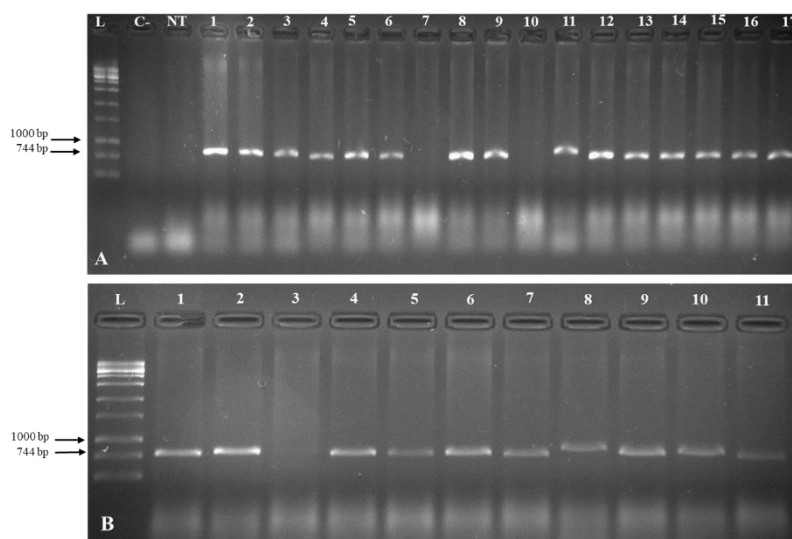


Figure 1 - Transgene detection by PCR analysis. PCR products of *SnOLP* gene (774 bp) amplified from DNA extracted of transgenic plants and negative controls (NT, Bragg). (A) L = 1 kb DNA Ladder; 1-17 = B1 plants; NT = negative control; C- = reaction negative control. (B) L = 1 kb DNA Ladder; 1-11 = B3 plants;

3.2 Physiological analyses

In order to evaluate the response of transgenic and non-transgenic plants under water deficit, irrigation was suspended for seven days. After the stress period, transgenic plants presented better phenotypic appearance when compared to non-transgenic plants (Figure 2). Under water deficit NT plants showed a visible loss of turgidity, while transgenic plants (B1 and B3) did not present this phenotype. This phenotypic trait was also observed by Weber et al (2014). In addition, other studies have shown similar results in cotton, mulberry and tomato plants that overexpress a tobacco osmotin (Parkhi et al., 2009; Goel et al., 2010; Das et al., 2011) and in tomato plants that overexpress a *Solanum nigrum* (*OLP*) osmotin (Kumar et al., 2016).

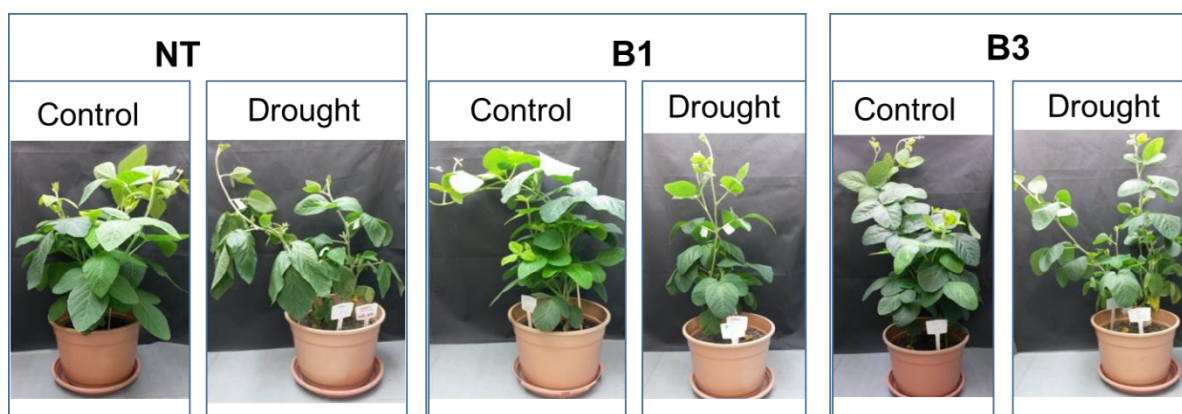


Figure 2 - Plants under water deficit. Non-transgenic (NT) and transgenic plants expressing SnOLP after withholding water for seven days. Non-transgenic Bragg plants and T8 transgenic homozygous plants from each transgenic line (B1 and B3) were grown in 1-L plastic pots for 26 days in a growth chamber. Plants were monitored daily for wilting.

Physiological variables (relative leaf water content and chlorophyll fluorescence) were monitored under control conditions and under water deficit. Relative water content (RWC) is an important indicator of the plant health status, as the water potential is an important determinant of metabolic activity and leaf survival (Kumar et al., 2016). The loss of water under the stress condition was noted in both transgenic and NT plants. However, the RWC reduction was significantly lower in B1 (32.9%) and B3 (34.6%) than in NT (45.7%) plants under water stress. No significant differences were observed among B1, B3 and NT plants under control conditions (Table 2). This result indicates that the SnOLP expression helps plants to retain water under drought conditions, as

previously demonstrated by Weber et al. (2014). Similarly, Kumar et al. (2016) also showed that transgenic tomatoes presented higher RWC values (%) when compared to non-transgenic plants during some treatments, including drought. According to the author the higher RWC may be due to membrane stability or osmotic adjustment, which would facilitate the accumulation or compartmentation of solutes or otherwise regulate metabolic or structural changes. Therefore, further analyses are necessary to confirm this relation in soybean transgenic plants.

Table 2 - Relative leaf water content – RWC.

Plant	Condition	RWC	%Reduction (Relative to control)
B1	Control	89,2 a	
	Stress	59,8 b	32.9
B3	Control	89,5 a	
	Stress	58,5 b	34.6
NT	Control	91,9 a	
	Stress	49,9 c	45.7

Significance of analysis of variance - F test ($P < 0,01$)
Multiple comparison of means - Duncan test ($P = 0,05$)

Regarding the photosynthetic capacity, B1 and B3 presented higher maximum quantum yield of photosystem II (ϕ PSII) under water deficit (Figure 3). This result confirmed the better performance of the transgenic plants when compared to the NT plants. Drought stress can reduce the quantum yield of PSII and this reduction was more pronounced in NT plants. According to Mathobo et al. (2017) plants under water stress demonstrate a decline in PSII quantum yield due to the effects of water shortage on the stomatal aperture and photosynthetic rate, reducing the efficiency of the photosynthesis photochemical phase and increasing the energy dissipation, to avoid oxidative stress. In addition, a decrease of F_v/F_m in leaves adapted to dark is related to the increase of the water stress and decrease of the foliar photosynthetic potential (Zhang et al., 2016). Weber et al. (2014) have demonstrated that B1 and B3 plants, besides presenting better photosynthetic capacity, also presented greater net assimilation of CO_2 , stomatal conductance and transpiration when compared to NT plants, suggesting that these plants are able to better utilize the internal carbon dioxide. Tomato plants overexpressing an

osmotin, presented higher proline content, water content, PSII activity, endocytosis activity, K^+ content, better vascular conductivity, better biomass roots and fruit numbers when compared to non-transformed plants, under abiotic and biotic stress conditions (Kumar et al., 2016).

As recently reviewed by Hakim et al. (2018) the main effects of osmotins expression in transgenic plants submitted to salt and drought stresses are: (i) osmolarity cellular maintenance by metabolic and structural alterations or by the solutes compartmentalization; (ii) increases in leaf expansion; (iii) increases in chlorophyll; (iv) increases in the RWC. According to the author, these effects indicate the capacity of osmotin to protect chlorophyll and photosynthetic machinery under drought conditions.

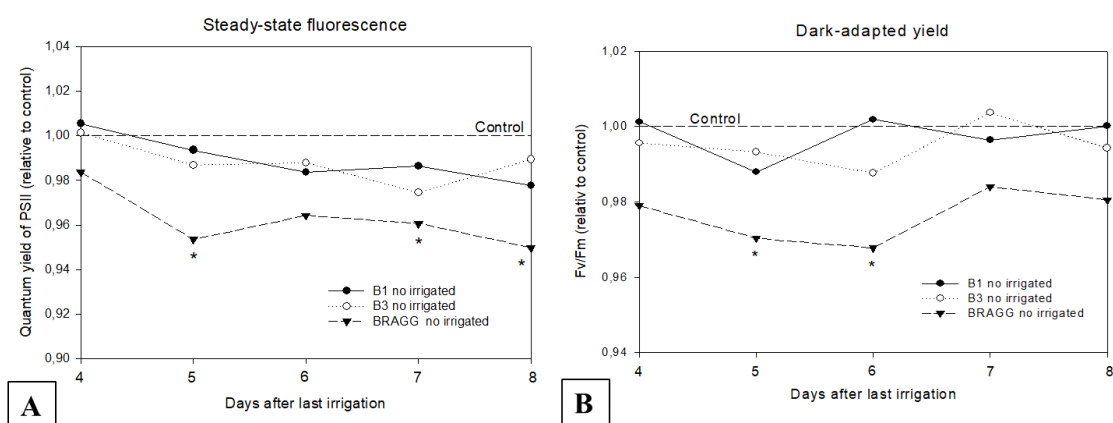


Figure 3 - Fluorescence measurement under (A) light and (B) dark. The B1 and B3 plants presented higher photosynthetic capacity when compared to control plants.

3.3 Transcriptome analyses

3.3.1 Transcriptome analyses by DESeq2

Transcriptome analyses are becoming more prevalent. RNA-seq is one of those techniques which besides having high yield has advantages such as: (i) not limiting the detection of transcripts, (ii) low background signal, (iii) quantification of high accuracy of expression levels and (iv) relatively low sample requirement (Yin et al., 2014). In soybean, several analyzes of the transcriptome in response to drought have already been reported (Chen et al., 2013; Belamkar et al., 2014; Shin et al., 2015; Rodrigues et al., 2015; Arraes et al., 2015; Tripathi et al., 2016).

The present study was carried out aiming the investigation of the drought tolerance mechanism in soybean transgenic plants that overexpress an osmotin from *S. nigrum*. The transcriptomes of transgenic and non-transgenic soybean plants were obtained from leaf pools of four plants/treatment in duplicate, totalizing 12 libraries: NTC1 and NTC2 (Non-transgenic plants in control condition); B1C3 and B1C4 (B1 transgenic plants in control condition); B3C5 and B3C6 (B3 transgenic plants in control condition); NTS7 and NTS8 (Non-transgenic plants in drought condition); B1S9 and B1S10 (B1 transgenic plants in drought condition); B3S11 and B3S12 (B3 transgenic plants in drought condition). A total of 672.635.892 millions of reads was obtained. Read counts (obtained by FastQC software and Trim Galore! Software), as well as the signature for each library are described in Table S1.

To perform the analyzes, raw gene counts were normalized using the DESeq2. The use of Principal Component Analysis (PCA), a mathematical algorithm, represents the samples graphically, making it possible to visually assess similarities and differences between samples and to determine if samples can be grouped (Ringnér, 2008). The PCA graph (Figure 4) clearly shows the formation of four distinct groups of samples. It is possible to observe that three groups are formed by similar samples (i) NTC1 and NTC2; (ii) B1C3, B1C4, B3C5, B3C6; (iii) B1S9, B1S10, B3S11 and B3S12. It was also verified that, unexpectedly, the NTS7 sample did not cluster with the NTS8 sample. In addition, the NTS7 library presented inferior quality (lower read counts) when compared to the other libraries (Table S1).

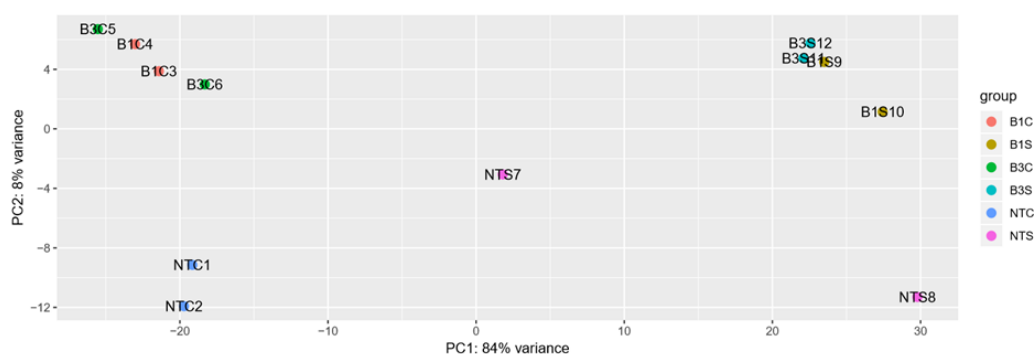


Figure 4 - Principal Component Analysis (PCA) per sample. Analysis of the NT, B1 and B3 samples in control and drought condition.

Although it is possible to use the DESeq2 software to analyze experiments without repetitions, the results often require validation of candidate genes. Therefore, the analysis was performed after removing the NTS7 library. From this analysis, 115 differentially expressed genes (DEGs) were detected when transgenic plants were compared to NT plants ($\log_2\text{FoldChange} \geq 2$ and $\text{padj} \leq 0.001$). In the comparison of B1 versus NT plants in the drought condition, 87 DEGs were detected, in which 43 up and 44 downregulated (Table S2). In the same way, comparing B3 versus NT plants under drought, 98 DEGs were detected, in which 54 up and 44 downregulated (Table S3).

Thirty six out of the 61 upregulated (Figure 5A) and 34 out of 54 downregulated (Figure 5B) genes are shared by B1 and B3 events.

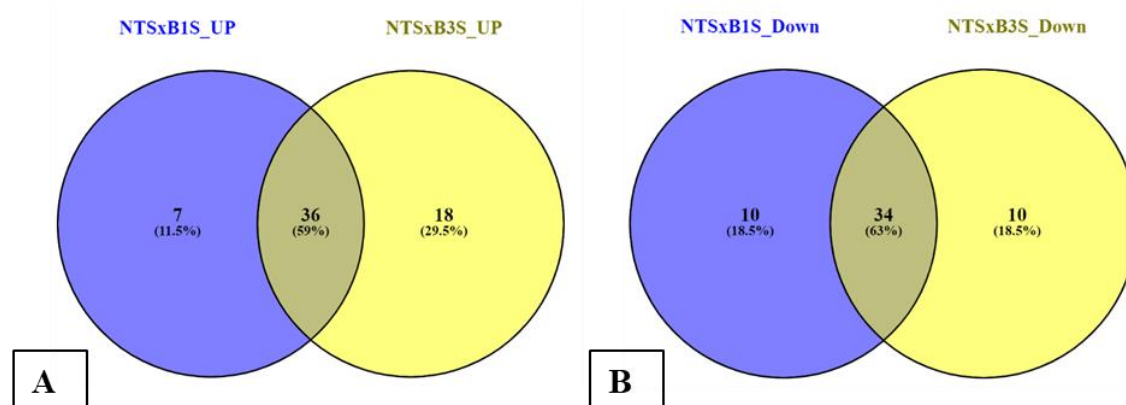


Figure 5 - Venn diagram analysis of drought-responsive genes identified using DESeq2. (A) Number of upregulated genes that are exclusive and overlapped between NTSxB1S and NTSxB3S. (B) Number of downregulated genes that are exclusive and overlapped between NTSxB1S and NTSxB3S.

The genes differentially expressed identified are annotated for photosynthesis, arginine biosynthesis and nicotinate and nicotinamide metabolism metabolic pathways.

The comparison between each transgenic event versus its control in the drought situation was also performed. A total of 2044 and 1505 DEGs were detected in B1 and B3 transgenic plants, respectively ($\log_2\text{FoldChange} \geq 2$ and $\text{padj} \leq 0.001$). Regarding B1 DEGs 769 genes were up and 1275 downregulated (Table S4). The same analysis for B3 revealed 541 up and 964 downregulated genes (Table S5). The exclusion of the DEGs in commun with non-transgenic plants resulted in 395 upregulated (Figure 6A) and 234 downregulated (Figure 6B) genes shared by B1 and B3 events. From this, 261

and 58 were upregulated in B1 and B3, respectively; 251 and 91 were downregulated in B1 and B3, respectively.

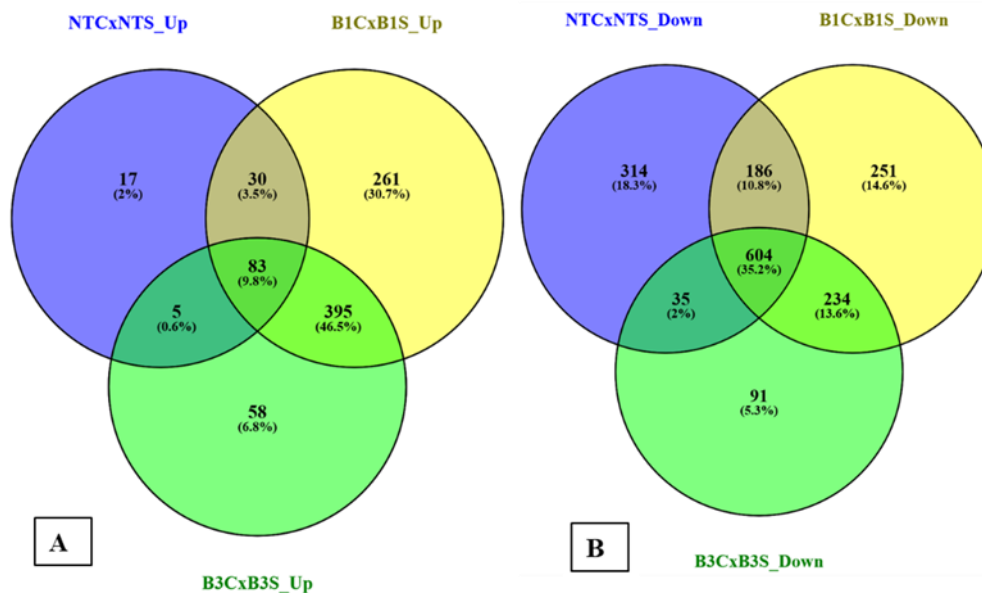


Figure 6 - Venn diagram analysis of drought-responsive genes identified using DESeq2. (A) Number of upregulated genes that are exclusive and overlapped among NTS, B1S and B3S. (B) Number of downregulated genes that are exclusive and overlapped among NTS, B1S and B3S.

Verification of which metabolic pathways differentially expressed genes have annotations is being performed.

3.3.2 Transcriptome analyses by Transcriptogramer

In the present study, the objective was not to compare DESeq2 and Transcriptogramer methods. The purpose of using both strategies was to obtain complementary data.

The Transcriptogramer strategy focuses on the analysis of large sets of genes with distinct metabolic functions (categories of gene ontologies) in order to identify the DEGs. Thus, this method presents and analyses transcription data on a wide genomic scale, which reduces noise and facilitates biological interpretation, providing a hierarchical relation of the data (Rybarczyk-Filho et al., 2011). Moreover, the usefulness and advantage of this tool is the easy identification of genome regions with

the greatest transcriptional changes allowing the projection of that altered regions to the graph of ordering and biological processes.

The set of data obtained by the DESeq2 software standardization was used to construct transcriptograms calculated using windows with radius (R) = 125 (Figure 7).

Differentially of DESeq2, the NTS7 library, which presented lower read counts in comparison to the other libraries, was included in the Transcriptogramer analysis. This was possible because the total read counts were used to normalize the reads of each gene and than the transcriptional profile of this library was similar to the other libraries (Figure 7).

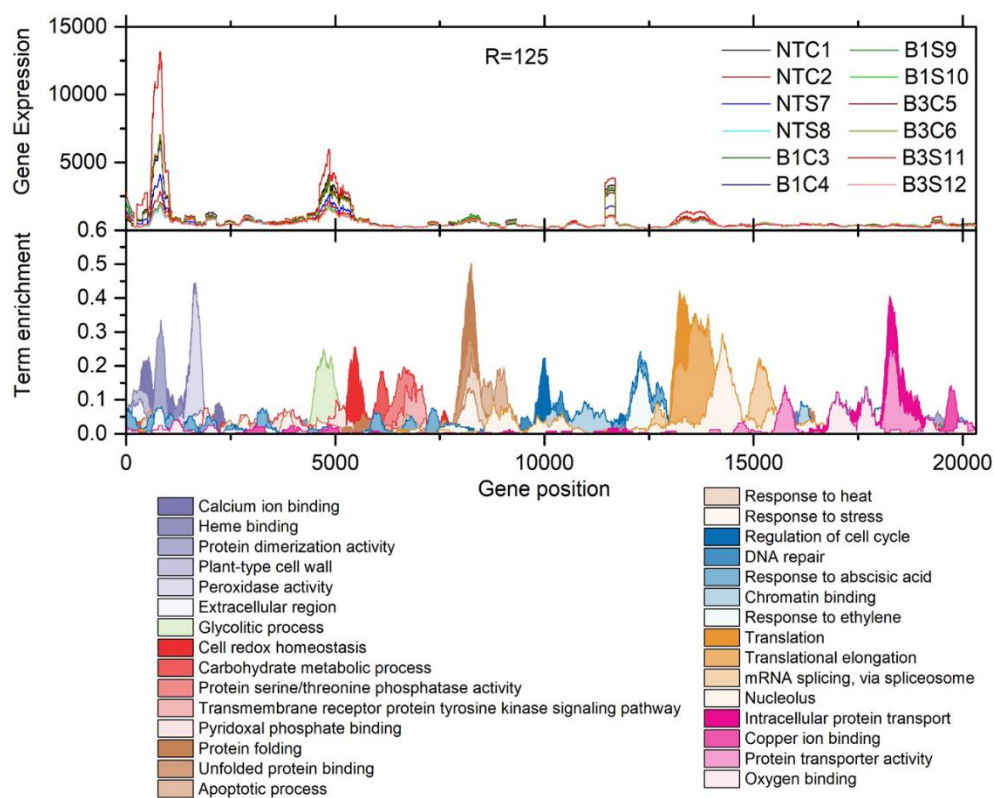


Figure 7 - Transcriptogram profile. Twelve libraries obtained by the Transcriptogramer analyses projected on the graph of ordering and biological processes.

The transcriptional profiles of NT, B1 and B3 plants submitted to water stress versus their respective control (irrigated) are represented in Figure 8. The figure contains four graphs. Three of them are the relative transcriptograms (A-C) and the fourth (D) is the density profile of GO terms. The x-axis is the same for the four graphs representing the ordered gene list. The averages of the relative transcriptogram for each gene position were calculated by dividing the mean of the sample transcriptogram value

by the value of its control sample. Therefore, the control (irrigated) is always represented by a straight horizontal line in the graphs. The standard error for the transcriptogram at each gene position was obtained and is represented as a shadow around the horizontal control line and around the mean of the relative transcriptogram value for each sample (Figure 8). These results indicate a general scenario of the biological processes involved in the soybean response to drought.

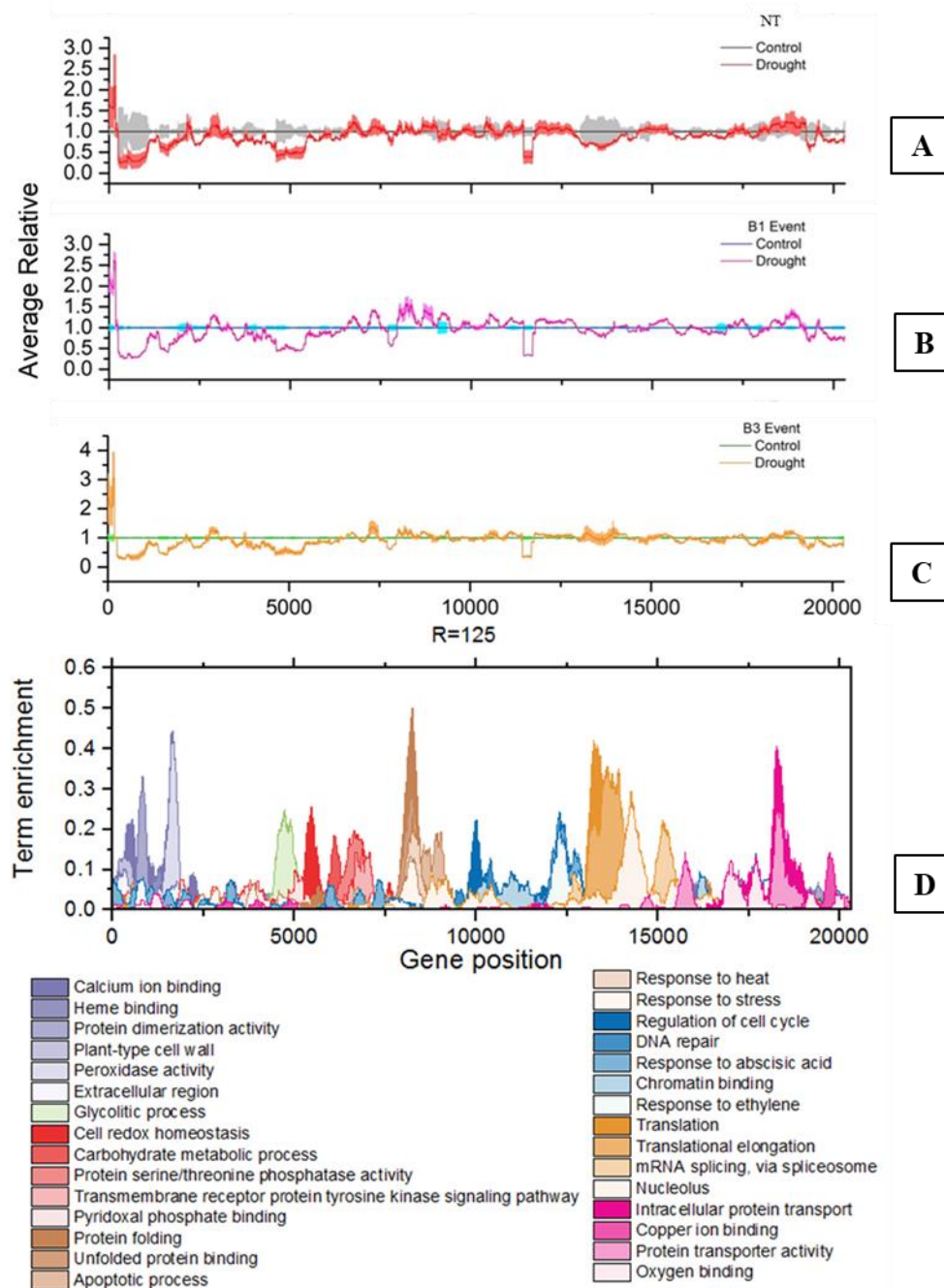


Figure 8 - Transcriptogram analyses projected on the graph of ordering and biological processes. (A) NT, (B) B1, (C) B3 transcriptograms and (D) Soybean ordered list and gene ontology categories. The expression profiles are relative to the respective irrigated control (average relative = 1). The standard error for the

transcriptogram at each gene position was obtained and is represented by the gray, blue and green shadow around the control horizontal line and by the red (a), magenta (b), and orange (c) around the mean relative transcriptogram values for each sample.

The NTS, B1S and B3S relative transcriptograms show changes at various intervals, which overlap in the different genotypes. In Figure 9, is possible to verify the highlighted regions with downregulated (salmon) and upregulated (green) genes. These result confirmed those obtained by DESeq2 analysis, showing a higher number of downregulated genes due to water stress. It is noteworthy that the different genotypes present a very similar transcriptogram pattern.

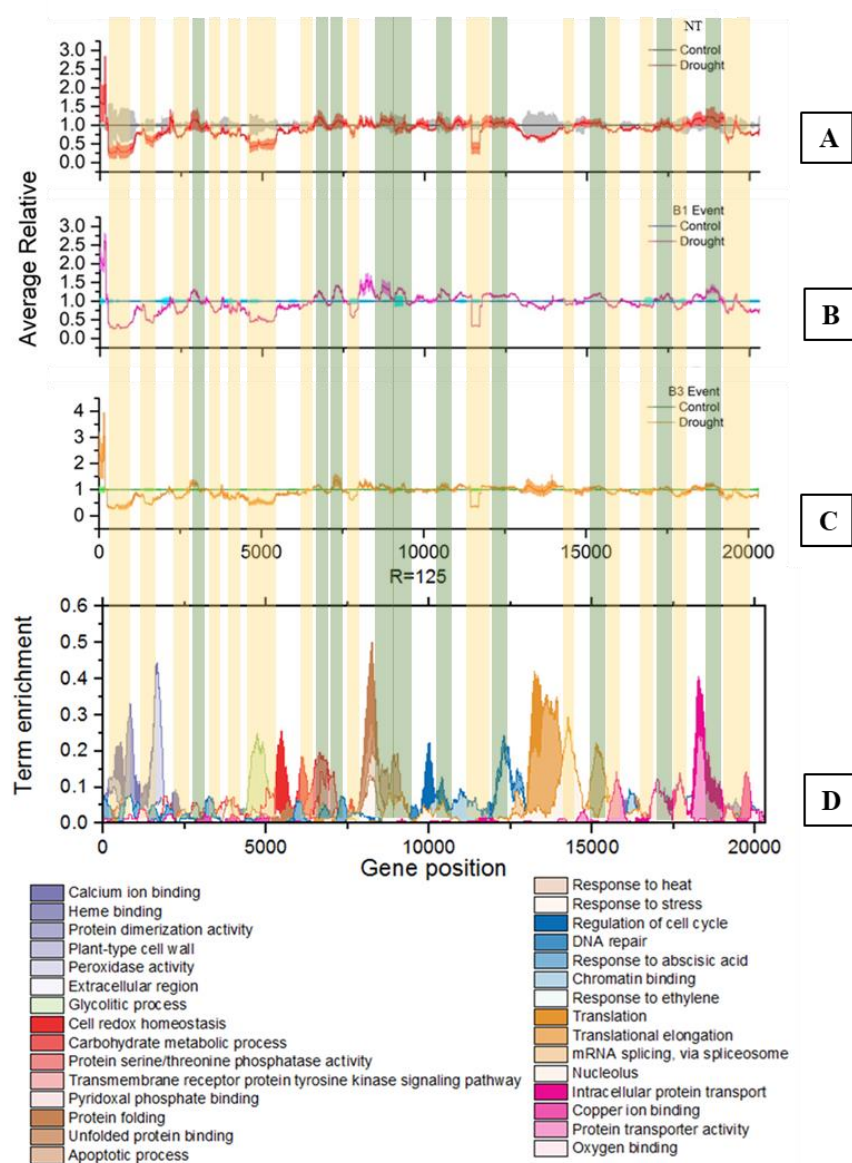


Figure 9 Average transcriptograms of each genotype relative to control. The salmon and green vertical regions highlight areas with down and up regulated genes respectively.

Fourty six GO categories were detected as differentially expressed when comparing B1S and B3S *versus* NTS. The cut-off point established for this analysis was $p \leq 0.08$. No GO categories were found as differentially expressed in B1S *versus* B3S comparison (Table S6). Only 10 out of the 46 GO categories contained differentially expressed genes: (i) establishment of chromosome localization; (ii) delta DNA polymerase complex; (iii) preribosome, large subunit precursor; (iv) oxidoreductase activity, acting on the CH-CH group of donors, NAD or NADP as acceptor; (v) Cajal body; (vi) single-stranded DNA 3'-5' exodeoxyribonuclease activity; (vii) base-excision repair, gap-filling; (viii) cholinesterase activity; (ix) diphosphomevalonate decarboxylase activity; and (x) mitotic metaphase. Some genes are present in more than one GO category. Thus, the final list comprised seven DEGs ($p \leq 0.06$): (i) Myosin heavy chain-related protein (*Glyma.14G115900*); (ii) DNA polymerase delta subunit 1 (POLD1) (*Glyma.06G155300*); (iii) Pescadillo (PES1, NOP7) (*Glyma.03G234900*); (iv) Hexadecanal dehydrogenase (acylating)/Fatty acyl-CoA reductase (*Glyma.02G169100*); (v) H/ACA ribonucleoprotein complex subunit 4 (DKC1, NOLA4, CBF5) (*Glyma.15G056100*); (vi) choline/ethanolamine kinase (CHK) (*Glyma.02G078500*); and (vii) Diphosphomevalonate decarboxylase / Mevalonate pyrophosphate decarboxylase (*Glyma.10G279800*) (Table 3).

Table 3 – Description of GOs categories and genes differentially expressed. GOs categories differentially expressed with DEGs differentially expressed when comparing B1S and B3S *versus* NTS.

GO	Description GO	<i>p</i> -value (NTSxB1S)	<i>p</i> -value (NTSxB3S)	Gene	Description Gene	<i>p</i> -value (NTSxB1S)	<i>p</i> -value (NTSxB3S)
GO0051303	Establishment of chromosome localization	0.0140624	0.0077311	<i>Glyma.14G115900</i>	PTHR23160:SF7 - Myosin heavy chain-related protein	0.00704222	0.00412314
GO0043625	Delta DNA polymerase complex	0.0649354	0.022635	<i>Glyma.06G155300</i>	K02327 - DNA polymerase delta subunit 1 (POLD1)	0.0252804	0.0165645
GO0030687	Preribosome, large subunit precursor	0.0661505	0.0476978	<i>Glyma.03G234900</i>	K14843 - Pescadillo (PES1, NOP7)	0.0661505	0.0476978
GO0016628	Oxidoreductase activity, acting on the CH-CH group of donors, NAD or NADP as acceptor	0.01966	0.048281	<i>Glyma.02G169100</i>	1.2.1.42 - Hexadecanal dehydrogenase (acylating) / Fatty acyl-CoA reductase	0.01966	0.048281
GO0015030	Cajal body	0.0475678	0.0653382	<i>Glyma.15G056100</i>	K11131 - H/ACA Ribonucleoprotein complex subunit 4 (DKC1, NOLA4, CBF5)	0.0421236	0.0692319
GO0008310	Single-stranded DNA 3'-5' exodeoxyribonuclease activity	0.0393401	0.0550234	<i>Glyma.06G155300</i>	K02327 - DNA polymerase delta subunit 1 (POLD1)	0.0252804	0.0165645
GO0006287	Base-excision repair, gap-filling	0.0567875	0.0882764	<i>Glyma.06G155300</i>	K02327 - DNA polymerase delta subunit 1 (POLD1)	0.0252804	0.0165645
GO0004104	Cholinesterase activity	0.0174741	0.00699706	<i>Glyma.02G078500</i>	K14156 - choline/ethanolamine kinase (CHK)	0.0107382	0.0101243
GO0004163	Diphosphomevalonate decarboxylase activity	0.0545798	0.0266303	<i>Glyma.10G279800</i>	4.1.1.33 - Diphosphomevalonate decarboxylase / Mevalonate pyrophosphate decarboxylase	0.0545798	0.0266303
GO0000089	Mitotic metafase	0.0661505	0.0476978	<i>Glyma.03G234900</i>	K14843 - Pescadillo (PES1, NOP7)	0.0661505	0.0476978

Taking in account the mean of reads, five out of the seven DEGs presented higher number of transcripts in the transgenic plants when compared with NT plants, under drought (Figure 10). This was the criterion used for the selection of DEGs for further analyses.

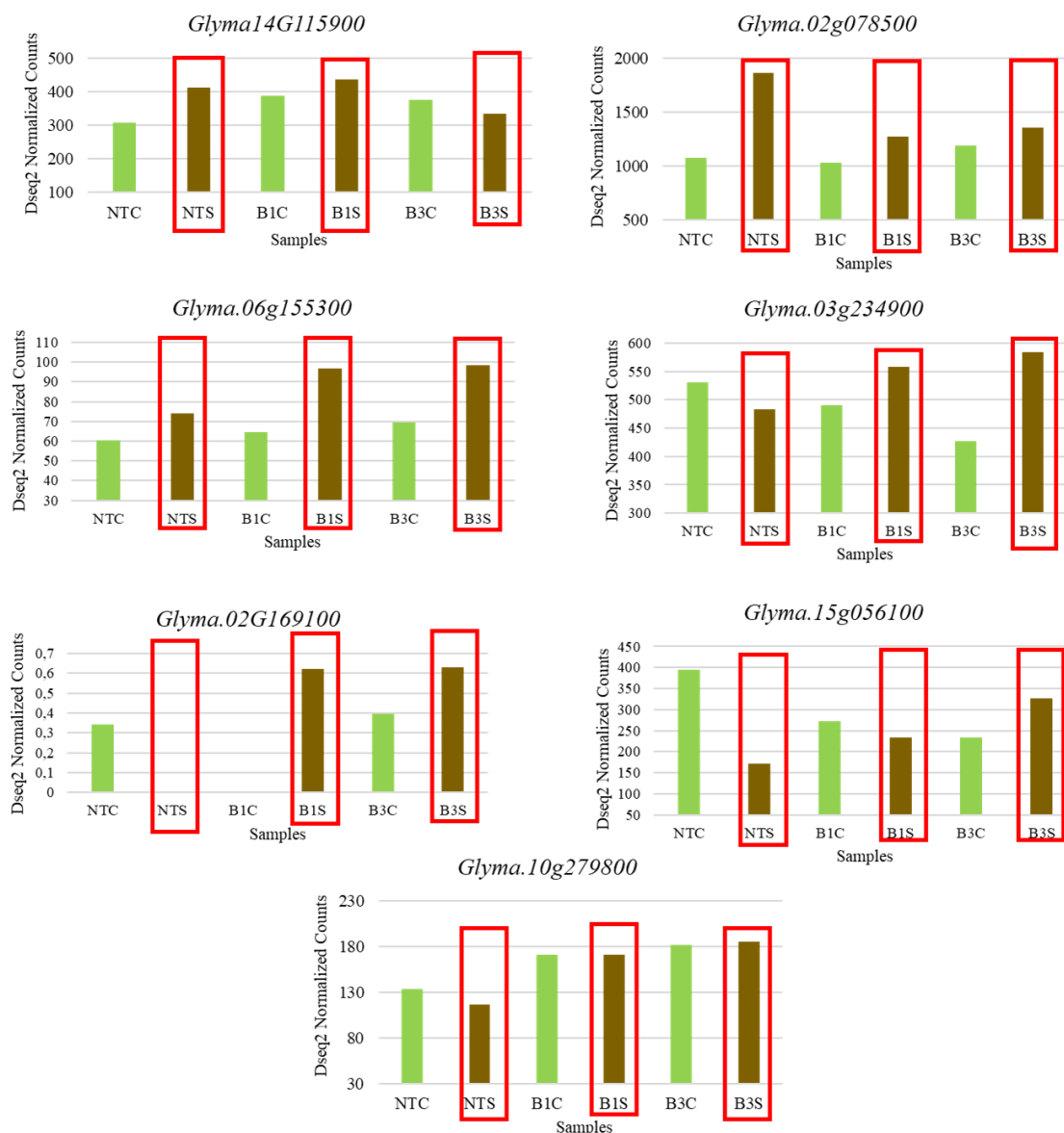


Figure 10 - Number of reads in the transgenic and NT plants. Five out of the seven DEGs presented higher number of reads in the transgenic plants when compared with NT plants, under drought. Red boxes highlight the samples analysed.

Pescadillo (PES1) and H/ACA ribonucleoprotein complex subunit 4 are involved in the ribosome biogenesis. This is a complex process with high demands in energy supply

and resources, co-ordinated with cell growth and proliferation. It is still unknown how the variable metabolic and environmental situations regulate the plant ribosome biogenesis, which is essential for plant growth and environmental acclimation. A screening performed in rice plants grown under limited water conditions revealed the activation of Ribosomal Protein Large (RPL) subunit genes in two mutants that exhibited high water-use efficiency (Moin et al., 2016). The Pescadillo (PES) protein controls the 60S ribosomal subunit biogenesis in plants (Ahn et al., 2016) and impacts cell cycle and ribosome biogenesis. It is a conserved nucleolar protein that is indispensable for the viability of higher eukaryotes and yeast (Lerch-Gaggl et al., 2002). Plant cells deficient in PES demonstrated delayed maturation of 25S ribosomal RNA and suppressed global translation (Cho et al., 2013). Thus, PES has a crucial role in cell growth and survival through regulation of mitotic progression and ribosome biogenesis.

H/ACA ribonucleoprotein subunits are part of a complex, which catalyses the pseudouridylation of rRNA. The nucleus possesses a great number of small nuclear RNAs (snRNAs), which present various cellular functions (Yu et al., 1999). Likewise, the RNA processing, the snRNAs are also involved in the regulation of transcription elongation and in the synthesis of telomeric DNA repeats (Blackburn E., 1999; Yang et al., 2001). The snRNAs associate with other specific proteins and form small nuclear or nucleolar ribonucleoprotein particles (snRNPs or snoRNPs) (Yu et al., 1999). snRNPs predominantly function as core constituents of the spliceosome, the molecular machinery for pre-mRNA splicing. Thus, snRNP biogenesis is a crucial step for plants, indispensable for the determination of a cell's activity through the regulation of gene expression (Ohtani, M. 2017). It is already known that plants can obtain tolerance to environmental stresses via transcriptome reprogramming at transcriptional and alternative splicing (AS) levels (Reddy et al., 2013). As an example, the DREB2B, an abiotic stress regulator was modulated by AS in rice (Matsukura et al., 2010). The data obtained using the large-scale sequencing platforms allowed a genome-wide profiling of AS events in plant species and revealed that ~33%–60% mRNA are alternatively spliced (reviewed by Liu et al., 2018).

Replication of DNA is an essential property of all living organisms. In multicellular organisms, the integration of replicative activity with proofreading functions is increasingly valued as fundamental for normal development. POLD1, the catalytic subunit of the replicative DNA polymerase delta (δ) includes two domains, the DNA

polymerase and 3' to 5' proofreading exonuclease (Nicolas et al., 2016). The replicative polymerases (DNA Pol α , δ , and ϵ) are associated to a large protein complex named “replisome” that covers all the core activities required for DNA replication (Kurth; O'Donnell, 2013). Besides lagging strand synthesis, Pol δ also participates in several other processes that repair DNA lesions necessary to protect the genome integrity. In plants, the role of the Pol δ in epigenetic inheritance has not been understood.

Diphosphomevalonate decarboxylase is an important enzyme of the mevalonic acid (MVA) pathway (Krepkiy and Miziorko, 2004). It is involved in the synthesis of isopentenyl diphosphate (IPP), required for the biosynthesis of isoprenoids, an important cellular intermediate (Bergès et al., 1997). In plants, the isoprenoids function in the protection from and adaptation to stresses. Drought stress inhibits photosynthesis, improves the production of reactive oxygen species (ROS), and consequently promotes photooxidative stress (Hernández et al., 2012). Isoprenoids and phenylpropanoids, may have the potential to complement the functional roles of antioxidant enzymes under persistent patterns of high temperature and light, and drought stress. It was already shown that under high-light and thermic stress the quantity of carotenoids (isoprenoid) is increased, as well as, the composition changes (Singsaas et al., 1997) enhancing tolerance. It was previously reported also that carotenoids play an important role in drought-stressed plants suffering severe inhibition of photosynthesis (Wujeska et al., 2013).

Hexadecanal dehydrogenase (acylating) or Fatty acyl-CoA reductase (*Glyma.02G169100*) is an enzyme that catalyzes the chemical reaction for the formation of hexadecanoyl-CoA (Palmitoyl-CoA), a long-chain fatty acyl-CoA. The reduction of fatty acid proceeds through the ATP-dependent formation of acyl intermediates such as acyl-CoA and acyl protein, followed by reduction to aldehyde and alcohol with NAD(P)H. Fatty alcohols and aldehydes play important roles in the metabolic pathways since they can generate fatty acids along with formation of the reduced forms of enzyme cofactors. The germinating seeds contain high levels of esterified fatty alcohols which serve as an essential source of energy for plant growth (Riendeau; Meighen, 1985).

Since fatty acyl-CoA reductase functions at a branch point in fatty acid metabolism, this enzyme may have important control functions. For example, the level of this enzyme may control the channeling of fatty acids into alkyl glyceryl ethers and wax esters which tend to accumulate in certain tumors and other tissues (Kolattukudy, 1971).

The *A. thaliana CER4* encodes an alcohol-forming fatty acyl-coenzymeA reductase involved in cuticular wax production. Waxes and cutin are the main components of the cuticle, which is synthesized by plant epidermal cells. The main function of the cuticle is to serve as a waterproof barrier, restricting nonstomatal water loss (Rowland et al., 2006).

Considering that 6875 GO categories are available for soybean at AmiGO 2 (<http://amigo.geneontology.org/amigo>) a screening was performed looking for DEGs in the other 6829 GO categories not differentially expressed. In this analysis 121 DEGs were identified ($p \leq 0.05$). Five out of the 121 DEGs were chosen for further analyses based on the criterion of lower p -value: (i) Haem-binding uptake, Tiki superfamily, ChaN (*Glyma.11G233800*); (ii) No description (*Glyma.14G188500*); (iii) Myosin heavy chain-related protein (*Glyma.14G115900*); (iv) Domain of unknown function (DUF966) (*Glyma.11G193000*); and (v) Kinesin-like protein KIN12A-related (*Glyma.08G106400*) (Table 4).

Table 4 – DEGs in the other 6829 GOs categories not differentially expressed when comparing B1S and B3S versus NTS.

Gene	Description Gene	p -value (NTSxB1S)	p -value (NTSxB3S)
<i>Glyma.11G233800</i>	PF04187 - Haem-binding uptake, Tiki superfamily, ChaN (Cofac_haem_bdg)	0.00476016	0.00662356
<i>Glyma.14G188500</i>	No description	0.00644416	0.0164108
<i>Glyma.14G115900</i>	PTHR23160:SF7 - Myosin heavy chain-related protein	0.00704222	0.00412314
<i>Glyma.11G193000</i>	PF06136 - Domain of unknown function (DUF966)	0.0415648	0.0209327
<i>Glyma.08G106400</i>	PTHR24115:SF413 - Kinesin-like protein KIN12A-related	0.0415648	0.0209327

Comparing the list of DEGs obtained from DESeq2 and The Transcriptogramer, only one gene, *Glyma.14G065100*, was common. This gene encodes a WD repeat and SOF domain-containing protein 1 (WDSOF1). It possesses orthologous genes in *A. thaliana* (*AT4G28450*) and rice (*Os01g13730*) with 87.8% similarity. Lee et al. (2008) reported that both genes have Sof1 domains that are a component of the nucleolar RNA processing machinery and are therefore essential for cell growth.

Since both methodologies used the same input data, it was expected to find larger number of differentially expressed genes in common. However, to identify other genes

in common, it would be necessary to use a larger p -value, which could result in weak evidence and lose the results reliability. Anyway our goal was to use The Transcriptogramer as a complementary tool to DeSeq2.

4. Conclusions

In the present study we confirmed by physiological analyses that the overexpression of SnOLP confers tolerance to drought. The responses to stress are complex and focused on hormonal, photosynthetic, respiration and transcriptional regulation (Chen et al., 2016). The transcriptome analyses carried out in the soybean transgenic plants submitted to water deficit provide an opportunity to investigate the molecular mechanism of drought tolerance.

The RNA sequencing analyses (RNA-Seq) often provide a high number of differentially expressed genes among the samples (Conesa et al., 2016). However, in our transcriptome analyses relatively few DEGs were identified. This can be accounted to the comparison performed between transgenic and NT plants under stress, which excluded the genes involved in drought response independent on the SnOLP presence. In addition, due to the lack of DNA binding motifs, osmotins do not act as transcription factors (Abdin et al., 2011). The GOs categories and genes differentially expressed identified are mainly involved in biological processes related to cell cycle regulation and maintenance of internal cell environment. DNA replication, repair and methylation; ribosome biosynthesis and secondary metabolic synthesis were also modulated. Validation of gene expression patterns and the identification of metabolic pathways in which the DEGs may interact are necessary for the elucidation of the osmotin role in drought tolerance.

References:

Abdin, M.Z., Kiran, U., Alam, A. (2011). Analysis of osmotin, a PR protein as metabolic modulator in plants. *Bioinformation*, 5, 336–340.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 25 (17), 3389-3402.

- Ahn, C. S., Cho, H. K., Lee, D. H., Sim, H. J., Kim, S. G., & Pai, H. S. (2016). Functional characterization of the ribosome biogenesis factors PES, BOP1, and WDR12 (PeBoW), and mechanisms of defective cell growth and proliferation caused by PeBoW deficiency in Arabidopsis. *Journal of experimental botany*, 67 (17), 5217-5232.
- Arraes, F. B. M., Beneventi, A., De Sa, M. E. L., Paixao, J. F. R., Albuquerque, E. V. S., Marin, S. R. R., Purgatto, E., Nepomuceno, A. L., GrossI-De-Sa, M. F. (2015). Implications of ethylene biosynthesis and signaling in soybean drought stress tolerance. *BMC Plant Biology*, 15 (1), 213.
- Barthakur, S., Babu, V., Bansal, K. C. (2001). Overexpression of osmotin induces proline accumulation and confers drought- and low-temperature stress tolerance in tobacco by gene transfer. *Plant Cell Physiol*, 45, 346–350.
- Baker, N. R. (2008). Chlorophyll fluorescence: a probe of photosynthesis in vivo. *Annu. Rev. Plant Biol*, 59, 89-113.
- Belamkar, V., Weeks, N. T., Bharti, A. K., Farmer, A. D., Graham, M. A., Cannon, S. B. (2014). Comprehensive characterization and RNA-Seq profiling of the HD-Zip transcription factor family in soybean (*Glycine max*) during dehydration and salt stress. *BMC Genomics*, 15 (1), 950.
- Bergès, T., Guyonnet, D., and Karst, F. (1997). The *Saccharomyces cerevisiae* mevalonate diphosphate decarboxylase is essential for viability, and a single Leu-to-Pro mutation in a conserved sequence leads to thermosensitivity. *Journal of bacteriology*, 179 (15), 4664-4670.
- Blackburn E. (1999). Telomerase. In Gesteland, R.F, Cech, T. R., Atkins, J. F. (eds.), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp. 609–635.
- Chen, L. M., Zhou, X. A., Li, W. B., Chang, W., Zhou, R., Wang, C., Sha, A. H., Shan, Z. H., Zhang, C. J., Qiu, D. Z., Yang, Z. L., Chen, A. L. (2013). Genome-wide transcriptional analysis of two soybean genotypes under dehydration and rehydration conditions. *BMC Genomics*, 14 (1), 687.
- Chen, W., Yao, Q., Patil, G. B., Agarwal, G., Deshmukh, R. K., Lin, L., Wang, B., Wang, Y., Prince, S. J., Song, L., Xu, D., An, Y. C., Valliyodan, B., Varshney, R. K., Nguyen, H. T. (2016). 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*, 7, 1044.
- Cho, H. K., Ahn C. S., Lee, H. S., Kim, J. K., Pai, H. S. (2013) Pescadillo plays an essential role in plant cell growth and survival by modulating ribosome biogenesis. *The Plant Journal*, 76 (3), 393-405.
- Conesa, A., Madrigal, P., Tarazona, S., Gomez-cabrero, D., Cervera, A., Mcpherson, A., Szczesniak, M. W., Gaffney, D. J., Elo, L. L., Zhang, X., Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, 17 (1), 13.

CONAB - Companhia Nacional de Abastecimento, Acompanhamento da Safra Brasileira de Grãos, Safra 2016/17, disponível em: <www.conab.gov.br>

Darzi, Y., Letunic, I., Bork, P., Yamada, T. (2018). iPath3. 0: interactive pathways explorer v3. *Nucleic acids research*, 46 (W1), W510-W513.

Das M., Chauhan, H., Chhibbar, A., Rizwanul Haq, Q. M., Khurana, P. (2011). High-efficiency transformation and selective tolerance against biotic and abiotic stress in mulberry, *Morus indica* cv. K2, by constitutive and inducible expression of tobacco osmotin. *Transgenic research*, 20 (2), 231-246.

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

Fàbregas, N., Fernie, A. R. (2019). The metabolic response to drought. *Journal of experimental botany*, 70 (4), 1077-1085.

Faillace, G. R., Turchetto-Zolet, A. C., Guzman, F. L., de Oliveira-Busatto, L. A., & Bodanese-Zanettini, M. H. (2019). Genome-wide analysis and evolution of plant thaumatin-like proteins: a focus on the origin and diversification of osmotins. *Molecular Genetics and Genomics*, 1-21.

Goel, D., Singh, A. K., Yadav, V., Babbar, S. B., Bansal, K. C. (2010). Overexpression of osmotin gene confers tolerance to salt and drought stresses in transgenic tomato (*Solanum lycopersicum* L.). *Protoplasma*, 245, 133-141.

Guedes, R. L. M., Prosdocimi, F., Fernandes, G. R., Moura, L. K., Ribeiro, H. A. L., Ortega, J. M. (2011). Amino acids biosynthesis and nitrogen assimilation pathways: a great genomic deletion during eukaryotes evolution. *BMC genomics*, 12 (4), 1.

Hakim, A. U., Amjad, H., Muhammad, S., Aamir, H. K., Muna, A., Summia, G., Zhang, J., Sun, L., Jianying, Li., Shuangxia, Jin., Muhammad, F. H. M. (2018). Osmotin: A plant defense tool against biotic and abiotic stresses. *Plant Physiology and Biochemistry*, 123, 149-159.

Hernández, I., Cela, J., Alegre, L., Munné-Bosch, S. (2012). Antioxidant defenses against drought stress. In: Aroca R, ed. *Plant responses to drought stress*. Berlin, Germany: Springer, 231–258.

Kurth, I., O'Donnell, M. (2013). New insights into replisome fluidity during chromosome replication. *Trends in biochemical sciences*, 38 (4), 195-203.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology*, 14 (4), 1.

Kolattukudy, P. E. (1971). Enzymatic synthesis of fatty alcohols in *Brassica oleracea*. *Archives of biochemistry and biophysics*, 142 (2), 701-709.

- Koressaar, T., Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics*, 23 (10), 1289-1291
- Kumar, S. A., Kumari, P. H., Jawahar, G., Prashanth, S., Suravajhala, P., Katam, R., Sivan, P., Rao, K. S., Kirti, P. B., Kishor, P. K. (2016). Beyond just being foot soldiers—osmotin like protein (OLP) and chitinase (Chi11) genes act as sentinels to confront salt, drought, and fungal stress tolerance in tomato. *Environmental and experimental botany*, 132, 53-65.
- Krepkiy, D., and Miziorko, H. M. (2004). Identification of active site residues in mevalonate diphosphate decarboxylase: implications for a family of phospho transferases. *Protein Science*, 13 (7), 1875-1881.
- Le, D. T., Nishiyama, R., Watanabe, Y., Tanaka, M., Seki, M., Yamaguchi-Shinozaki, K., Tran, L. S. P. (2012). Differential gene expression in soybean leaf tissues at late developmental stages under drought stress revealed by genome-wide transcriptome analysis. *PLoS one*, 7 (11), e49522.
- Lee, J. H., Terzaghi, W., Gusmaroli, G., Charron, J. B. F., Yoon, H. J., Chen, H., He, Y. J., Xiong, Y., Deng, X. W. (2008). Characterization of Arabidopsis and rice DWD proteins and their roles as substrate receptors for CUL4-RING E3 ubiquitin ligases. *The Plant Cell*, 20 (1), 152-167.
- Lerch-Gaggl, A., Haque, J., Li, J., Ning, G., Traktman, P., Duncan, S. A. (2002). Pescadillo is essential for nucleolar assembly, ribosome biogenesis, and mammalian cell proliferation. *Journal of Biological Chemistry*, 277, 45347–45355.
- Liao, Y., Smyth, G. K., Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic acids research*, 41 (10), e108.
- Livak, K. J., Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ Ct method. *Methods*, 25 (4), 402-408.
- Liu, Z., Qin, J., Tian, X., Xu, S., Wang, Y., Li, H., Wang, X., Peng, H., Yao, Y., Hu, Z., Ni, Z., Xin, M., Sun, Q. (2018). Global profiling of alternative splicing landscape responsive to drought, heat and their combination in wheat (*Triticum aestivum* L.). *Plant Biotechnology Journal*, 16 (3), 714-726.
- Love, M. I., Huber, W., Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, 15 (12), 550.
- Mathobo, R., Marais, D., Steyn, J. M. (2017). The effect of drought stress on yield, leaf gaseous exchange and chlorophyll fluorescence of dry beans (*Phaseolus vulgaris* L.). *Agricultural Water Management*, 180, 118-125.
- Matsukura, S., Mizoi, J., Yoshida, T., Todaka, D., Ito, Y., Maruyama, K., Shinozaki, K. et al. (2010). Comprehensive analysis of rice DREB2-type genes that encode transcription factors involved in the expression of abiotic stress-responsive genes. *Molecular Genetics and Genomics*, 283(2), 185-196.

- Moin, M., Bakshi, A., Saha, A., Udaya, Kumar, M., Reddy, A. R., Rao, K. V., Siddiq, E. A., Kirti, P. B. (2016). Activation tagging in indica rice identifies ribosomal proteins as potential targets for manipulation of water-use efficiency and abiotic stress tolerance in plants. *Plant Cell Environ*, 39 (11), 2440-2459.
- Nicolas, E., Golemis, E. A., Arora, S. (2016). POLD1: Central mediator of DNA replication and repair, and implication in cancer and other pathologies. *Gene*, 590 (1), 128-141.
- Ohtani, M. (2017). Transcriptional regulation of snRNAs and its significance for plant development. *Journal of plant research*, 130 (1), 57-66.
- Parkhi, V., Kumar, V., Sunilkumar, G., Campbell, L. M., Singh, N. K., Rathore, K. S. (2009). Expression of apoplastically secreted tobacco osmotin in cotton confers drought tolerance. *Molecular breeding*, 23 (4), 625-639.
- Reddy, A. S., Marquez, Y., Kalyna, M., Barta, A. (2013). Complexity of the alternative splicing landscape in plants. *Plant Cell*, 25, 3657–3683.
- Riendeau, D., Meighen, E. (1985). Enzymatic reduction of fatty acids and acyl-CoAs to long chain aldehydes and alcohols. *Experientia*, 41 (6), 707-713.
- Ringnér, M. (2008). What is principal component analysis?. *Nature biotechnology*, 26 (3), 303.
- Rodrigues, F. A., Fuganti-Pagliarini, R., Marcolino-Gomes, J., Nakayama, T. J., Molinari, H. B. C., Lobo, F. P., Harmon, F. G., Nepomuceno, A. L. (2015). Daytime soybean transcriptome fluctuations during water deficit stress. *BMC Genomics*, 16 (1), 505.
- Rowland, O., Zheng, H., Hepworth, S. R., Lam, P., Jetter, R., Kunst, L. (2006). *CER4* encodes an alcohol-forming fatty acyl-coenzyme A reductase involved in cuticular wax production in Arabidopsis. *Plant Physiology*, 142 (3), 866-877.
- Rybarczyk-Filho, J. L., Castro, M. A. A., Dalmolin, R. J. S., Moreira, J. C. F., Brunnet, L. G., De Almeida, R. M. C. (2011). Towards a genome-wide transcriptogram: the *Saccharomyces cerevisiae* case. *Nucleic acids research*, 39 (8), 3005-3016.
- Shin, J. H., Vaughn, J. N., Abdel-Haleem, H., Chavarro, C., Abernathy, B., Kim, K. D., Jackson, S. A., LI, Z. (2015). Transcriptomic changes due to water deficit define a general soybean response and accession-specific pathways for drought avoidance. *BMC Plant Biology*, 15 (1), 26.
- Singsaas, E. L., Lerdau, M., Winter, K., Sharkey, T. D. (1997). Isoprene Increases Thermotolerance of Isoprene-Emitting Species. *Plant Physiology*, 115 (4), 1413-1420.
- Tripathi, P.; Rabara, R. C.; Reese, R. N.; Miller, M. A.; Rohila, J. S.; Subramanian, S.; Qingxi, J. S.; Morandi, D.; Bücking, H.; Shulaev, V.; Rushton, P. J. (2016). A toolbox of genes, proteins, metabolites and promoters for improving drought tolerance in

soybean includes the metabolite coumestrol and stomatal development genes. *BMC Genomics*, 17 (1), 102.

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

Virginia, S., Pagan, M., Cooper, M., Kantartzi, S., K., Lightfoot, D., A., Meksem, K., Kassem, M., A. (2012). Genetic analysis of relative water content (RWC) in two recombinant inbred line populations of soybean [*Glycine max* (L.) merr.]. *Journal of Plant Genome Sciences*, 1 (2), 46–53.

Wang, L., Wang, S., Li, W. (2012). RSeQC: quality control of RNA-seq experiments. *Bioinformatics*, 28 (16), 2184-2185.

Weber, R. L. M., Wiebke-Strohm, B., Bredemeier, C., Margis-Pinheiro, M., Brito, G. G., Rechenmacher, C., Bertagnolli, P. F., Sá, M. E. L., Campos, M. A., Amorim, R. M. S., Beneventi, M. A. Margis, R., Grossi-De-Se, M. F., Bodanese-Zanettini, M. H. (2014). Expression of an osmotinlike protein from *Solanum nigrum* confers drought tolerance in transgenic soybean. *BMC Plant Biology*, 14, 343-352.

Wujeska, A., Bossinger, G., Tausz, M. (2013). Responses of foliar antioxidative and photoprotective defence systems of trees to drought: a meta-analysis. *Tree Physiology*, 33, 1018–1029.

Yin, G., Xu, H., Liu, J., Gao, C., Sun, J., Yan, Y., Hu, Y. (2014). Screening and identification of soybean seed-specific genes by using integrated bioinformatics of digital differential display, microarray, and RNA-seq data. *Gene*, 546 (2): 177-186.

Yu Y.T., Scharl, E. C., Smith, C. M., Steitz, J. A. (1999). The growing world of small nuclear ribonucleoproteins. In Gesteland, R. F., Cech, T. R., Atkins, J. F. (eds), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp. 487–524.

Yang, Z., Zhu, Q., Luo, K., Zhou, Q. (2001). The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature*, 414 (6861), 317-22.

Zhang, H., Song, B. (2017). RNA-seq data comparisons of wild soybean genotypes in response to soybean cyst nematode (*Heterodera glycines*). *Genomics Data*, 14, 36-39.

Zhang, J., Liu, J., Yang, C., Du, S., Yang, W. (2016). Photosynthetic performance of soybean plants to water deficit under high and low light intensity. *South African journal of botany*, 105, 279-287.

Supplementary material:

Table S1 - Description of the 12 libraries used in this study.

Table S2 - Heat map of differentially expressed genes between NTS and B1S based on DESeq2 results. Scale according to the logfoldChange value.

Table S3 - Heat map of differentially expressed genes between NTS and B3S based on DESeq2 results. Scale according to the logfoldChange value.

Table S4 - Heat map of differentially expressed genes between B1C and B1S based on DESeq2 results. Scale according to the logfoldChange value.

Table S5 - Heat map of differentially expressed genes between B3C and B3S based on DESeq2 results. Scale according to the logfoldChange value.

Table S6 - Fourty six GO categories differentially expressed when comparing B1S and B3S *versus* NTS. Only 10 out of the 46 GO categories ($p \leq 0.08$) contained differentially expressed genes (DEGs).

Table S1 - Description of the 12 libraries used in this study.

<i>Sample Identification</i>		<i>Read counts (FastQC)</i>	<i>Read counts (Trim Galore!)</i>	<i>Assigned</i>
<i>NTC1</i>	Non-transgenic plant in the control condition (Library 1)	80722088	69259732	30113630
<i>NTC2</i>	Non-transgenic plant in the control condition (Library 2)	76407818	56933068	24017664
<i>B1C3</i>	B1 event in the control condition (Library 3)	65469938	47940462	20978862
<i>B1C4</i>	B1 event in the control condition (Library 4)	107616144	79818172	34472157
<i>B3C5</i>	B3 event in the control condition (Library 5)	43879904	35198506	15228903
<i>B3C6</i>	B3 event in the control condition (Library 6)	73469384	57931092	25139831
<i>NTS7</i>	Non-transgenic plant in the drought condition (Library 7)	1390818	1071698	473239
<i>NTS8</i>	Non-transgenic plant in the drought condition (Library 8)	98156046	85559850	38483488
<i>B1S9</i>	B1 event in the drought condition (Library 9)	77439700	61048026	27454608
<i>B1S10</i>	B1 event in the drought condition (Library 10)	73206406	58345506	26234488
<i>B3S11</i>	B3 event in the drought condition (Library 11)	71367168	58175620	26084508
<i>B3S12</i>	B3 event in the drought condition (Library 12)	75830714	61354160	27469924

Table S2 - Heat map of differentially expressed genes between NTS and B1S based on DESeq2 results. Scale according to the logfoldChange value.

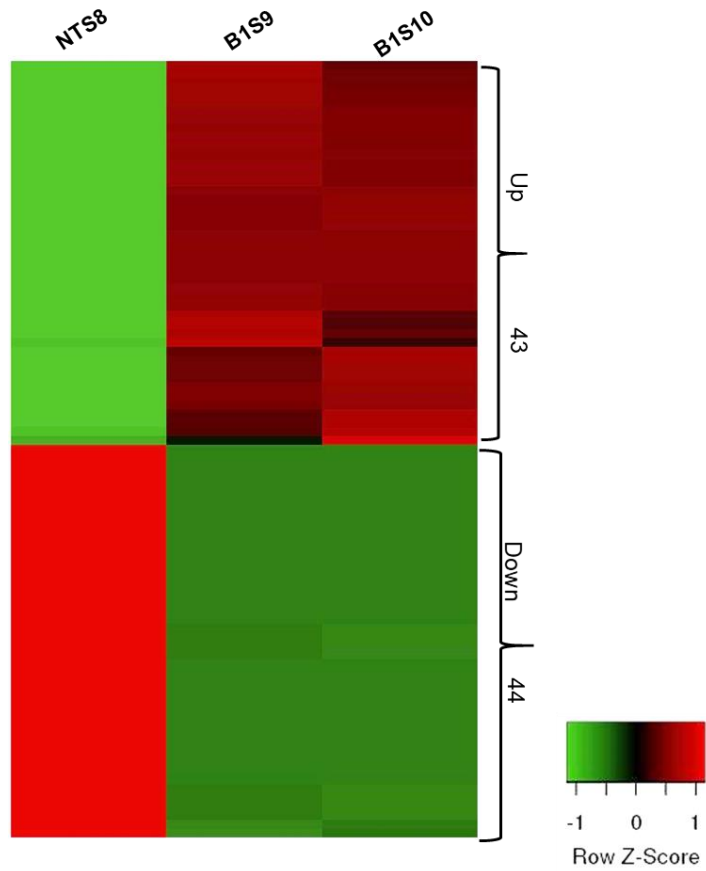


Table S3 - Heat map of differentially expressed genes between NTS and B3S based on DESeq2 results. Scale according to the logfoldChange value.

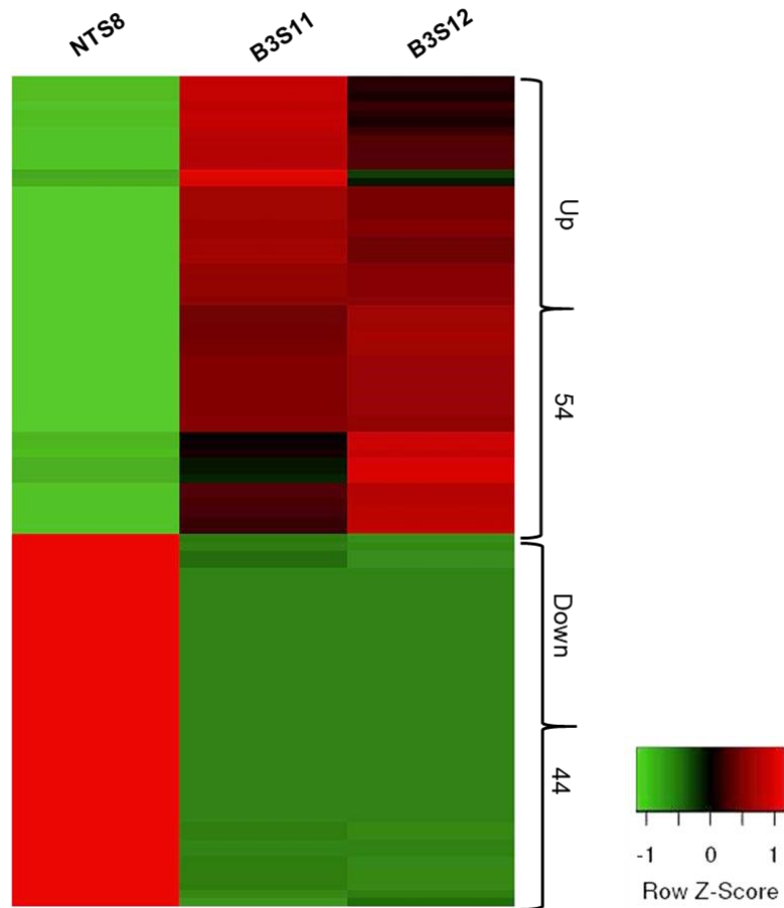


Table S4 - Heat map of differentially expressed genes between B1C and B1S based on DESeq2 results. Scale according to the logfoldChange value.

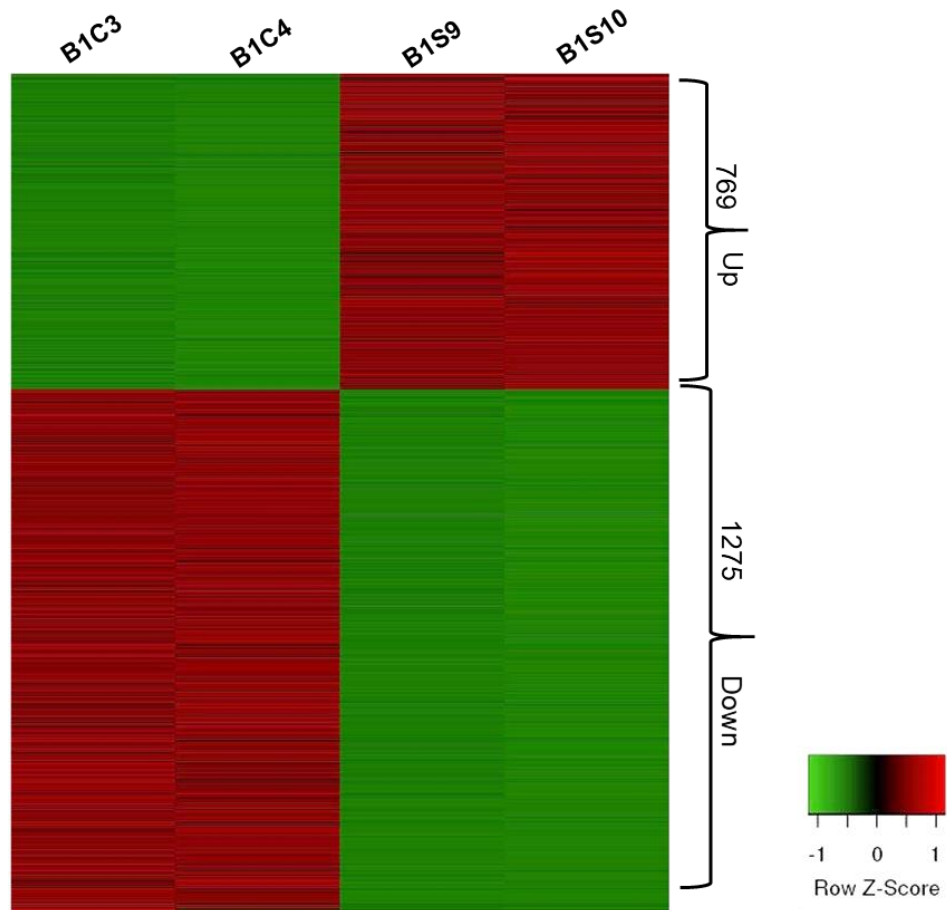


Table S5 - Heat map of differentially expressed genes between B3C and B3S based on DESeq2 results. Scale according to the logfoldChange value.

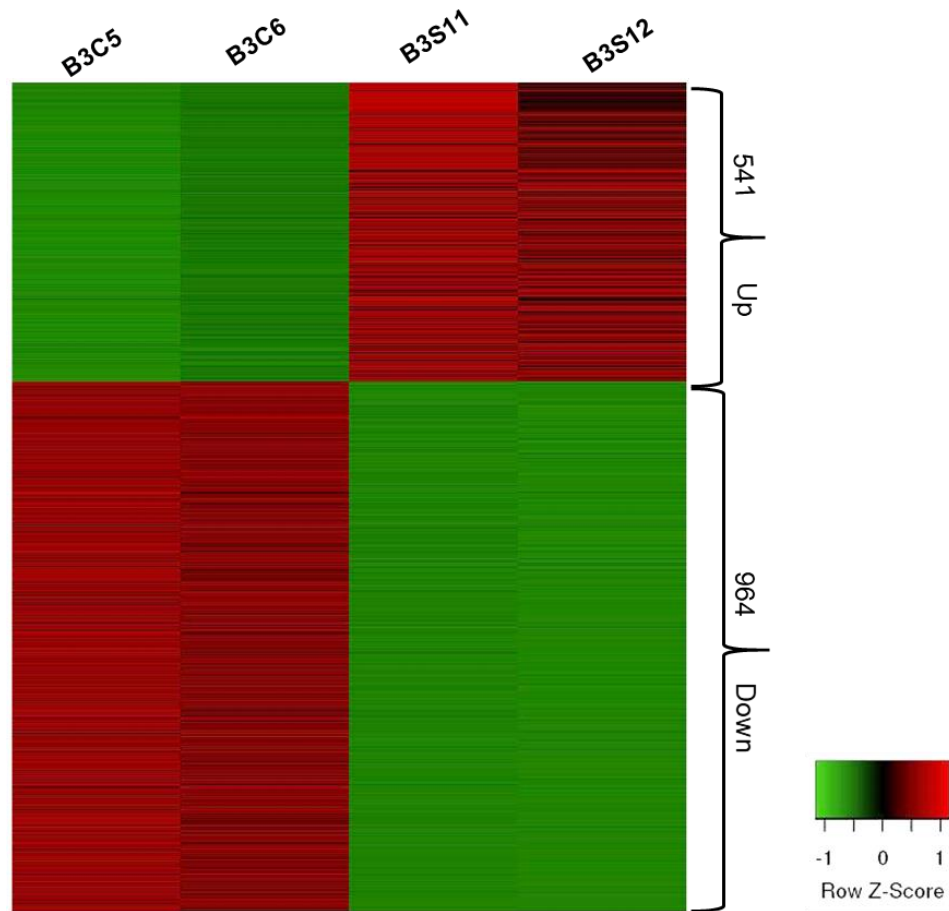


Table S6 - Fourty six GO categories differentially expressed when comparing B1S and B3S versus NTS. Only 10 out of the 46 GO categories ($p \leq 0.08$) contained differentially expressed genes (DEGs).

GO	Description	<i>p</i> -value (NTSxB1S)	<i>p</i> -value (NTSxB3S)	DEGs
GO0051563	Smooth endoplasmic reticulum calcium ion homeostasis	0,0113929	0,0313269	
GO0051966	Regulation of synaptic transmission, glutamatergic	0,0113929	0,0313269	
GO0060075	Regulation of resting membrane potential	0,0113929	0,0313269	
GO0050673	Epithelial cell proliferation	0,0113929	0,0313269	
GO0050771	Negative regulation of axonogenesis	0,0113929	0,0313269	
GO0048705	Skeletal system morphogenesis	0,0113929	0,0313269	
GO0051303	Establishment of chromosome localization	0,0140624	0,0077311	<i>glyma14g14290</i>
GO0048471	Perinuclear region of cytoplasm	0,0631415	0,0158733	
GO0045296	Cadherin binding	0,0113929	0,0313269	
GO0043406	Positive regulation of MAP kinase activity	0,0113929	0,0313269	
GO0045436	Lycopene beta cyclase activity	0,0412186	0,0690183	
GO0043625	Delta DNA polymerase complex	0,0649354	0,022635	<i>glyma06g16110</i>
GO0034100	Não tem acesso	0,0535311	0,0584337	
GO0021795	Cerebral cortex cell migration	0,0113929	0,0313269	
GO0021870	Cajal-Retzius cell differentiation	0,0113929	0,0313269	
GO0030838	Positive regulation of actin filament polymerization	0,0470622	0,0506017	
GO0031540	Regulation of anthocyanin biosynthetic process	0,0477145	0,0567988	
GO0030687	Preribosome, large subunit precursor	0,0661505	0,0476978	<i>glyma03g39460</i>

Table S6 – continuation.

GO	Description	<i>p</i> -value (NTSxB1S)	<i>p</i> -value (NTSxB3S)	DEGs
GO0016080	Synaptic vesicle targeting	0,0113929	0,0313269	
GO0015871	Choline transport	0,0113929	0,0313269	
GO0015148	D-xylose transmembrane transporter activity	0,0195392	0,0203114	
GO0015168	Glycerol transmembrane transporter activity	0,0195392	0,0203114	
GO0015575	Mannitol transmembrane transporter activity	0,0195392	0,0203114	
GO0015576	Sorbitol transmembrane transporter activity	0,0195392	0,0203114	
GO0015591	D-ribose transmembrane transporter activity	0,0195392	0,0203114	
GO0015030	Cajal body	0,0475678	0,0653382	<i>glyma15g06160</i>
GO0015174	Basic amino acid transmembrane transporter activity	0,0549658	0,047881	
GO0008793	Aromatic-amino-acid:2-oxoglutarate aminotransferase activity	0,0351703	0,026185	
GO0008310	Single-stranded DNA 3'-5' exodeoxyribonuclease activity	0,0393401	0,0550234	<i>glyma06g16110</i>
GO0008615	Pyridoxine biosynthetic process	0,0486372	0,0413245	
GO0009051	Pentose-phosphate shunt, oxidative branch	0,0533892	0,0515207	
GO0009045	Xylose isomerase activity	0,0689042	0,0573739	
GO0007427	Epithelial cell migration, open tracheal system	0,0192273	0,0642245	
GO0006304	DNA modification	0,0322007	0,071284	
GO0006108	Malate metabolic process	0,0394405	0,0952481	
GO0006287	Base-excision repair, gap-filling	0,0567875	0,0882764	<i>glyma06g16110</i>
GO0005365	Myo-inositol transmembrane transporter activity	0,0195392	0,0203114	
GO0016628	oxidoreductase activity, acting on the CH-CH group of donors, NAD or NADP as acceptor	0,01966	0,048281	<i>glyma02g26670</i>

Table S6 – continuation.

GO	Description	<i>p</i> -value (NTSxB1S)	<i>p</i> -value (NTSxB3S)	DEGs
GO0005548	Phospholipid transporter activity	0,0492278	0,0565257	
GO0004104	Cholinesterase activity	0,0174741	0,00699706	<i>glyma02g08600</i>
GO0004471	Malate dehydrogenase (decarboxylating) (NAD ⁺) activity	0,0463755	0,0530061	
GO0004163	Diphosphomevalonate decarboxylase activity	0,0545798	0,0266303	<i>glyma10g42630</i>
GO0000186	Activation of MAPKK activity	0,0113929	0,0313269	
GO0002098	tRNA wobble uridine modification	0,0263236	0,0241391	
GO0000089	Mitotic metaphase	0,0661505	0,0476978	<i>glyma03g39460</i>
GO0005354	Galactose transmembrane transporter activity	0,0245705	0,0241379	

5 CONSIDERAÇÕES FINAIS

As plantas estão continuamente expostas a mudanças das condições ambientais que podem ameaçar a sua sobrevivência. Mecanismos complexos evoluíram para monitorar com precisão o ambiente e reprogramar o seu metabolismo e crescimento. A disponibilidade de água, que pode ser limitada pela seca, salinidade ou congelamento, é um dos principais fatores que limitam o crescimento e o desenvolvimento das plantas em ambientes agrícolas. Dada a sua importância para a agricultura, os efeitos da seca no desenvolvimento das plantas têm sido extensivamente estudados nas últimas décadas. Isso contribuiu significativamente para a compreensão das respostas fisiológicas e moleculares à limitação de água.

O objetivo do presente trabalho foi investigar o mecanismo de tolerância das plantas transgênicas de sojas que superexpressam o gene *SnOLP* de *Solanum nigrum* var. *americanum*. Apesar de ainda não ser possível propor um modelo de tolerância desencadeada pela osmotina, os resultados obtidos no presente estudo permitiram a identificação de potenciais genes candidatos.

A próxima etapa será a validação do perfil de expressão dos genes candidatos identificados pelas duas técnicas utilizadas. Está previsto também um experimento para completar as análises das respostas fisiológicas, fenotípicas e bioquímicas das plantas transgênicas quando submetidas ao estresse hídrico. As seguintes características serão avaliadas: data da emergência, data do florescimento, cor da flor, variação do índice de vegetação por diferença normalizada (NDVI), fluorescência da clorofila, hábito de crescimento, altura da planta, número de entrenós, peso seco da raiz, peso seco da parte aérea, número de sementes e peso de 100 sementes. Durante a realização do experimento serão coletadas folhas das plantas transgênicas e controles. A partir do extrato de proteína solúvel total será avaliado o conteúdo de peróxido de hidrogênio (H_2O_2) e atividade de enzimas de detoxificação de espécies reativas de oxigênio (APX, SOD, CAT). Serão também avaliadas variáveis de ajustamento osmótico, como a concentração de osmoreguladores (Carboidratos solúveis, Prolina, Glicina-betaína e amino ácidos livres totais).

Além disso, serão utilizadas ferramentas de Biologia de Sistemas, para análise de redes de interações proteicas, visando à elaboração de um modelo biológico que poderá auxiliar no entendimento das respostas moleculares, celulares e fisiológicas induzidas nas plantas transgênicas sob condições de estresse.

6 APÊNDICE: RESULTADOS PRELIMINARES

6.1 Caracterização do promotor do gene WRKY106 de soja

Os promotores mais comumente utilizados nos experimentos de transformação genética de plantas são protegidos por propriedade intelectual, o que restringe seu uso. A identificação de novos promotores, ainda não protegidos, ou o aperfeiçoamento de promotores já conhecidos e que tenham mais alta eficiência na regulação da expressão gênica poderá facilitar a obtenção de novos eventos transgênicos.

Na maioria das plantas transgênicas disponíveis atualmente no mercado, o transgene é expresso sob a regulação de promotores fortes que induzem uma expressão constitutiva, isto é, a todas as partes do vegetal e ao longo de todas as etapas do desenvolvimento, o que causa gastos energéticos desnecessários, dificulta a recuperação (regeneração) de indivíduos transgênicos e aumenta as exigências com relação às comprovações de biossegurança. Em alguns casos, o alto nível de expressão da proteína recombinante pode ser detrimental para a planta, devido a sua natureza tóxica. Em outros casos, a proteína pode não ser necessária em certos tecidos e/ou estádios do desenvolvimento. Portanto, seria desejável ter a expressão do transgene regulada por promotores tecido-específicos ou induzíveis. Promotores induzíveis são ativados por fatores bióticos ou abióticos, podendo-se evitar a expressão do gene em estágios que interferem no crescimento, regeneração ou reprodução da planta transgênica.

As proteínas WRKY formam uma grande família de fatores de transcrição envolvidos em importantes processos fisiológicos e bioquímicos em plantas, incluindo a resposta à deficiência hídrica. Em um estudo prévio do nosso grupo, o padrão de expressão determinado por RT-qPCR mostrou que o gene *GmWRKY106* é diferencialmente expresso entre um genótipo de soja tolerante (Embrapa 48) e um sensível (BR 16) à seca, em condições de estresse hídrico. O gene *GmWRKY106* se destacou pela maior expressão nas raízes da cultivar tolerante.

O presente estudo tem por objetivo o isolamento e a caracterização funcional do promotor do gene *GmWRKY106* de ambos os genótipos. A região promotora de 1 kb a montante do sítio de início de transcrição (TSS) foi utilizada para identificar *cis*-elementos. A análise foi realizada usando o banco de dados Plant Pan (<http://plantpan2.itps.ncku.edu.tw/>). Somente dois *cis*-elementos foram identificados: “GGTCGT” e “CTCC”, envolvidos na resposta à luz.

Fragmentos da sequência promotora (~ 0,5, 1 e 2 Kb) do gene *WRKY106* foram amplificados, clonados no vetor pENTR™ / D-TOPO® e recombinados no vetor pHGWFS7 (Karimi et al., 2002), usando o sistema de clonagem Gateway (Figura 1). Células de *Agrobacterium rhizogenes* da linhagem K599 foram transformadas por eletroporação. Como controle positivo, o promotor CaMV35S foi submetido ao mesmo processo de clonagem.

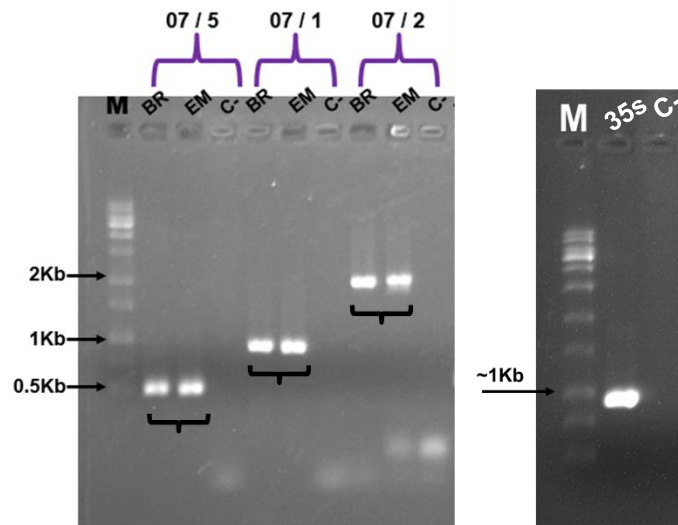


Figura 1- Fragmentos de 0,5kb, 1kb e 2kb da região promotora do gene *WRKY106*. Produto de PCR visualizados em gel de agarose 1,2%. M = Marcador de DNA 1Kb; BR = Fragmentos do genótipo sensível BR16; EM = Fragmentos do genótipo tolerante Embrapa 48; 07/5 = Fragmento de ~0.5Kb a montante do gene *GmWRKY106*; 07/1 = Fragmento de ~1Kb a montante do gene *GmWRKY106*; 07/2 = Fragmento de ~2Kb a montante do gene *GmWRKY106*; 35S = Controle positivo; C- = controle negativo da reação.

Sementes da cultivar IAS5 foram esterilizadas com etanol 70% durante 1min e solução de hipoclorito 3% por 20 min. Foram enxaguadas três vezes com água destilada autoclavada e imediatamente colocadas em um meio de cultura ½ MS (Murashige and Skoog 1962) sólido para a germinação. Colônias isoladas de *A. rhizogenes* K599 contendo os vetores pH7WG2D::35S ou pH7WG2D::1kb (BR16) foram inoculadas e incubadas em 10 mL de meio de extrato de levedura peptona (YEP) contendo estreptomicina 100 mg/L e espectinomicina 100 mg/L, a 28 °C durante 24 h, sob agitação. Os pré-inóculos foram inoculados em 500 mL do mesmo meio por mais 24 h. Após, as culturas foram centrifugadas a 500 rpm a 4 °C por 10 minutos e os *pellets* foram ressuspensos em 8 mL de água mili-Q autoclavada.

Para a obtenção de plantas compostas (com raízes transgênicas e parte aérea não transformada) foram realizadas transformações com plântulas com cinco dias após a germinação. A radícula foi removida e cada plântula foi imediatamente inoculada com 0,1 mL de células de *A. rhizogenes* contendo a construção. As plântulas inoculadas foram transferidas para um meio de co-cultivo (6 g/L de ágar, ½ MS, vitamina B5 de Borg, 1,90 g/L MES, 1 mL/L Ditiotreitolo, tiosulfato de sódio 1 M, pH 7,0). As plantas não transformadas foram submetidas aos mesmos procedimentos, porém injetadas com água em vez da cultura bacteriana.

Após quatro dias, as plantas foram transferidas para um meio de seleção contendo Cefotaxima (100 µg/mL) e Higromicina (15 µg/mL), permanecendo por 10 dias. A primeira seleção de raízes transformadas foi realizada pela observação da expressão GFP sob um estereomicroscópio de fluorescência Leica MZIII (Leica Microsystems GmbH, Wetzlar, Alemanha). Após este período, as plantas foram colocadas em meio hidropônico contendo meio ½ MS. Foi realizado um bioensaio para verificar a resposta do promotor à seca. O estresse foi estabelecido retirando-se as plantas do seu cultivo hidropônico, durante 45 minutos.

Este ensaio preliminar demonstrou aumento da expressão de GFP nas raízes transgênicas, quando o gene está sob o controle do promotor BR16 de 1kb (Figura 2A). O promotor CaMV35S foi utilizado como controle positivo em raízes transformadas (Figura 2B). Nas raízes não transformadas não houve expressão de GFP (Figura 2C).

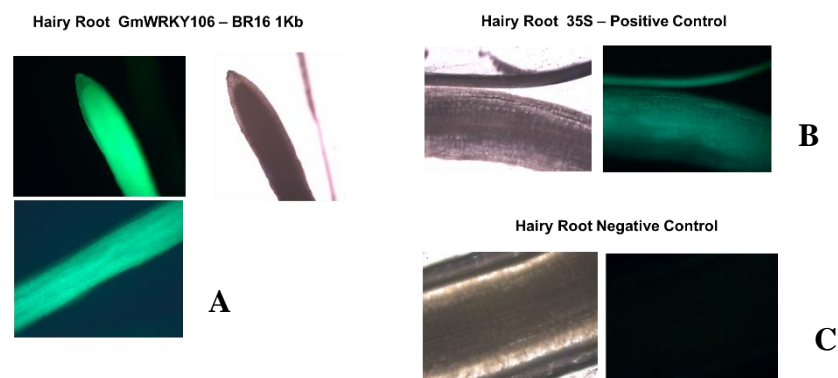


Figura 2: Análise das raízes transformadas sob estresse hídrico após 45 minutos de tratamento. (A) Expressão de GFP quando o gene estava sob o controle do promotor BR16 1kb. (B) Expressão de GFP quando o gene estava sob o controle do promotor CaMV35S. (C) Ausência da expressão de GFP em raízes não transformadas.

Foi realizado também, um bioensaio preliminar em meio de cultura MS contendo NaCl (150mM). Os resultados demonstraram que quando o gene estava sob o controle do promotor CaMV35S, foi observada a expressão de GFP (Figura 3A). Não foi observado aumento da expressão de GFP quando o gene estava sob o controle do promotor BR16 de 1Kb (Figura 3B). Nas raízes não transformadas não houve expressão de GFP (Figura 3C).

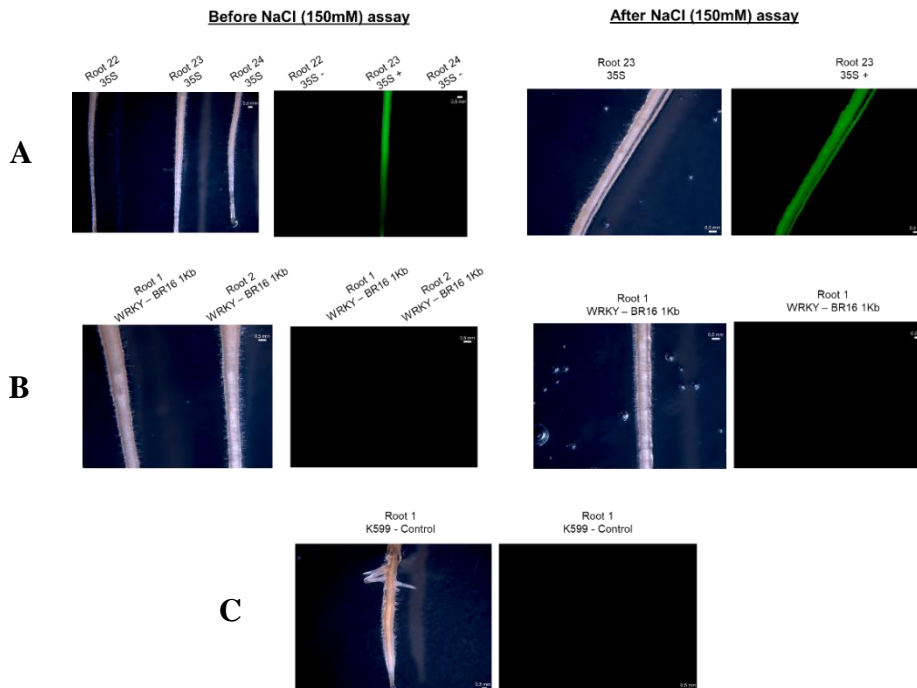


Figura 3: Análise das raízes transformadas em meio MS contendo NaCl (150Mm), após 4 dias de tratamento. (A) Expressão de GFP quando o gene estava sob o controle do promotor CaMV35S. (B) Não houve expressão de GFP quando o gene estava sob o controle do promotor BR16 1kb. (C) Ausência da expressão de GFP em raízes não transformadas.

Visando a obtenção de plantas completas transformadas, embriões somáticos de soja, foram transformados por bombardeamento, com a utilização de um acelerador de partículas de baixa pressão, modelo PIG (Finer et al., 1992), seguindo protocolo estabelecido por Droste et al. (2002). As construções utilizadas continham: fragmentos de 0,5 ou 1kb de WRKY106, de ambos os genótipos, ou o promotor CaMV35S, no vetor pH7WG2D. Como controle, alguns conjuntos de embriões foram bombardeados com partículas de tungstênio livres de DNA.

Após o bombardeamento, os embriões foram mantidos nas mesmas placas contendo meio D20 por dez dias. Os conjuntos de embriões foram transferidos para meio D20 contendo 12,5

mg/L de higromicina. Após 21 dias, os conjuntos foram transferidos novamente para meio D20 contendo 25 mg/L de higromicina por 90 dias, com troca de meio a cada 15 dias.

Após o período de seleção, cada conjunto de embriões resistentes à higromicina e proliferantes, foi separado e proliferado individualmente, por mais 30 dias, em meio D20 sem o antibiótico. Os conjuntos de embriões higromicina resistentes foram transferidos para meio de histodiferenciação - MSM6 modificado (Finer & McMullen, 1991), composto por sais MS, vitaminas B5, 0,3% Phytigel, pH 6,4, 1% carvão ativado e 6% sacarose. Após 30 dias, os embriões foram cuidadosamente individualizados e transferidos para o meio de histodiferenciação (o mesmo MSM6, sem carvão ativado) onde permanecerão por mais 30 dias. O trabalho encontra-se nesta etapa de desenvolvimento.

As próximas etapas serão: (i) dessecação parcial dos embriões histodiferenciados, o que contribui com o aumento da frequência e da velocidade de regeneração; (ii) transferência para o meio de regeneração MS0 (MSM6, com 3% de sacarose e sem carvão ativado); (iii) após a emissão de raízes e folhas, as plântulas serão transferidas para vidros contendo o mesmo meio de regeneração, onde permanecerão até atingirem cerca de 4 cm; (iv) transferência para copos plásticos contendo vermiculita, para aclimação; (v) transplante para o solo.

Para avaliar a atividade do promotor em estudo, a expressão do gene repórter *GFP* será analisada nas plantas transgênicas obtidas e submetidas a diferentes estresses, tais como: estresse hídrico, tratamento com sal, tratamento com frio, inducibilidade por ABA. A expressão do gene repórter *GFP* será observada sob microscopia de fluorescência (Estereomicroscópio Leica MZIII - Leica Microsystems GmbH, Wetzlar, Germany).

REFERÊNCIAS

- Abdin, M.Z., Kiran, U., Alam, A. (2011). Analysis of osmotin, a PR protein as metabolic modulator in plants. *Bioinformation*, 5, 336–340.
- Arraes, F. B. M., Beneventi, A., De Sa, M. E. L., Paixao, J. F. R., Albuquerque, E. V. S., Marin, S. R. R., Purgatto, E., Nepomuceno, A. L., GrossI-De-Sa, M. F. (2015). Implications of ethylene biosynthesis and signaling in soybean drought stress tolerance. *BMC Plant Biology*, 15 (1), 213.
- Balbinot Junior, A. A.; Hirakuri, M. H.; Franchini, J. C.; Debiasi, H.; Ribeiro, R. H. (2017). Análise da área, produção e produtividade da soja no Brasil em duas décadas (1997-2016). Boletim de Pesquisa Número 11, [s. l.], p. 21.
- Barthakur, S., Babu, V., Bansal, K. C. (2001). Overexpression of osmotin induces proline accumulation and confers drought- and low-temperature stress tolerance in tobacco by gene transfer. *Plant Cell Physiol*, 45, 346–350.
- Belamkar, V., Weeks, N. T., Bharti, A. K., Farmer, A. D., Graham, M. A., Cannon, S. B. (2014). Comprehensive characterization and RNA-Seq profiling of the HD-Zip transcription factor family in soybean (*Glycine max*) during dehydration and salt stress. *BMC Genomics*, 15 (1), 950.
- Campos, M. A., Ribeiro, S. G., Rigden, D. J., & Monte, D. C. (2002). Putative pathogenesis-related genes within *Solanum nigrum L. var. americanum* genome: isolation of two genes coding for PR5-like proteins, phylogenetic and sequence analysis. *Physiological and Molecular Plant Pathology*, 61 (4), 205-216.
- Chen, L. M., Zhou, X. A., Li, W. B., Chang, W., Zhou, R., Wang, C., Sha, A. H., Shan, Z. H., Zhang, C. J., Qiu, D. Z., Yang, Z. L., Chen, A. L. (2013). Genome-wide transcriptional analysis of two soybean genotypes under dehydration and rehydration conditions. *BMC Genomics*, 14 (1), 687.
- Chen, L., Cai, Y., Liu, X., Guo, C., Sun, S., Wu, C., Jiang, B., Han, T., Hou, W. (2018). Soybean hairy roots produced in vitro by *Agrobacterium rhizogenes*-mediated transformation. *The Crop Journal*, 6 (2), 162-171.
- Chen, L. M., Fang, Y. S., Zhang, C. J., Hao, Q. N., Cao, D., Yuan, S. L., Chen, H. F., Yang, Z. L., Chen, S. L., Shan, Z. H., Liu, B. H., Wang, J., Zhan, Y., Zhang, X. J., Qiu, D. Z., Li, W. B., Zhou, X. A. (2019). GmSYP24, a putative syntaxin gene, confers osmotic/drought, salt stress tolerances and ABA signal pathway. *Scientific Reports*, 9 (1), 5990.
- CONAB - Companhia Nacional de Abastecimento, Acompanhamento da Safra Brasileira de Grãos, Safra 2018/19, Análise Mensal, disponível em: <www.conab.gov.br>
- CONAB - Companhia Nacional de Abastecimento, Acompanhamento da Safra Brasileira de Grãos, Safra 2016/17, disponível em: <www.conab.gov.br>

- Conesa, A., Madrigal, P., Tarazona, S., Gomez-cabrero, D., Cervera, A., Mcpherson, A., Szczesniak, M. W., Gaffney, D. J., Elo, L. L., Zhang, X., Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, 17 (1), 13.
- Cramer, G. R., Urano, K., Deltrot S., Pezzotti, M., Shinozaki, K. (2011). Effects of abiotic stress on plants: a systems biology perspective. *BMC plant biology*, 11 (1), 163.
- Das M., Chauhan, H., Chhibbar, A., Rizwanul Haq, Q. M., Khurana, P. (2011). High-efficiency transformation and selective tolerance against biotic and abiotic stress in mulberry, *Morus indica* cv. K2, by constitutive and inducible expression of tobacco osmotin. *Transgenic research*, 20 (2), 231-246.
- Droste, A., Pasquali, G., Bodanese- Zanettini, M. H. (2002). Transgenic fertile plants of soybean [*Glycine max* (L.) Merrill] obtained from bombarded embryogenic tissue. *Euphytica*, 127 (3), 367-376.
- Fàbregas, N., Fernie, A. R. (2019). The metabolic response to drought. *Journal of experimental botany*, 70 (4), 1077-1085.
- Finer, J. J., Vain, P., Jones, M. W., & McMullen, M. D. (1992). Development of the particle inflow gun for DNA delivery to plant cells. *Plant Cell Reports*, 11 (7), 323-328.
- Goel, D., Singh, A. K., Yadav, V., Babbar, S. B., Bansal, K. C. (2010). Overexpression of osmotin gene confers tolerance to salt and drought stresses in transgenic tomato (*Solanum lycopersicum l.*). *Protoplasma*, 245, 133-141.
- Hakim, A. U., Amjad, H., Muhammad, S., Aamir, H. K., Muna, A., Summia, G., Zhang, J., Sun, L., Jianying, Li., Shuangxia, Jin., Muhammad, F. H. M. (2018). Osmotin: A plant defense tool against biotic and abiotic stresses. *Plant Physiology and Biochemistry*, 123, 149-159.
- James, A. T., Lawn, R. J., and Cooper, M. (2008). Genotypic variation for drought stress response traits in soybean. Inter-relations between epidermal conductance, osmotic potential, relative water content, and plant survival. *Australian Journal of Agricultural Research*, 59 (7), 670-678.
- Karimi, M., Inzé, D., & Depicker, A. (2002). GATEWAY™ vectors for Agrobacterium-mediated plant transformation. *Trends in plant science*, 7 (5), 193-195.
- Koiwa, H., Kato, H., Nakatsu, T., Oda, J., Yamada, Y., Sato, F. (1999). Crystal structure of tobacco PR-5d protein at 1.8 Å resolution reveals a conserved acidic cleft structure in antifungal thaumatin-like proteins. *Journal of molecular biology*, 286 (4), 1137–1145.
- Li, Y., Ye, W., Wang, M., & Yan, X. (2009). Climate change and drought: a risk assessment of crop-yield impacts. *Climate research*, 39 (1), 31-46.

Liu, F., Andersen, M. N., and Jensen, C. R. (2003). Loss of pod set caused by drought stress is associated with water status and ABA content of reproductive structures in soybean. *Functional Plant Biology*, 30 (3), 271–280.

Liu, J. J., Sturrock, R., & Ekramoddoullah, A. K. (2010). The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. *Plant cell reports*, 29 (5), 419-436.

Manavalan, L. P., Guttikonda, S. K., Tran, L. S. P., Nguyen, H. T. (2009). Physiological and molecular approaches to improve drought resistance in soybean. *Plant e Cell Physiology*. 50: 1260-1276.

Min, K., Ha, S. C., Hasegawa, P. M., Bressan, R. A., Yun, D. J., Kim, K. K., (2004). Crystal structure of osmotin, a plant antifungal protein. *Proteins: Structure, Function, and Bioinformatics*, 54, 170–173.

Nawaz, M. A., Rehman, H. M., Baloch, F. S., Ijaz, B., Ali, M. A., Khan, I. A., Lee, J. D., Chung, G., Yang, S. H. (2017). Genome and transcriptome-wide analyses of cellulose synthase gene superfamily in soybean. *Journal of Plant Physiology*, 215: 163-175. 2017.

Noori, S. A. S., Sokhansanj, A. (2008). Wheat plants containing an osmotin gene show enhanced ability to produce roots at high NaCl concentration. *Russian journal of plant physiology*, 55 (2), 256-258.

Oh, M. W., Komatsu, S. (2015). Characterization of proteins in soybean roots under flooding and drought stresses. *Journal of Proteomics*, 114, 161–181.

Parkhi, V., Kumar, V., Sunilkumar, G., Campbell, L. M., Singh, N. K., Rathore, K. S. (2009). Expression of apoplastically secreted tobacco osmotin in cotton confers drought tolerance. *Molecular breeding*, 23 (4), 625-639.

Rodrigues, F. A., Fuganti-Pagliarini, R., Marcolino-Gomes, J., Nakayama, T. J., Molinari, H. B. C., Lobo, F. P., Harmon, F. G., Nepomuceno, A. L. (2015). Daytime soybean transcriptome fluctuations during water deficit stress. *BMC Genomics*, 16 (1), 505.

Rybarczyk-Filho, J. L., Castro, M. A. A., Dalmolin, R. J. S., Moreira, J. C. F., Brunnet, L. G., De Almeida, R. M. C. (2011). Towards a genome-wide transcriptogram: the *Saccharomyces cerevisiae* case. *Nucleic acids research*, 39 (8), 3005-3016.

Sharma, M., Gupta, S. K., Majumder, B., Maurya, V. K., Deeba, F., Alam, A., Pandey, V. (2018). Proteomics unravel the regulating role of salicylic acid in soybean under yield limiting drought stress. *Plant Physiology and Biochemistry*, 130, 529-541.

Shin, J. H., Vaughn, J. N., Abdel-Haleem, H., Chavarro, C., Abernathy, B., Kim, K. D., Jackson, S. A., LI, Z. (2015). Transcriptomic changes due to water deficit define a general soybean response and accession-specific pathways for drought avoidance. *BMC Plant Biology*, 15 (1), 26.

- Stintzi A., Heitz T., Kauffmann S., Legrand M., Fritig B. (1991). Identification of a basic pathogenesis-related, thaumatin-like protein of virus-infected tobacco as osmotin. *Physiological and Molecular Plant Pathology*, 38 (2), 137-146.
- Subramanyam, K., Sailaja, K. V., Subramanyam, K., Muralidhara R, D., Lakshmidevi, K. (2011). Ectopic expression of an osmotin gene leads to enhanced salt tolerance in transgenic chilli pepper (*Capsicum annum L.*). *Plant Cell, Tissue and Organ Culture (PCTOC)*, 105 (2), 181-192.
- Subramanyam, K., Arun, M., Mariashibu, T. S., Theboral, J., Rajesh, M., Singh, N. K., Manickavasagam, M., Ganapathi, A. (2012). Overexpression of tobacco osmotin (Tbosm) in soybean conferred resistance to salinity stress and fungal infections. *Planta*, 236 (6), 1909-1925.
- Tripathi, P.; Rabara, R. C.; Reese, R. N.; Miller, M. A.; Rohila, J. S.; Subramanian, S.; Qingxi, J. S.; Morandi, D.; Bücking, H.; Shulaev, V.; Rushton, P. J. (2016). A toolbox of genes, proteins, metabolites and promoters for improving drought tolerance in soybean includes the metabolite coumestrol and stomatal development genes. *BMC Genomics*, 17 (1), 102.
- Viktorova, J., Krasny, L., Kamlar, M., Novakova, M., Mackova, M., & Macek, T. (2012). Osmotin, a pathogenesis-related protein. *Current Protein and Peptide Science*, 13 (7), 672-681.
- Xu, K., Zhang, X., Fan, C., Chen, F., Zhu, J., Zhang, S., Chen, Q., Fu, Y. (2017). A callus transformation system for gene functional studies in soybean. *Journal of Integrative Agriculture*, 16 (9), 1913-1922.
- Yin, G., Xu, H., Liu, J., Gao, C., Sun, J., Yan, Y., Hu, Y. (2014). Screening and identification of soybean seed-specific genes by using integrated bioinformatics of digital differential display, microarray, and RNA-seq data. *Gene*, 546 (2): 177-186.
- Weber, R. L. M., Wiebke-Strohm, B., Bredemeier, C., Margis-pinheiro, M., Brito, G. G., Rechenmacher, C., Bertagnolli, P. F., Sá, M. E. L., Campos, M. A., Amorim, R. M. S., Beneventi, M. A. Margis, R., Grossi-De-Se, M. F., Bodanese-Zanettini, M. H. (2014). Expression of an osmotinlike protein from *Solanum nigrum* confers drought tolerance in transgenic soybean. *BMC Plant Biology*, 14, 343-352.
- Wijewardana, C., Reddy, K. R., Bellaloui, N. (2018). Soybean seed physiology, quality, and chemical composition under soil moisture stress. *Food Chemistry*, 278, 92–100.
- Zhang, H., Song, B. (2017). RNA-seq data comparisons of wild soybean genotypes in response to soybean cyst nematode (*Heterodera glycines*). *Genomics Data*, 14, 36-39.