



**UNIVERSIDADE
ESTADUAL DE LONDRINA**

JULIANA MARCOLINO RIBEIRO

**PROSPECÇÃO DE GENES DE SOJA EXPRESSOS SOB
CONDIÇÕES DE DÉFICIT HÍDRICO**

JULIANA MARCOLINO RIBEIRO

**PROSPECÇÃO DE GENES DE SOJA EXPRESSOS SOB
CONDIÇÕES DE DÉFICIT HÍDRICO**

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

Orientador: Dr. Alexandre Lima Nepomuceno

Co-orientadora: Dra. Fabiana Aparecida Rodrigues

Londrina
2011

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

Dados Internacionais de Catalogação-na-Publicação (CIP)

R484pT Ribeiro, Marcolino Juliana.
Prospecção de genes de soja expressos sob condições de déficit hídrico/ Juliana
Marcolino Ribeiro. – Londrina, 2011.
83 f. : il.

Orientador: Alexandre Lima Nepomuceno.
Dissertação (Mestrado em Genética e biologia molecular) – Universidade
Estadual de Londrina, Centro de Ciências Biológicas, Programa de Pós-Graduação em
Ciências Biológicas, 2011.
Inclui bibliografia.

1. Transcritoma – Teses. 2. Expressão gênica – Teses. 3. Seca – Teses. I.
Nepomuceno.II. Universidade Estadual de Londrina. Centro de Ciências
Biológicas. Programa de Pós-Graduação em Ciências Biológicas. III. Título.

CDU 631.5234

JULIANA MARCOLINO RIBEIRO

**PROSPECÇÃO DE GENES DE SOJA EXPRESSOS SOB
CONDIÇÕES DE DÉFICIT HÍDRICO**

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

BANCA EXAMINADORA

Prof. Dr. Alexandre Lima Nepomuceno
EMBRAPA – Soja – UEL

Prof. Dr. Luiz Filipe Protasio Pereira
EMBRAPA - Café – UEL

Dr. Hugo Bruno Correa Molinari
EMBRAPA - Agroenergia

Londrina, 23 de fevereiro de 2011.

*Aos meus Pais, que me
fizeram a mulher que sou
hoje; Ao meu
querido esposo, que faz
cada dia valer à pena.*

AGRADECIMENTOS

Agradeço primeiramente a Deus, pela vida, pela saúde e por ter preenchido meu caminho com pessoas e oportunidades maravilhosas, sem as quais este trabalho não poderia ser realizado;

Aos meus pais, os maiores exemplos de vida que já conheci, e que me ensinaram a superar todos os desafios e a não ter medo de correr atrás dos meus sonhos;

Ao meu esposo, Douglas, por ter me ajudado em todos os momentos da minha vida, me animando, ensinando e, principalmente, por ter o dom de preencher meu caminho de amor, alegria e esperança;

Agradeço ao meu orientador, Dr. Alexandre Lima Nepomuceno, pelo voto de confiança ao acolher-me em sua equipe, possibilitando que eu realizasse este trabalho;

À minha co-orientadora, Dra Fabiana Aparecida Rodrigues, pela paciência e dedicação que me fizeram crescer, não só na vida acadêmica, mas como ser humano. É um privilégio desfrutar de sua amizade;

À Dr. Ana Lúcia Dias, pelo apoio e incentivo que me acompanham desde a Graduação;

À Dr. Maria Cristina, pela ajuda fundamental nas análises estatísticas;

À Dr. Mariangela Hungria, pela possibilidade de utilizar as instalações de seu laboratório nas análises de PCR quantitativo;

Às instituições UEL, CAPES, EMBRAPA e ao Programa de Pós-graduação em Genética e Biologia Molecular, pela oportunidade de aprendizado.

Aos meus amigos do Laboratório de Biotecnologia Vegetal – Embrapa Soja, por tantas vezes terem me ajudado a superar as dificuldades: vocês estão no meu coração.

*“Lord, Don't Move the Mountain,
Just Give me the Strength to
Climb”*

Mahalia Jackson

RIBEIRO, Juliana Marcolino. **Prospecção de genes de soja expressos sob condições de déficit hídrico**. 83f.2011.Dissertação (Mestrado em Genética e Biologia Molecular) – Universidade Estadual de Londrina. Londrina, 2011.

RESUMO

A soja (*Glycine max* (L.) Merrill) é uma cultura economicamente importante, sendo utilizada principalmente na alimentação humana e animal, mas também com um grande potencial para se tornar uma importante fonte de biocombustíveis. Entre os fatores que influenciam a produtividade, a seca é o principal fator limitante e o de mais difícil controle. Neste estudo, foram analisadas duas cultivares de soja contrastantes quanto a tolerância / sensibilidade à seca: BR 16 (sensível) e Embrapa 48 (tolerante). As plantas foram cultivadas em sistema hidropônico até o estágio V3, quando o tratamento de déficit hídrico foi imposto, consistindo na exposição das raízes à desidratação. Bibliotecas subtrativas foram construídas e sequenciadas por equipamento *Genome Analyzer IIe* (Illumina). Análises *in silico* das sequências obtidas permitiu a identificação de vários genes diferencialmente expressos em raízes das duas cultivares durante o déficit hídrico. A categorização de acordo com os processos biológicos envolvidos revelou as alterações celulares em resposta à seca e as diferenças nos perfis transcricionais entre as cultivares. Além disso, foram encontrados diversos fatores de transcrição, dentre eles, genes da família AP2/EREBP. Análises filogenéticas destes genes revelaram estreita relação com genes DREB de Fabaceae. Ensaios de qPCR confirmaram a expressão de seis candidatos AP2/EREBP em resposta à seca, validando os resultados obtidos nas bibliotecas subtrativas. Curiosamente, em nossas bibliotecas, foram identificados 12 fatores de transcrição ainda não relacionados com a resposta ao déficit hídrico e formulamos hipóteses sobre seu papel neste tipo de estresse.

Palavras-chave: Transcriptoma. Expressão gênica. Seca.

RIBEIRO, Juliana Marcolino. **Search for genes expressed in soybean under water deficit.**2011. 83 p. Dissertation (Master's degree in Genetics and Molecular Biology) –Universidade Estadual de Londrina, Londrina, 2011.

ABSTRACT

Soybean (*Glycine max* (L.) Merrill) is an economically important crop, being used mainly in human nutrition and animal feed, but also with a great potential to become an important source of biofuel. Among factors influencing productivity, drought is the major limiting factor and the most difficult to control. In this study, we analyzed two soybean cultivars contrasting in tolerance/sensitivity to drought: BR 16 (sensitive) and Embrapa 48 (tolerant). Plants were grown in a hydroponic system until the V3 stage, when the water deficit treatment was imposed, consisting in root exposure to dehydration. Subtractive libraries were constructed and sequenced by *Genome Analyzer IIe* (Illumina). *In silico* analyses of these sequences enabled the identification of several differentially expressed genes in roots of both cultivars during water deficit. The categorization according to biological processes (Gene Ontology) revealed cellular changes in response to drought and the differences in the cultivars transcriptional profiles. In addition, many transcription factor were found, among than, AP2/EREBP family genes. Phylogenetic analysis of these genes revealed close relationship with the Fabaceae DREB genes family. qPCR assays confirmed the up regulation of six putative soybean AP2/EREBP genes in response to drought. These results validate the results obtained in the subtractive libraries. Interestingly, in our libraries, 12 transcription factors not related to drought response were identified and we formulated hypotheses about its role in this type of stress.

Key words: Soybean transcriptome. Differential gene expression. Drought.

LISTA DE FIGURAS

- Figura 1 -** Diagrama de Venn mostrando a distribuição dos transcritos identificados nas bibliotecas subtrativas. Os números de genes exclusivos e comuns aos diferentes períodos do déficit hídrico são apresentados para as bibliotecas R1 (25-50 min sob estresse), R2 (75-100 min) e R3 (125-150 min). **(A)** BR 16 e **(B)** Embrapa 48.....45
- Figura 2 -** Distribuição dos Resultados de BLASTX por espécies (% sequências). Porcentagem de sequências por espécies nas cultivares **(A)** BR 16 e **(B)** Embrapa 48, distribuídos nas bibliotecas R1 (T25-50 min sob estresse), R2 (T75-100 min) e R3 (T125-150 min); **(C)** árvore taxonômica com as espécies mais frequentes no resultado de BLASTX de ambas as cultivares47
- Figura 3 -** Categorização funcional dos transcritos. Porcentagem de sequências (%) atribuídas aos respectivos processos biológicos (Gene Ontology) para as cultivares **(A)** BR 16 e **(B)** Embrapa 48, distribuídas nas bibliotecas R1 (T25-50 min sob estresse), R2 (T75-100 min) e R3 (T125-150 min)49
- Figura 4 -** Árvore filogenética com proteínas codificadas pelos genes candidatos e os genes *DREB* descritos no banco de dados NCBI. Genes candidatos são representados pelos respectivos *genemodels* (Glyma) e as sequências de *DREBs* homólogos da família Fabaceae (*Glycine max*, *Medicago truncatula*, *Cypripedium arietinum*, *Trifolium repens*, *Glycine soja*, *Caragana korshinskii*). A árvore foi construída utilizando algoritmo Clustal W no programa MEGA 4.0 com o teste *Bootstrap* (1000 repetições)54

Figura 5 - PCR quantitativo de genes de família AP2/EREB. Expressão relativa dos genes na cultivar BR 16, durante os seis períodos de estresse (25 a 150 min.de déficit hídrico- T25 a T150) em relação ao controle (0 min.- condição sem déficit hídrico). (A)alvo 1 (Glyma10g2); (B)alvo 2 (Glyma14g09); (C)alvo 3 (Glyma04g1); (D)alvo 4 (Glyma14g3); (E)alvo 5 (Glyma13g01) avaliado em folha; (F)alvo 5 (Glyma13g01) avaliado em raiz; (G)alvo 6 (Glyma13g3) avaliado em folha; (H)alvo 6 (Glyma13g3) avaliado em raiz. Barras representam erro padrão da média56

Figura 6 - PCR quantitativo de genes de família AP2/EREB. Expressão relativa dos genes na cultivar Embrapa 48, durante os seis períodos de estresse (25 a 150 min.de déficit hídrico- T25 a T150) em relação ao controle (0 min.-condição sem déficit hídrico). (A)alvo 1 (Glyma10g2); (B)alvo 2 (Glyma14g09); (C)alvo 3 (Glyma04g1); (D)alvo 4 (Glyma14g3); (E)alvo 5 (Glyma13g01) avaliado em folha; (F)alvo 5 (Glyma13g01) avaliado em raiz; (G)alvo 6 (Glyma13g3) avaliado em folha; (H)alvo 6 (Glyma13g3) avaliado em raiz. Barras representam erro padrão da média57

LISTA DE TABELAS

- Tabela 1** -Primers específicos para os genes alvo. Número do gene alvo e respectivo *genemodel* (gene predito em soja). Sequências de bases e temperatura de desnaturação (Tm) dos primers Forward e Reverse. Eficiência refere-se à eficiência das reações de amplificação dos primers, calculada a partir da construção de curvas padrão.....44
- Tabela 2** -Análise dos *contigs* utilizando o BLASTX apresentando o número de *reads* geradas por biblioteca (R1, R2, R3) das cultivares BR 16 e Embrapa 48 e número de *contigs* montados a partir da sobreposição dos *reads* (*Contigs*); sequências sem similaridade (*no hits*); sequências similares a proteínas descritas (*Hits*); *hits* com proteínas conhecidas (Sequências conhecidas) e desconhecidas (Sequências desconhecidas).....46
- Tabela 3**- Número de genes responsivos ao déficit hídrico nas bibliotecas R1 (T25-50 min sob estresse), R2 (T75-100 min) e R3 (T125-150 min), das cultivares (BR 16 e Embrapa 48).....50
- Tabela 4** -Principais fatores de transcrição relacionados ao déficit hídrico. Número de sequências (contigs únicos) identificadas nas bibliotecas (T25-50 min sob estresse), R2 (T75-100 min) e R3 (T125-150 min), das cultivares BR 16 e Embrapa 48.....51
- Tabela 5** -Transcritos expressos nas cultivares BR 16 e Embrapa 48 que apresentaram similaridade aos genes da família AP2/EREB. Os genes candidatos são representados pelo *genemodel* correspondente. As sequência com maior similaridade, o respectivo e-value e identidade (% de nucleotídeos idênticos) são referentes ao resultado do BLASTX contra o banco de dados NCBI.....52

Tabela 6 Genes avaliados por PCR quantitativo. Identificação do gene alvo, nome do gene predito em soja segundo Phytozome e Descrição do gene, segundo análises de similaridade com sequências do NCBI. A última coluna refere-se ao tecido onde o gene foi isolado (por meio das bibliotecas subtrativas), nos quais os genes serão validados via PCR quantitativo.....	55
Tabela 7 -Teste de hipóteses baseado no programa Rest 2009. Genes alvo: alvo 1 (Glyma10g2); alvo 2 (Glyma14g09); alvo 3 (Glyma04g1); alvo 4 (Glyma14g3); alvo 5 (Glyma13g01); alvo 6 (Glyma13g3). P(H1) refere-se à probabilidade do resultado de expressão das plantas tratadas (25 a 150 min.de déficit hídrico) ser diferente das plantas controle (0 min.de déficit hídrico) devido ao acaso. Result. demonstra o resultado da expressão relativa, onde considera $P < 0,05$ para genes significativamente induzidos (<i>UP</i>)	58
Tabela 8 -Novos Fatores de transcrição relacionados ao déficit hídrico em soja. Nome dos genes e organismo que apresentam maior similaridade (segundo análises por BLASTX e AutoFact) com os transcritos obtidos nas bibliotecas de raízes submetidas ao déficit hídrico. Função do fator de transcrição identificado segundo literatura e respectivas referências.....	59

SUMÁRIO

1	INTRODUÇÃO	15
2	OBJETIVO	17
2.1	OBJETIVOS ESPECÍFICOS.....	17
3	REVISÃO BIBLIOGRÁFICA	18
3.1	A SOJA.....	18
3.1.1	Origem e Dispersão.....	18
3.1.2	Importância Econômica	19
3.1.3	O Genoma da Soja	20
3.2.	DÉFICIT HÍDRICO.....	20
3.2.1	Visão Geral.....	20
3.2.2	Mecanismos Morfo-Fisiológicos de Resposta ao Déficit Hídrico	22
3.2.3	Respostas Moleculares ao Déficit Hídrico	25
3.2.4	Técnicas Moleculares e Melhoramento Genético	29
3.2.5	Sequenciamento de DNA de Última Geração (NGS)	31
3.3	REFERÊNCIAS BIBLIOGRÁFICAS	33
4	ARTIGO: ARTIGO A SER SUBMETIDO À REVISTA PLANT SCIENCE (ISSN: 0168-9452)	37
1	INTRODUÇÃO	37
2	MATERIAL E MÉTODOS	40
2.1	DESENHO EXPERIMENTAL E MATERIAL BIOLÓGICO.....	40
2.2	ANÁLISES FISIOLÓGICAS	40
2.3	EXTRAÇÃO DE RNA E ISOLAMENTO DE MRNA.....	41
2.4	CONSTRUÇÃO DAS BIBLIOTECAS SUBTRATIVAS.....	41
2.5	SEQUENCIAMENTO E BIOINFORMÁTICA.....	42
2.6	VALIDAÇÃO DAS BIBLIOTECAS SUBTRATIVAS E ANÁLISE DE EXPRESSÃO GÊNICA VIA PCR QUANTITATIVO	43

3	RESULTADOS	45
3.1	BIBLIOTECAS SUBTRATIVAS DE CDNA.....	45
3.2	ANÁLISES DE BIOINFORMÁTICA	45
3.3	ALTERAÇÕES MOLECULARES EM DECORRÊNCIA DO DÉFICIT HÍDRICO	48
3.4	GENES RELACIONADOS AO DÉFICIT HÍDRICO	49
3.5	FATORES DE TRANSCRIÇÃO	51
3.5.1	Fatores de Transcrição Relacionados Ao Déficit Hídrico.....	51
3.5.2	Fatores de Transcrição AP2/ERE BP	51
3.6	VALIDAÇÃO DAS BIBLIOTECAS SUBTRATIVAS E ANÁLISE DE EXPRESSÃO DE GENES DA FAMÍLIA AP2/ERE BP	55
3.7	NOVOS FATORES DE TRANSCRIÇÃO NA RESPOSTA AO DÉFICIT HÍDRICO	58
4	DISCUSSÃO	59
4.1	SEQUENCIAMENTO E ANÁLISES DE BIOINFORMÁTICA	59
4.2	DINÂMICA DOS PROCESSOS BIOLÓGICOS DURANTE O DÉFICIT HÍDRICO EM BR 16 E EMBRAPA	61
4.3	GENES DIFERENCIALMENTE EXPRESSOS E SUA RELAÇÃO COM O DÉFICIT HÍDRICO	63
4.4	FATORES DE TRANSCRIÇÃO DESCRITOS NA RESPOSTA AO DÉFICIT HÍDRICO.....	65
4.5	VALIDAÇÃO DAS BIBLIOTECAS SUBTRATIVAS E AVALIAÇÃO DA EXPRESSÃO DE GENES DA FAMÍLIA AP2/ERE BP.....	66
4.6	NOVOS FATORES DE TRANSCRIÇÃO NA RESPOSTA AO DÉFICIT HÍDRICO.....	69
5	CONCLUSÃO	73
6	REFERÊNCIAS BIBLIOGRÁFICAS	74
	ANEXOS	79
	ANEXO A – Experimento de déficit hídrico	80
	ANEXO B – Análises Fisiológicas	81
	ANEXO B 2 –	82

1 INTRODUÇÃO

O Brasil é o segundo maior produtor e exportador mundial de soja, com uma produção média de 56,3 milhões (mi) de toneladas (ton) (CONAB, 2010). Apesar da alta produtividade desta cultura, a ocorrência de estresses, como o déficit hídrico, pode reduzir significativamente os rendimentos das lavouras, restringindo também as latitudes e os solos que podem ser cultivados (NEPOMUCENO, 2001). Segundo levantamento da Companhia Nacional de Abastecimento (CONAB), a soja foi uma das culturas mais prejudicadas pelas adversidades climáticas dos últimos anos (CONAB, 2009). No Rio Grande do Sul a estiagem levou a uma redução média de 22% da produtividade na safra 2007/2008, gerando prejuízos em torno de R\$ 1,6 bilhão (EMBRAPA, 2008).

Considerando o real cenário de mudanças climáticas, que sugere um aumento na aridez de diversas áreas do globo terrestre (BATES et al., 2008), estudos sobre as respostas das plantas ao déficit hídrico são cada vez mais importantes. A ocorrência de estresses, como o déficit hídrico, leva a uma série de alterações fisiológicas, capazes de ativar complexos enzimáticos, que, por sua vez, desencadeiam cascatas de eventos moleculares de sinalização e resposta ao estresse, induzindo a expressão de várias categorias de genes envolvidos de alguma forma na defesa contra o estresse (NEPOMUCENO, 2001).

O advento de novas técnicas moleculares como sequenciamento em massa e o PCR quantitativo tem permitido a identificação de novos genes e rotas metabólicas relacionadas a diversos tipos de estresse, dentre eles o déficit hídrico. Neste contexto, muitos esforços têm sido empregados no intuito de melhorar a produção agrícola sob condições limitadas de água, dentre eles destacam-se o melhoramento genético e a seleção assistida, que têm buscado a obtenção de plantas com características morfo-fisiológicas favoráveis à tolerância ao déficit hídrico (MANAVALAN et al., 2009). Visando a identificação de novos genes, a construção de bibliotecas subtrativas (DIATCHENKO et al., 1996) tem sido uma importante ferramenta, permitindo elucidar os perfis de transcrição que embasam processos biológicos envolvidos na resposta a estresses bióticos (CHOI et al., 2008) e abióticos, incluindo o déficit hídrico (CLEMENT et al., 2008).

O presente trabalho se insere como parte de um grande projeto, o projeto GENOSoja, que visa integrar esforços dos principais grupos que trabalham

com pesquisas genômicas em soja no país, para mitigar danos causados por situações de estresses bióticos e abióticos. Como parte integrante do plano de ação Transcriptoma, este trabalho objetiva gerar informações sobre a expressão diferencial de genes em cultivares de soja submetidas ao déficit hídrico. Adicionalmente, em uma das linhas de pesquisa, nosso grupo estuda como os fatores de transcrição da família DREB atuam na tolerância das plantas de soja ao déficit hídrico. Diferentes genes envolvidos nesta cascata de sinalização estão sendo introduzidos na soja via engenharia genética. Dentre os genes identificados neste trabalho encontramos alguns potenciais candidatos a novos fatores de transcrição da família AP2/EREB, além de outros genes responsivos ao estresse. A obtenção destas informações contribuirá para a melhor compreensão do funcionamento do metabolismo vegetal sob condições limitantes de água, assim como os mecanismos de sensibilidade e tolerância ao déficit hídrico em plantas. Assim, mais estratégias de engenharia genética e seleção assistida por marcadores moleculares poderão ser utilizadas visando o desenvolvimento de genótipos de soja mais tolerantes ao déficit hídrico.

2 OBJETIVO

Identificar genes envolvidos nos mecanismos de resposta ao déficit hídrico em raízes de soja submetidas a diferentes níveis de estresse.

2.1 OBJETIVOS ESPECÍFICOS

- Identificar, através da construção de bibliotecas subtrativas, genes envolvidos nos mecanismos de resposta ao estresse, e *in silico*, identificar as possíveis proteínas codificadas por estes genes.
- Categorizar os genes diferencialmente expressos para compreender as alterações dos processos biológicos em decorrência do estresse e comparar as respostas das cultivares sensível e tolerante ao estresse.
- Buscar genes ainda não descritos na resposta ao déficit hídrico e estabelecer suas relações filogenéticas com sequências já descritas, assim como seu papel na resposta a este tipo de estresse.
- Validar os resultados obtidos nas bibliotecas subtrativas via PCR quantitativo, através da avaliação do perfil de expressão de alguns dos genes estresse-responsivos.

3 REVISÃO BIBLIOGRÁFICA

3.1 A SOJA

3.1.1 Origem e Dispersão

O cultivo da soja permaneceu restrito ao continente asiático por milhares de anos, só sendo introduzida na Europa no final do século XV, como curiosidade nos jardins botânicos da Inglaterra, França e Alemanha. Na segunda década do século XX, o teor de óleo e proteína do grão começou a despertar o interesse das indústrias mundiais (SHURTLEFF e AOYAGI, 2009a). No entanto, as tentativas de introdução comercial do cultivo do grão na Rússia, Inglaterra e Alemanha fracassaram, provavelmente, devido às condições climáticas desfavoráveis (EMBRAPA, 2010a). Nos Estados Unidos a produção tornou-se intensa após a Primeira Guerra Mundial, levando o país ao status de maior produtor e exportador mundial (SHURTLEFF e AOYAGI, 2009a). Diferentes autores discordam quanto ao centro de origem específico da soja atualmente cultivada, entretanto todos concordam que a área mais provável se localiza na região leste da Ásia. http://www.cnpso.embrapa.br/index.php?op_page=112&cod_pai=33

No Brasil, a soja foi introduzida na Bahia no ano de 1882. Nas décadas seguintes, ela foi utilizada como cultura experimental em algumas instituições de pesquisa e como planta hortícola entre os descendentes de imigrantes japoneses. Na década de 60, o trigo era a principal cultura do Sul do Brasil e a soja surgia como uma opção de verão, em sucessão ao trigo. Com o cultivo sucessivo de soja e trigo a produção de soja teve seu crescimento mais expressivo. A explosão do preço da soja no mercado mundial, em meados de 1970, despertou ainda mais os agricultores e o próprio governo brasileiro para o plantio desta cultura. A soja atualmente cultivada apresenta um grande número de cultivares adaptadas a diferentes regiões, fruto do intenso processo de melhoramento genético aplicado a esta cultura. (EMBRAPA, 2010b; SHURTLEFF e AOYAGI, 2009b).

3.1.2 Importância Econômica

A soja é uma cultura muito importante economicamente, pois é utilizada na alimentação humana e, principalmente, na alimentação animal, e tem sido cada vez mais aplicada na produção de biocombustíveis (FAO, 2008).

O óleo de soja é um dos produtos mais utilizados na alimentação humana e a sua participação no mercado mundial de óleos vegetais comestíveis é de 27,5% (SEDIYAMA et al., 2005). A proteína de soja é abundantemente empregada como complemento em diferentes formas de alimento com o objetivo de melhorar sua qualidade proteica, com a vantagem de ser mais barata que as proteínas de origem animal (MARTIN et al., 2010). Além das funções nutricionais básicas, atualmente a soja tem sido apontada por seus efeitos benéficos à saúde humana. Estudos indicam seu benefício em casos de diabetes (KWONA et al., 2010), hipertensão (NAKAHARA et al., 2010), sintomas do climatério e câncer de mama (BARNES, 2010).

Atualmente o Brasil é o segundo maior produtor mundial do grão com uma produção média de 56,3 milhões de toneladas (mi ton) nos últimos cinco anos, em uma área plantada que atinge 21,7 milhões de hectares (mi ha), (CONAB, 2010). O país ocupa também a posição de segundo maior exportador do produto, sendo que em 2009 exportou-se 28,7 milhões de toneladas de grãos, 16,5% a mais do que no ano anterior (CONAB, 2010).

A ocorrência de períodos de estiagem vem acarretando diminuições severas na produtividade desta cultura. A safra 2004/2005 foi uma das mais afetadas pela estiagem nos últimos anos, sendo que no Rio Grande do Sul, estado mais atingido, houve uma queda de 78% na produção da cultura em relação à safra anterior (FARIAS, NEUMAIER e NEPOMUCENO, 2009). Neste mesmo estado a estiagem levou a uma redução média de 22% na produtividade da safra 2007/2008, gerando prejuízos em torno de R\$ 1,6 bilhão (EMBRAPA, 2008). Segundo a FAO (2009), o aumento da produtividade agrícola é fundamental para o crescimento econômico e segurança alimentar. Neste panorama, estudos no intuito de melhorar a produtividade sob condições ambientais desfavoráveis, como o déficit hídrico, tornam-se de suma importância.

3.1.3 O Genoma da Soja

A soja é uma planta herbácea, pertencente à família das leguminosas. O genoma diplóide, que derivou de um ancestral alotetraplóide (GILL et al., 2009), é o maior genoma completo de planta já sequenciado, consistindo em 950 megabases (Mb), 85% do genoma de soja inicialmente predito (1.115 Mb) (SCHMUTZ et al., 2010). Está organizado em 20 cromossomos, de acordo com os 20 grupos de ligação de soja, e possui 46.430 genes preditos como codificadores de proteína, em uma estrutura gênica exon–intron altamente conservada com *Populus trichocarpa* e *Vitis vinifera*.

Outra característica marcante deste genoma é a presença de sequências altamente repetitivas, heterocromáticas com baixa recombinação (57% do genoma), e um grande número de elementos transponíveis, dos quais *long terminal repeat* (LTR) retrotransposons correspondem à classe mais abundante (42% do genoma) (SCHMUTZ et al., 2010).

A alta taxa de duplicação encontrada neste genoma (75% dos genes apresentam múltiplas cópias) pode ser explicada por dois eventos de duplicação, que ocorreram cerca de 59 e 13 mi de anos atrás. Entretanto, é importante salientar a importância destas duplicações para a evolução do genoma, pois possibilitou a diversificação e perda de genes, além de inúmeros eventos de rearranjos cromossômicos (McCLEAN et al., 2010).

3.2 DÉFICIT HÍDRICO

3.2.1 Visão Geral

O déficit hídrico, em conjunto com o aumento da temperatura e radiação, torna-se a maior adversidade ambiental para a sobrevivência das plantas e para a produtividade agrícola. Considerando o projetado cenário de mudanças climáticas, que sugerem um aumento na aridez de diversas áreas do globo terrestre (BATES et al., 2008), estudos sobre as respostas das plantas ao déficit hídrico são cada vez mais importantes.

Segundo trabalhos de NOBRE, ASSAD e OYAMA (2005), o pior impacto do déficit hídrico no Brasil é estimado para lavouras em solos arenosos,

onde, nos próximos cem anos, o aumento das temperaturas médias do planeta poderá alcançar até 5,8°C. Nessas condições, o cultivo de soja no Brasil diminuiria dos 340 mil ha de área potencialmente produtiva atualmente, para aproximadamente 120 mil ha, uma redução de 60%. Com o aumento de temperatura, a primeira consequência é o aumento nas taxas de evapotranspiração, favorecendo um maior consumo de água pelas plantas e, portanto, esgotando o recurso do “reservatório solo” mais rapidamente. A segunda consequência seria a redução do ciclo das culturas, tornando-as mais eficientes em termos de assimilação e transformação energética, porém mais sensíveis à deficiência hídrica (NOBRE, ASSAD e OYAMA, 2005).

A soja é considerada uma planta sensível ao déficit hídrico, cujo desenvolvimento pode ser dividido em dois grandes estádios: vegetativo (V) e reprodutivo (R) (FEHR e CAVINESS, 1977). A duração dos estádios é controlada por fatores genéticos e ambientais, principalmente fotoperíodo e temperatura. O estágio vegetativo compreende as fases VE até as fases Vn, onde VE representa a emergência dos cotilédones; V1 representa o surgimento das duas primeiras folhas unifolioladas e os demais estádios V2, V3, ...Vn são numerados de acordo com a presença e o desenvolvimento dos trifólios. O estágio reprodutivo engloba as fases de floração – estádios R1 e R2, formação de vagens - R3 e R4, e enchimento e maturação de grãos - R5 a R8. O período crítico de sensibilidade ao déficit hídrico em soja encontra-se, principalmente, durante o período de emergência, floração e enchimento de vagens. Desta maneira, a ocorrência de déficit hídrico durante estas etapas pode levar a diminuições drásticas na produtividade da cultura (PEDERSON, 2004).

Embora seja considerada sensível, a soja é capaz de desenvolver mecanismos de aclimação que permitem tolerar o déficit hídrico e conseqüentemente sobreviver a determinados períodos de estresse. Existem evidências de que a ocorrência de déficit hídrico em diferentes estádios do desenvolvimento pode aumentar sua eficiência em tolerar o déficit hídrico (KRON et al., 2008). Segundo os autores, a ocorrência de déficit hídrico no estágio vegetativo anterior ao florescimento e enchimento de grãos permitiu que as plantas desenvolvessem mecanismos de aclimação que resultaram na maior tolerância a um segundo período de falta de água, sendo a redução do crescimento provavelmente relacionada à manutenção de uma estratégia de conservação de

energia, além das limitações físico-químicas do processo de crescimento celular desencadeadas pela falta de água.

3.2.2 Mecanismos Morfo-Fisiológicos de Resposta ao Déficit Hídrico

Visto a capacidade de percepção do estresse, as plantas, quando submetidas ao déficit hídrico, apresentam mecanismos morfológicos e fisiológicos de aclimação que reduzem a perda de água para o ambiente e mantêm sua hidratação. Estas alterações ocorrem tanto nas raízes como na parte aérea da planta (CATTIVELLI et al., 2008).

A plasticidade das raízes é um fator crítico para a adaptação das plantas ao déficit hídrico. Isto porque nestas condições limitantes, a água disponível encontra-se, em camadas mais profundas do solo, tornando o desenvolvimento de raízes mais profundas uma importante adaptação das plantas ao déficit hídrico. Sabe-se que existe uma correlação positiva significativa entre a tolerância ao déficit hídrico em soja e fatores relacionados à raiz como peso seco, comprimento total, volume e número de raízes laterais (LIU et al., 2005b). Além dos fatores acima citados, a ocorrência de raízes fibrosas, como no caso da soja PI416937, conferiu maior tolerância ao déficit hídrico através da limitação da transpiração durante os períodos de maior demanda evaporativa do solo, além de possibilitar o crescimento em solos mais compactados (FLETCHER e SINCLAIR, 2007).

A fixação biológica de nitrogênio (N_2) nas raízes é um processo extremamente vulnerável ao déficit hídrico. Com o declínio da fixação de N_2 há uma redução no suprimento de N_2 da planta para síntese proteica dificultando a manutenção do desenvolvimento de células e tecidos, acarretando uma diminuição na formação de grãos, implicando diretamente na produtividade. O impacto negativo do déficit hídrico na fixação biológica de N_2 fica claro em experimentos nos quais, com a adição de fertilizantes nitrogenados no solo sob déficit hídrico, foi eliminada a dependência de fixação de N_2 , e houve um aumento de 15 a 20% na produtividade em relação aos tratamentos em que não houve adição do fertilizante (RAY et al., 2006). Atualmente, os estudos para identificação e desenvolvimento de cultivares de soja com fixação de N_2 menos sensível ao déficit hídrico objetivam a expansão das zonas agricultáveis a regiões sujeitas ao déficit hídrico, mantendo-se a produtividade

sem a necessidade da utilização de fertilizantes nitrogenados (SINCLAIR et al., 2007).

Características relacionadas às raízes demonstram um grande potencial para aumentar a tolerância ao déficit hídrico em soja. Entretanto, a seleção de cultivares tolerantes baseada em características fenotípicas das raízes é difícil. Por outro lado, o emprego de marcadores moleculares possibilita o melhoramento genético para esta característica (CATTIVELLI et al., 2008). Outra estratégia interessante é a busca por genes candidatos para maior tolerância ao déficit hídrico, abordagem que foi utilizada com sucesso na identificação de genes para número de raízes em arroz submetido à baixa disponibilidade de água (VINOD et al., 2006). Os genes identificados podem ser aplicados em estratégias de seleção assistida por marcadores moleculares e engenharia genética, visando o desenvolvimento de genótipos mais tolerantes ao déficit hídrico.

Desse modo, entender os mecanismos morfológicos, fisiológicos e genéticos de plasticidade das raízes ajudará a identificar genes e rotas metabólicas específicas relacionadas ao déficit hídrico, fornecendo subsídios tanto para a seleção assistida por marcadores como para o desenvolvimento de cultivares mais tolerantes através da engenharia genética (MANAVALAN et al., 2009).

Com referência às modificações da parte aérea da planta à condição de déficit hídrico, um dos fatores mais importantes é a condutância estomática, visto que ela é um evento chave no controle da troca de gases e vapor de água entre a folha e o ambiente (TAIZ, 2010). Uma das primeiras respostas ao déficit hídrico é o fechamento estomático. O fechamento estomático pode ocorrer por meio de sinais hidráulicos (diminuição do turgor celular) ou químicos (ação do hormônio ácido abscísico – ABA). Autores relatam um decréscimo na condutância estomática coincidente com o aumento de ABA no xilema, mesmo antes de ocorrerem mudanças significativas no turgor celular, indicando que o ABA produzido nas raízes ascende via xilema e está envolvido no fechamento estomático (LIU et al., 2005a). Sua ação é mediada por meio da formação de complexos ternários com proteínas START da família PIR/PYL/RCAR e fosfatases tipo 2C (PP2C), inibindo a atividade das últimas (MELCHER, 2009). Estudos sugerem ainda que ambos os sinais (químicos e hidráulicos) podem agir juntos ou em tempos separados no fechamento estomático (KALEFETOGLU e EKMEKÇI, 2005). Este mecanismo reduz a perda de água, contribuindo para a manutenção do conteúdo de água na célula, mas limita a

difusão de CO_2 para os cloroplastos, levando à diminuição da atividade fotossintética. Desta maneira o controle da abertura estomática é essencial para manter um balanço ótimo entre o potencial hídrico adequado e as funções energéticas da planta (FLEXAS et al., 2006).

Outra importante característica relacionada à parte aérea da planta é a condutância epidermal. A condutância total da folha é uma soma da condutância estomática e cuticular. Quando os estômatos estão abertos a condutância cuticular é uma fração irrelevante de condutância total, mas quando o déficit hídrico leva ao fechamento estomático, o componente cuticular da condutância epidermal pode exceder a condutância estomática. Culturas com baixa condutância epidermal têm mostrado maior tolerância à deficiência hídrica, sendo que genótipos de soja adaptados a ambientes tropicais possuem menor condutância epidermal quando comparados a genótipos de regiões temperadas, sugerindo que esta pode ser uma adaptação dos genótipos tropicais a ambientes mais secos (JAMES et al. 2008). Em 2007, Hufstetler e colaboradores descreveram uma correlação negativa entre condutância epidermal e a eficiência no uso de água (EUA), o que poderia indicar a importância desta característica para a tolerância ao déficit hídrico.

A eficiência do uso de água é definida como a quantidade de biomassa acumulada por unidade de água utilizada. Nos últimos anos, associações positivas entre a EUA e a biomassa total têm sido feitas, sugerindo que aumentos na EUA resultem em uma melhor produtividade (WRIGHT, 1996). Entretanto, mais recentemente, argumenta-se que a seleção ou melhoramento genético para aumento do EUA, na maioria das vezes, leva à redução da produtividade e tolerância ao déficit hídrico (BLUM, 2009). Isto visto que a bioquímica da fotossíntese não pode ser melhorada geneticamente (PARRY et al., 2007), logo, maiores taxas de EUA podem ser alcançados apenas através da redução da transpiração e uso de água da planta, processos cruciais para a produtividade e tolerância sob condições de déficit hídrico (BLUM, 2009).

O ajuste osmótico ocorre pelo acúmulo ativo de solutos na célula vegetal, propiciando, em várias espécies, o aumento na capacidade de tolerar períodos curtos de déficit hídrico através da diminuição do potencial osmótico e da retenção de água dele decorrente. Isto possibilita que a planta mantenha a absorção de água e a pressão de turgor celular, contribuindo para sustentar a taxa fotossintética e o crescimento celular por expansão (CATTIVELLI et al., 2008).

Diversas moléculas já foram descritas no ajuste osmótico como aminoácidos (prolina e ectoína), polióis (manitol, sorbitol e D-ononitol) e açúcares (trehalose) (ALMEIDA et al., 2007). Além de contribuir para o turgor celular e absorção de água, as moléculas osmoprotetoras participam na proteção e estabilização de macromoléculas e estruturas celulares (proteínas, membranas, cloroplastos) contra o dano desencadeado pelo estresse osmótico (MARTÍNEZ et al., 2004).

A considerável variação da capacidade de ajuste osmótico entre diferentes culturas pode ser um indício da habilidade da cultura em suportar o déficit hídrico. A soja normalmente apresenta baixa capacidade de ajuste osmótico. Apesar disso, um estudo identificou genótipos com menor declínio de teor relativo de água (TRA) devido a um maior potencial osmótico, o qual ocorre em função do ajuste osmótico, evidenciando a habilidade destes genótipos em manter o turgor celular durante o déficit hídrico (JAMES et al., 2008).

A densidade de tricomas nas folhas é uma característica adaptativa relacionada às plantas xerófilas e que, em soja, tem sido relacionada à tolerância ao déficit hídrico. A presença de tricomas nas folhas aumenta a reflectância da folha e reduz a temperatura foliar, além de restringir a perda de água por transpiração, otimizando a fotossíntese (DU et al., 2009).

3.2.3 Respostas Moleculares ao Déficit Hídrico

Sabe-se que todas as alterações morfo-fisiológicas estresse-induzidas possuem uma base molecular. A primeira etapa na resposta molecular ao déficit hídrico é a percepção do estresse via receptores específicos. Uma vez ativados, estes receptores iniciam (ou suprimem) uma cascata de eventos moleculares para transmitir a informação através de complexas vias de transdução de sinais (SEKI et al., 2007). A perda de água causa uma diminuição no turgor celular, levando à percepção mecânica do estresse (KALEFETOGLU e EKMEKÇI, 2005). Em *Arabidopsis thaliana* foram identificadas proteínas receptoras histidina quinase, codificadas pelos genes *AHK1/ATHK1*, *AHK2*, *AHK3*, e receptores de citocininas, como o codificado pelo gene *CRE1*. Estas proteínas possuem domínios receptores e transmembranas e são capazes de perceber as mudanças no potencial osmótico da célula em resposta ao déficit hídrico e osmótico. Estudos de repressão

ou superexpressão dos genes *AHK1* demonstram que ele é um regulador positivo de genes que atuam em resposta ao déficit hídrico, estresse salino e ao ABA (TRAN et al., 2007).

O gene *NtC7* também tem sido descrito como sensível às alterações osmóticas em plantas (NAKAMURA et al., 2009). A superexpressão deste gene, que codifica uma proteína similar a um receptor de membrana, levou ao aumento da tolerância ao estresse osmótico, sugerindo um importante papel na sinalização deste tipo de estresse (NAKAMURA et al., 2009). O próximo passo na transdução do sinal, após a percepção por osmosensores, é a ativação de cascatas de fosforilação, envolvendo a participação de MAPKs (*Mitogen-Activated Protein Kinase*). A rota das MAPKs consiste numa cascata de proteínas quinases que fosforilam de maneira ordenada, a partir de uma *MAPK kinase kinase* (MAPKKK) para uma *MAPK kinase* (MAPKK), e então para uma MAPK (ZHI-MING et al., 2006). Em plantas, um grande número de MAPKs tem sido identificado. Sabe-se que o genoma de *Arabidopsis* codifica aproximadamente 60 MAPKKKs, 10 MAPKKs e 20 MAPKs. Foram também identificados 20 MAPKs, 10 MAPKKs e 60 MAPKKKs possíveis ortólogos destes genes em *Medicago truncatula*, tabaco (*Nicotiana tabacum*) e arroz (*Oryza sativa*) (NAKAGAMI et al., 2005), respectivamente.

Embora a função da cascata de fosforilação em resposta a estresses ainda não esteja totalmente esclarecida, sabe-se que os membros da cascata de MAPKs são ativados por mais de um tipo de estresse, o que sugere que esta via atue como ponto de convergência na sinalização de estresses (YOSHIOKA e SHINOZAKI, 2009). Em algodão o gene GhMAPK foi superexpresso em resposta a diversos estresses bióticos e abióticos tais como frio, salinidade, injúria, ácido salicílico, peróxido de hidrogênio e ataque de patógenos (WANG et al., 2007). Em relação ao déficit hídrico, recentemente uma nova MAPKKK (DSM1) reguladora da resposta de eliminação de espécies reativas de oxigênio (ERO) foi descrita em arroz (NING et al., 2010). Em soja, diversos tipos de MAPK que têm sido descritos, denominados como *GmMPKs*, já foram relatados em mecanismos de defesa contra patógenos (DAXBERGER et al., 2007).

Outra importante classe de sinalizadores são os derivados de fosfolipídeos. Sabe-se que a estrutura fluida da membrana plasmática é consequência do ambiente aquoso da célula, onde temos as caudas hidrofóbicas dos fosfolipídeos sendo repelidas pela água, formando a bicamada lipídica. Quando

a água deixa a célula, esta estrutura é alterada (KALEFETOGLU e EKMEKÇI, 2005). Hipóteses, atualmente aceitas, indicam que os fosfolípidos presentes na membrana submetida ao déficit hídrico são clivados por fosfolipases, produzindo derivados que agem como mensageiros secundários. Os principais derivados de fosfolípidos relacionados ao estresse osmótico são o inositol 1,4,5-trifosfato (IP3), diacilglicerol (DAG) e ácido fosfatídico (PA) (BARTELS e SUNKAR, 2005).

Ainda com relação ao processo de sinalização, sabe-se que estresses abióticos, como o déficit hídrico, levam ao acúmulo de hormônio ácido abscísico - ABA. A sinalização do estresse em plantas tem sido dividida em vias dependentes e independentes de ABA. Na sinalização dependente de ABA este hormônio liga-se a proteínas START, que são receptores protéicos, denominados PYR/PYL/RCARs. Esta recém descoberta classe de proteínas START inibe a ação de fosfatases, permitindo a ativação de SnRK2 quinases que fosforilam efetores *downstream*, como *basic leucine zipper transcription factors* chamados ABFs/AREBs (PARK, et al., 2009). Estes efetores, por sua vez, ligam-se aos elementos-cis ABRE para induzir a expressão gênica em resposta ao déficit hídrico (MIZOGUCHI et al., 2010).

Na via de sinalização mediante ABA, o cálcio citosólico (Ca^{+2}) atua como um dos principais mensageiros secundários. Esta via ocorre mediante rápidas mudanças na concentração do cálcio citosólico, desencadeadas pela participação dos canais de cálcio, que regulam os níveis deste elemento (SONG et al., 2008). Em 2005, o gene *TaTPC1* (*two-pore channel 1*), codificante de um canal de cálcio voltagem-dependente, foi clonado de trigo (*Triticum aestivum*) e sua superexpressão em plantas transgênicas de *Arabidopsis* acelerou o fechamento estomático na presença de cálcio (WANG et al., 2005).

Como molécula sinalizadora, o cálcio está estreitamente relacionado com diversos estresses abióticos, dentre eles o déficit hídrico (SONG et al., 2008). A rede de sinalização via cálcio é regulada por proteínas decodificadoras de sinais de cálcio. Três classes de moléculas sensíveis a alterações na concentração de cálcio foram caracterizadas em plantas: calmodulinas, CDPKs (calcium-dependent protein kinase) e CBLs (calcineurin B-like proteins) (BATISTIČ e KUDLA, 2009). CDPKs são proteínas quinases que possuem um domínio auto-regulatório ao qual se liga o cálcio. Estas proteínas são amplamente descritas na sinalização via ABA e, em alguns casos, genes codificantes de CDPKs possuem elementos ABA-responsivos

em seus promotores, mostrando-se induzidos em nível transcricional na presença de ABA (HEY et al., 2010). As proteínas CBL foram inicialmente descritas em *Arabidopsis thaliana*. Elas interagem com CDPKs (CBL-interacting protein kinases) formando uma rede interligada de sinalização. Estudos com mutantes para genes de CBL e CDPK revelaram a importância destas proteínas na regulação das respostas das plantas a diferentes tipos de estresses (BATISTIC e KUDLA, 2004). Análises do genoma de *Arabidopsis* e arroz identificaram 10 CBLs em ambos os genomas, e 25 (*Arabidopsis*) e 30 (arroz) CDPKs (BATISTIČ e KUDLA, 2009).

Entretanto, existem também vias de sinalização independentes de ABA para ativação da expressão gênica induzida pelo déficit hídrico. Nesta via a expressão gênica é regulada através de elementos-*cis* DRE (do inglês, Dehydration Responsive Element). O elemento responsivo à desidratação (DRE) foi identificado como importante elemento *cis*-atuante na regulação da expressão gênica em resposta ao déficit hídrico, alta salinidade e frio (YAMAGUCHI-SHINOZAKI e SHINOZAKI, 2005).

Durante estresses abióticos, fatores de transcrição do tipo AP2/EREBP ativam os elementos *cis*-atuantes DRE de genes estresse-induzidos (CHINNUSAMY, SCHUMAKER e ZHU, 2004). Estes fatores de transcrição foram isolados e divididos em *CBF/DREB1* e *DREB2* (YAMAGUCHI-SHINOZAKI e SHINOZAKI, 2005). Em uma análise *in silico*, WANG e seus colaboradores (2009) identificaram 160 genes responsivos a estresses abióticos, alvo dos fatores de transcrição DREB em *Arabidopsis thaliana* (WANG, et al.). Os genes *DREB1* e *DREB2* desempenham importante papel na via de resposta independente de ABA, sendo que os genes *DREB1* foram inicialmente identificados na resposta ao frio, sendo mais tarde também descritos na resposta à seca e alta salinidade. Os genes *DREB2* têm sido descritos principalmente na resposta ao déficit hídrico e estresse osmótico (YAMAGUCHI-SHINOZAKI e SHINOZAKI, 2005). A superexpressão dos genes *DREB* em diferentes espécies de plantas como girassol, arroz e milho, aumentou a tolerância das plantas transgênicas a estresses abióticos como altas temperaturas e déficit hídrico (ALMOGUERA et al., 2009; OH et al., 2009; QIN et al., 2004). Em soja, análises de ESTs (Expressed Sequence Tags) indicam a presença de 290 membros da família AP2/EREBP (TIAN et al., 2004), sendo que já foram isolados três genes homólogos de DREB: *GmDREBa*, *GmDREBb*, e *GmDREBc* (LI et al., 2005). Em tabaco, a superexpressão do gene endógeno de soja

GmDREB2 conferiu maior tolerância ao déficit hídrico e estresse salino nas plantas transgênicas do que o verificado em plantas não transformadas (CHEN et al., 2007).

3.2.4 Técnicas Moleculares e Melhoramento Genético

Diversas características têm sido associadas à capacidade de adaptação das plantas a ambientes de baixa disponibilidade de água, dentre elas, redução no tamanho da planta, redução da área foliar, maturidade precoce, ajuste osmótico, retardo da senescência e fechamento estomático prolongado (CATTIVELLI et al., 2008). Com base neste conhecimento, nos últimos anos, muitos estudos têm sido desenvolvidos no intuito de melhorar a produção agrícola sob condições limitadas de água, buscando plantas com características de maior tolerância ao déficit hídrico (MANAVALAN et al., 2009).

A disponibilidade do genoma completo da soja (SCHMUTZ et al., 2010), além de mapas físicos e genéticos (SONG et al., 2004), e as ferramentas de genômica funcional, integrados com estratégias de melhoramento e engenharia genética, trazem novas oportunidades para a pesquisa da tolerância ao déficit hídrico em soja. O fato do RNA ser diretamente transcrito do DNA tem levado pesquisadores a analisar os níveis de transcritos para entender a expressão gênica induzida pelo estresse. Embora ferramentas convencionais como *Northern blotting* (ALWINE et al., 1977) e RT-PCR (do inglês, Reverse Transcription-Polymerase Chain Reaction) (LEE et al., 1989) tenham sido muito utilizadas, elas são capazes de analisar poucos genes em um experimento. O uso de técnicas que possibilitam uma análise global da expressão gênica, como a técnica de SAGE (do inglês, *Serial Analysis Gene Expression* - VELCULESCU et al., 1995), microarranjos de DNA (SCHENA et al., 1995), ou ainda a construção de bibliotecas de cDNA, associadas aos sequenciamentos em larga escala, têm possibilitado uma visão ampla das respostas moleculares ao déficit hídrico, visto que possibilitam a identificação de diversos genes expressos em resposta ao estresse.

A construção de Bibliotecas de cDNA e o sequenciamento dos clones obtidos tem sido muito utilizado para a descoberta de novos genes (JEONG et al., 2007; WANG et al., 2008; NARAYANAN et al., 2009). Mais recentemente a associação da técnica de Supressão Subtrativa à construção de bibliotecas de cDNA, (DIATCHENKO et al. 1996) tornou possível identificar uma variedade de

transcritos diferencialmente expressos em diferentes tecidos e/ou condições específicas (CLEMENT et al., 2008; CHOI et al., 2008; ASLAM, et al., 2010). A obtenção de bibliotecas subtrativas visa eliminar os transcritos expressos tanto na situação de tratamento/estresse quanto em uma situação controle, permanecendo na biblioteca apenas os transcritos expressos diferencialmente pela imposição do tratamento. Esta subtração é feita através do anelamento de um *pool* de cDNAs da situação de interesse (*tester*) com um excesso de cDNAs da situação controle (*driver*). Desta maneira os transcritos presentes na amostra *driver* irão hibridizar com os transcritos em comum ao *tester* e serão eliminados durante a síntese da biblioteca, permanecendo apenas os transcritos dos *tester* considerados diferencialmente expressos. Um passo importante na construção desse tipo de biblioteca é a amplificação do cDNA diferencialmente expresso por meio de reações de PCR, o que possibilita a identificação de transcritos raros (DIATCHENKO et al. 1996).

Os genes diferencialmente expressos, identificados através da construção de bibliotecas subtrativas, por exemplo, são candidatos interessantes à tolerância ao déficit hídrico, podendo ser utilizados para o desenvolvimento de cultivares mais tolerantes ao déficit hídrico, seja através da Seleção Assistida por Marcadores Moleculares (SAM) (RIBAUT, et al. 2002), ou da transformação genética, a qual tem sido apontada como maior inovação no desenvolvimento de linhagens tolerantes a este tipo de estresse (MANAVALAN et al., 2009). Inúmeros trabalhos reportam o aumento da tolerância ao déficit hídrico por meio da alteração da expressão de genes candidatos, via engenharia genética. Em arroz e tabaco, por exemplo, através da transformação genética com o gene *SDR1*, envolvido na rota de sinalização do déficit hídrico via ácido abscísico (ABA), obtiveram-se plantas com maior tolerância ao déficit hídrico (ZHANG et al., 2008). Por sua vez, a superexpressão de um gene codificante de uma chaperona BiP (*soyBiPD*) em plantas de soja e tabaco resultou no retardo da senescência foliar durante o déficit hídrico. Verificou-se também uma menor expressão de genes relacionados ao déficit hídrico nas plantas de soja transformadas com este gene, indicando que estas estariam sentindo menos os efeitos do déficit hídrico (VALENTE et al., 2009).

Estudos com genes candidatos endógenos de soja também têm obtido resultados positivos quanto à tolerância ao déficit hídrico. Em plantas de *Arabidopsis thaliana* e *Nicotiana tabacum*, a superexpressão de um gene isolado de

soja (*GmDREB2*) possibilitou o aumento da tolerância ao déficit hídrico e à alta salinidade (CHEN et al., 2007). Em 2009, ZHANG e seus colaboradores obtiveram plantas de tabaco mais tolerantes à salinidade e déficit hídrico, através da transformação com o gene *GmERF3* de soja. De maneira semelhante, plantas de *Arabidopsis thaliana* transformadas com o gene endógeno de soja, *GmCHI*, também apresentaram maior tolerância à estresses abióticos como frio, déficit hídrico e salinidade (CHENG, et al., 2009).

3.2.5 Sequenciamento de DNA de Última Geração (NGS)

Os sequenciamentos de última geração (do inglês, *Next Generation Sequencing*) apresentam grande potencial para estudos de genômica estrutural, como na identificação de variações do genoma, e genômica funcional, através do sequenciamento de bibliotecas de cDNA, visto o aumento da sensibilidade da técnica e a rapidez na obtenção dos resultados. Dentre estas tecnologias comercialmente disponíveis estão o pirosequenciamento (Sequenciador 454, Roche), e os sequenciamentos de fragmentos curtos, como o SOLiD (Applied Biosystems), *Genome Analyzer IIe* (Illumina) e Heliscope (Helicos). As principais diferenças entre estes tipos de sequenciamento encontram-se no tipo de amplificação e métodos de detecção utilizados. As técnicas de sequenciamento tradicionais, baseadas em Sanger (SANGER et al., 1977), requerem a clonagem dos fragmentos dentro de vetores, onde os fragmentos são amplificados previamente à etapa de sequenciamento, o qual é realizado utilizando *primers* baseados na sequência do vetor. Já os sequenciamentos de última geração dispensam tais etapas de clonagem *in vivo*, resultando em economia de custos, tempo e mão de obra.

O sequenciamento por 454 utiliza a amplificação *in vitro* por meio de PCR em emulsão. Neste tipo de PCR, fragmentos individuais de DNA ou cDNA são ligados a pequenas esferas (*beads*) e são separadamente amplificados em uma gota de emulsão, que age como um reator de amplificação, gerando cerca de 10^7 cópias idênticas de cada sequência. A forma de detecção das bases fundamenta-se na reação com o pirofosfato liberado durante a incorporação de nucleotídeos, que, juntamente com a enzima luciferase, gera uma reação de quimio-luminescência (RONAGHI, 1996). De maneira similar a tecnologia SOLiD (síntese por ligação)

também utiliza a PCR em emulsão na clonagem dos fragmentos para o sequenciamento. Os produtos da amplificação são transferidos para uma superfície sólida e o sequenciamento ocorre por etapas sucessivas de hibridização e ligação de dinucleotídeos marcados com diferentes fluoróforos, utilizados para a detecção.

A tecnologia de sequenciamento desenvolvida pela Illumina, também dispensa as etapas de clonagem *in vivo*, utilizando PCR de moléculas ligadas a uma superfície sólida, que consiste em uma lâmina de vidro contendo 8 canais. Cada amostra, como por exemplo, uma única biblioteca de cDNA, pode ser adicionada a cada um destes canais, ou várias bibliotecas podem ser adicionadas em cada canal, desde que contenham adaptadores específicos para cada biblioteca, que funcionam como “código de barras” para diferenciá-las. Inicialmente, uma das extremidades da molécula de DNA liga-se a adaptadores contidos na lâmina, e durante o processo de amplificação a molécula dobra-se para que ocorra, por complementariedade, a ligação do adaptador da outra extremidade com o correspondente adaptador presente na lâmina, em um processo chamado de amplificação “em ponte”. Esta amplificação gera aproximadamente 40 milhões de *clusters*, cada um representante de uma única sequência, contendo aproximadamente um milhão de cópias, suficientes para a detecção do sinal de incorporação de bases durante o sequenciamento. Durante este processo são adicionados, simultaneamente, nucleotídeos referentes às quatro bases, cada um marcado com um tipo de fluoróforo com a extremidade 3’OH quimicamente bloqueada. Após cada etapa de incorporação, o bloqueio no grupo 3’OH é removido quimicamente para permitir a ligação da próxima base no ciclo seguinte. A cada etapa são capturadas imagens que posteriormente são sobrepostas para obtenção da sequência completa. O equipamento *Genome Analyzer IIe* (Illumina), é capaz de produzir até dois bilhões de reads por corrida, gerando até 25 gigabases por dia (ILLUMINA, 2010).

A tecnologia Helicos (MILOS, 2008), diferentemente das anteriores, não requer etapas de amplificação, realizando o sequenciamento de moléculas únicas, denominado *True Single Molecule Sequencing (tSMS)*[™]. Esta tecnologia permite o sequenciamento simultâneo de um grande número de moléculas de DNA ou mRNA utilizando o método de síntese com bases marcadas com fluoróforos, os quais são adicionados sequencialmente ao longo do processo. As imagens são capturadas e analisadas por um programa especialmente desenvolvido, o *HeliScope*[™] *Analysis Engine*.

As tecnologias acima abordadas são as mais comumente utilizadas e já estão disponíveis comercialmente, entretanto a tecnologia de sequenciamento continua a evoluir no intuito de desenvolver metodologias mais rápidas, precisas e com menor custo. Dentre as tecnologias em desenvolvimento está o sequenciamento por nanoporos, que estima ser capaz de sequenciar o genoma de mamíferos a um custo de aproximadamente \$1.000 em cerca de 24 h (BRANTON et al., 2008).

3.3 REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, A.M. et al. Trehalose and its applications in plant biotechnology. **In Vitro Cellular & Developmental Biology - Plant**, v.43, p.167-177, 2007.
- ALMOGUERA, et al. The HaDREB2 transcription factor enhances basal thermotolerance and longevity of seeds through functional interaction with HaHSFA9. **BMC Plant Biology**, v.9, p.75, 2009.
- ALWINE, J.C., KEMP, D.J., STARK, G.R. Method for detection of specific RNAs in agarose gels by transfer to diazobenzoyloxymethylpaper and hybridization with DNA probes. **Proceedings Of The National Academy Of Sciences Of The U.S.A.**, v.74, p.5350-5354, 1977.
- ASLAM, M. et al. Isolation of cold stress-responsive genes from *Lepidium latifolium* by suppressive subtraction hybridization. **Acta Physiologiae Plantarum**, v.32, p.205-210, 2010.
- BARNES, S. The biochemistry, chemistry and physiology of the isoflavones in soybeans and their food products. **Lymphatic Research and Biology**, v.8, p.89-98, 2010.
- BARTELS, D., SUNKAR, R. Drought and salt tolerance in plants. **Critical Reviews in Plant Sciences**, v.24, p.23-58, 2005.
- BATES, B.C. et al. **Climate Change and Water**. IPCC/ONU, Technical Paper VI, Eds. IPCC Secretaria, Geneva, p.210, 2008.
- BATISTIC, O., KUDLA, J. Integration and channeling of calcium signaling through the CBL calcium sensor/CIPK protein kinase network. **Planta**, v.219, p.915-924, 2004.
- BATISTIČ, O., KUDLA, J. Plant calcineurin B-like proteins and their interacting protein kinases. **Biochimica et Biophysica Acta**, v.1793, p.985-992, 2009.
- BLUM, A. Effective use of water (EUW) and not water-use efficiency (WUE) is the target of crop yield improvement under drought stress. **Field Crops Research**, v.112, p.119-123, 2009.
- BRANTON, D. et al. The potential and challenges of nanopore sequencing. **Nature biotechnology**, v.26, n.10, p.1146-1153, 2008.
- CATTIVELLI, L. et al. Drought tolerance improvement in crop plants: An integrated view from breeding to genomics. **Field Crops Research**, v.105, p.1-14, 2008.
- CHEN, M. et al. GmDREB2, a soybean DRE-binding transcription factor, conferred drought and high-salt tolerance in transgenic plants. **Biochemical and Biophysical Research Communications**, v.353, p.299-305, 2007.
- CHENG, L. et al. GMCHI, cloned from soybean [*Glycine max* (L.) Meer.] enhances survival in transgenic *Arabidopsis* under abiotic stress. **Plant Cell Reports**, v.28, p.145-153, 2009.
- CHINNUSAMY, V., SCHUMAKER, K., ZHU, J. Molecular genetic perspectives on cross-talk and specificity in abiotic stress signaling in plants. **Journal of Experimental Botany**, v.55, p.225-236, 2004.
- CHOI, J. J. et al. Expression patterns in soybean resistant to *Phakopsora pachyrhizi* reveal the importance of peroxidases and lipoxygenases. **Functional & Integrative Genomics**, v.8, p.341-359, 2008.
- CLEMENT, M. et al. Identification of new up-regulated genes under drought stress in soybean nodules. **Gene**, v.426, p.15-22, 2008.
- CONAB, 2009. Disponível em: <<http://www.conab.gov.br/conabweb/>> Acesso em: 22/04/09.

- CONAB, 2010. Acompanhamento de safra brasileira: grãos, sexto levantamento, março 2010. Disponível em <http://www.conab.gov.br/conabweb/download/safra/6graos_09.03.10.pdf> Acesso em 31/03/2010, 15:40h, 2010.
- DAXBERGER, A. et al. Activation of members of a MAPK module in glucan elicitor-mediated non-host resistance of soybean. **Planta**, v.225, p.1559-1571, 2007.
- DIATCHENKO et al. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. **Proceedings Of The National Academy Of Sciences Of The U.S.A.**, v.93, p.6025-6030, 1996.
- DU, W.J., FU, S.X., YU, D.Y. Genetic analysis for the leaf pubescence density and water status traits in soybean [*Glycine max* (L.) Merr.]. **Plant Breeding**, v.128, p.259-265, 2009.
- EMBRAPA, 2010^a. Disponível em: <http://www.cnpso.embrapa.br/index.php?op_page=112&cod_pai=33> Acesso em 31/03/2010, 14:56h.
- EMBRAPA, 2010^b. Disponível em: <http://www.cnpso.embrapa.br/index.php?op_page=113&cod_pai=35> Acesso em: 31/03/2010, 15:45h.
- EMBRAPA, 2008. Disponível em: <<http://www.embrapa.br/imprensa/noticias/2008/junho/2a-semana/levantamento-quantifica-perdas-na-soja-no-rs/>> Acesso em: 17/08/2009.
- FAO, 2008. Disponível em: <<https://www.fao.org.br/default.asp>> Acesso em Dez 2008.
- FAO, 2009. **El Estado Mundial De La Agricultura Y La Alimentación**. Disponível em: <<https://www.fao.org.br/download/i0680s.pdf>> Acesso em: 31/03/2010, 16:26
- FARIAS, J. R. B.; NEUMAIER, N., NEPOMUCENO, A. L., 2009. **Seca: soja em carência de água**. Disponível em: <<http://bioinfo.cnpso.embrapa.br/seca/index.php/ecofisiologia/exigencias-climaticas>> Acesso em 01/02/2011.
- FEHR, W.R., CAVINESS, C.E. **Stages of soybean development**. Ames: Iowa State University, 1977.
- FLETCHER, A.L., SINCLAIR, T.R. Transpiration responses to vapor pressure deficit in well watered 'slow-wilting' and commercial soybean. **Environmental Experimental Botany**, 2007.
- FLEXAS, J. et al. Keeping a positive carbon balance under adverse conditions: responses of photosynthesis and respiration to water stress. **Physiologia Plantarum**, 2006.
- GILL, N. et al. Molecular and chromosomal evidence for allopolyploidy in soybean *Glycine max* (L.) Merr. **Plant Physiology**, v.151, p.1167-1174, 2009.
- HEY, S. J., BYRNE, E., HALFORD, N.G. The interface between metabolic and stress signaling. **Annals of Botany**, v.105, p.197-203, 2010.
- HUFSTETLER, E.V. et al. Genotypic variation for three physiological traits affecting drought tolerance in soybean. **Crop Science**, v.47, p.25-35, 2007.
- ILLUMINA 2010. Disponível em: <http://www.illumina.com/technology/solexa_technology.ilmn> Acesso em: 23/09/2010.
- JAMES A. T., LAWN R. J., COOPER, M. Genotypic variation for drought stress response traits in soybean. I. Variation in soybean and wild *Glycine* spp. for epidermal conductance, osmotic potential and relative water content. **Australian journal of agricultural research**, v.59, p.656-669, 2008.
- JEONG, J. C. et al. Isolation and characterization of a novel Calcium/Calmodulin-Dependent protein kinase, *AtCK*, from *Arabidopsis*. **Molecules And Cells**, v.24, p.276-282, 2007.
- KALEFETOGLU, T., EKMEKÇI, Y. The affects of drought on plants and tolerance mechanisms. **Journal of Science**, v.18, p.723-740, 2005.
- KRON, A.P., SOUZA, G.M. e RIBEIRO, R.V. Water deficiency at different developmental stages of *Glycine max* can improve drought tolerance. **Bragantia**, v.67, p.43-49, 2008.
- KWONA, D.Y. et al. Antidiabetic effects of fermented soybean products on type 2 diabetes. **Nutrition Research**, v.30, p.13, 2010.
- LEE, M.S. et al. Detection of two alternative bcr/abl mRNA junctions and minimal residual disease in Philadelphia chromosome positive chronic myelogenous leukemia by polymerase chain reaction. **Blood**, v.73, p.2165-2170, 1989.
- LI, X.P. et al. Soybean DRE-binding transcription factors those are responsive to abiotic stress. **Theoretical Applied Genetics**, v.110, p.1355-1362, 2005.
- LIU, F. et al. ABA regulated stomatal control and photosynthetic water use efficiency of potato (*Solanum tuberosum* L.) during progressive soil drying. **Plant Science**, v.168, p.831-836, 2005a.
- LIU, Y. et al. Identification of drought tolerant germplasma and inheritance and QTL mapping of related roots traits in soybean (*Glycine max* (L.) Merr.). **Acta Genetica Sinica**, v.32, p.855- 863, 2005b.

- MANAVALAN, L.P. et al. Physiological and molecular approaches to improve drought resistance in soybean. **Plant Cell Physiology**, v.50, p.1260-1276, 2009.
- MARTIN, H., LASWAI, H., KULWA, K. Nutrient content and acceptability of soybean based complementary food. **Journal of Food Agriculture Nutrition and Development**, v.10, 2010.
- MARTÍNEZ, J.P. et al. Is osmotic adjustment required for water stress resistance in the Mediterranean shrub *Atriplex halimus* L.? **Journal Plant Physiology**, v.161, p.1041-1051, 2004.
- McCLEAN, P.E. et al. Synteny mapping between common bean and soybean reveals extensive blocks of shared loci. **BMC Genomics**, v.11, p.184, 2010.
- MELCHER, K., et al. Agate-latch-lock mechanism for hormone signalling by abscisic acid receptors. **Nature**, v.462, 2009.
- MILOS, P. Company Profile: Helicos BioSciences. **Pharmacogenomics**, v.9, n.4, p.477-480, 2008.
- MIZOGUCHI, M. et al. Two closely related subclass II SnRK2 protein kinases cooperatively regulate drought-inducible gene expression. **Plant and Cell Physiology**, v.51, p.842-847, 2010.
- NAKAGAMI, H., PITZSCHKE A.E., HIRT, H. Emerging MAP kinase pathways in plant stress signaling. **Trends in Plant Science**, v.10, 2005.
- NAKAHARA, T. et al. Antihypertensive effect of peptide-enriched soy sauce-like seasoning and identification of its angiotensin I-converting enzyme inhibitory substances. **Journal of Agricultural Food Chemistry**, v.58, p.821-827, 2010.
- NAKAMURA, T., SANO, H. A plasma-membrane linker for the phosphoinositide-specific phospholipase C in tobacco plants. **Plant Signaling & Behavior**, v.4, p.26-29, 2009.
- NARAYANAN, N. et al. Identification of candidate signaling genes including regulators of chromosome condensation protein family differentially expressed in the soybean-*Phytophthora sojae* interaction. **Theoretical and applied genetics**, v.118, p.399-412, 2009.
- NEPOMUCENO, A.L. Tolerância à seca em plantas. **Biociência**, n.23, 2001.
- NING, J. et al. A Raf-like MAPKKK gene DSM1 mediates drought resistance through reactive oxygen species scavenging in rice. **Plant Physiology**, v.152, p.876-90, 2010.
- NOBRE, C.A., ASSAD, E.D. e OYAMA, M.D. Mudança Ambiental no Brasil – O impacto do aquecimento global nos ecossistemas da Amazônia e na agricultura. In: **Scientific American Brasil**, n.12, Set 2005.
- OH, S.J. et al. Overexpression of the transcription factor AP37 in rice improves grain yield under drought conditions. **Plant Physiology**, v.150, p.1368-1379, 2009.
- PARK, S.Y. et al. Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. **Science**, v.324, p.1068-1071, 2009.
- PARRY, M.A.J. et al. Prospects for increasing photosynthesis by overcoming the limitations of Rubisco. **Journal of Agricultural Science**, v.145, p.31-43, 2007.
- PEDERSON, P. **Soybean growth and development**. Publ. PM-1945. Iowa State University Extension, 2004.
- QIN, F.Y. et al. Cloning and functional analysis of a novel DREB1/CBF transcription factor involved in cold-responsive gene expression in *Zea mays* L. **Plant Cell Physiology**, v.45, p.1042-1052, 2004.
- RAY, J.D., HEATHERLY, L.G., FRITSCHI, F.B. Influence of large amounts of nitrogen on nonirrigated and irrigated soybean. **Crop Science**, v.46, p.52-60, 2006.
- RIBAUT, J.M. et al. Use of molecular markers in plant breeding: drought tolerance improvement in tropical maize. In: **Quantitative Genetics, Genomics and Plant Breeding**, p.85-99, 2002.
- RONAGHI, M., et al. Real-time DNA sequencing using detection of pyrophosphate release. **Analytical Biochemistry**, v.242, p.84-89, 1996.
- SANGER F., NICKLEN S., COULSON A.R. DNA sequencing with chain-terminating inhibitors. **Proceedings Of The National Academy Of Sciences Of The U.S.A.**, v.74, p.5463-5467, 1977.
- SCHENA, M. et al. Quantitative monitoring of gene expression patterns with complementary DNA microarray. **Science**, v.270, p.467-470, 1995.
- SCHMUTZ J. et al. Genome sequence of the palaeopolyploid soybean. **Nature**, v.463, 2010.
- SEDIYAMA, T., TEIXEIRA, R.C., REIS, M.S. Melhoramento da soja. In: BORÉM, A. **Melhoramento de espécies cultivadas**. Editora UFV. Viçosa, p.969, 2005.
- SEKI, M. et al. Regulatory metabolic networks in drought stress responses. **Current Opinion in Plant Biology**, v.10, p.296-302, 2007.
- SHURTLEFF, W., AOYAGI, A. **History Of Edamame, Green Vegetable Soybeans And Vegetable-Type Soybeans (1275 - 2009)**: Extensively annotated bibliography and sourcebook. Soyinfo Center, USA, 2009a.
- SHURTLEFF, W., AOYAGI, A. **History Of Soybeans And Soyfoods In South America (1882-2009)**: Extensively annotated bibliography and sourcebook. Soyinfo Center, USA, 2009b.

- SINCLAIR, R.T. et al. Drought tolerance and yield increase of soybean resulting from improved symbiotic N₂ fixation. **Field Crops Research**, v.101, p.68-71, 2007.
- SONG, Q.J. et al. A new integrated genetic linkage map of the soybean. **Theoretical and Applied Genetics**, v.109, p.122-128, 2004.
- SONG, W.Y. et al. Relationship between calcium decoding elements and plant abiotic-stress resistance. **International Journal of Biological Sciences**, v.4, p.116-125, 2008.
- TAIZ, L. e ZEIGER, E. **Fisiologia vegetal**. Porto Alegre: Artmed, 5ed, 2010.
- TIAN, A.G. et al. Characterization of soybean genomic features by analysis of its expressed sequence tags. **Theoretical and Applied Genetics**, v.108, p.903-913, 2004.
- TRAN, L.S.P. Functional analysis of AHK1/ATHK1 and cytokinin receptor histidine kinases in response to abscisic acid, drought, and salt stress in *Arabidopsis*. **Proceedings Of The National Academy Of Sciences Of The U.S.A.**, v.104, p.20623-20628, 2007.
- VALENTE, M.A.S. et al. The ER luminal binding protein (BiP) mediates an increase in drought tolerance in soybean and delays drought-induced leaf senescence in soybean and tobacco. **Journal of Experimental Botany**, v.60, p.533-546, 2009.
- VELCULESCU, V.E. et al. Serial analysis of gene expression. **Science**, v.270, p.484-487, 1995.
- VINOD, M.S. et al. Candidate genes for drought tolerance and improved productivity in rice (*Oryza sativa* L.). **Journal of Bioscience**, v.31, p.69-74, 2006.
- WANG Y.J. et al. Functional analysis of a putative Ca²⁺ channel gene TaTPC1 from wheat. **Journal of Experimental Botany**, v.56, p.3051-3060, 2005.
- WANG, M. et al. A novel MAP Kinase gene in cotton (*Gossypium hirsutum* L.), GhMAPK, is involved in response to diverse environmental stresses. **Journal of Biochemistry and Molecular Biology**, v.40, p.325-332, 2007.
- WANG, S. et al. An *in silico* strategy identified the target gene candidates regulated by dehydration responsive element binding proteins (DREBs) in *Arabidopsis* genome. **Plant Molecular Biology**, v.69, p.167-178, 2009.
- WANG, Y.P., CHEN, X.T., QIU, L.J. Novel alleles among soybean Bowman-Birk proteinase inhibitor gene families. **Science in china series life sciences**, v.51, p.687-692, 2008.
- WRIGHT, G.C. Review of ACIAR selection for water use efficiency in legumes project recommends further research. **ACIAR Food Legume Newsletter**, p.2-3, 1996.
- YAMAGUCHI-SHINOZAKI, K., SHINOZAKI, K. Organization of cis-acting regulatory elements in osmotic- and cold-stress responsive promoters. **Trends Plant Science**, v.10, p.88-94, 2005.
- YOSHIOKA, K., SHINOZAKI, K. **Signal crosstalk in plant stress responses**. Wiley-Blackwell, p.23-32, 2009.
- ZHANG, G. et al. Overexpression of the soybean GmERF3 gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. **Journal of Experimental Botany**, v.60, p.3781-3796, 2009.
- ZHANG, Y. et al. *Arabidopsis* SDR1 enhances drought tolerance in crop plants. **Bioscience, Biotechnology and Biochemistry**, v.72, p.2251-2254, 2008.
- ZHI-MING, H., YANG, X., FROMM, M.E. Activation of the NaCl- and drought-induced *RD29A* and *RD29B* promoters by constitutively active *Arabidopsis* MAPKK or MAPK proteins. **Plant, Cell and Environment**, v.29, p.1761-1770, 2006.

4 ARTIGO: ARTIGO A SER SUBMETIDO À REVISTA PLANT SCIENCE (ISSN: 0168-9452).

“GENES INDUZIDOS PELO DÉFICIT HÍDRICO EM SOJA: VISÃO GERAL E FATORES REGULADORES DA TRANSCRIÇÃO.”

1 INTRODUÇÃO

A soja é uma das mais importantes oleaginosas sob cultivo extensivo devido sua ampla aplicação na alimentação humana, animal e grande potencial para a produção de biocombustíveis e uso de seus subprodutos na indústria [1]. Anualmente no Brasil são produzidos em média 56,3 milhões de toneladas (mi ton) de grãos de soja, fazendo com que o país alcance o segundo lugar no ranking de produtores desta cultura [2]. A soja possui também grande participação na balança comercial brasileira, sendo o Brasil o segundo maior exportador do produto, onde em 2009 exportou-se 28,7 milhões de toneladas de grãos [2].

Entretanto, apesar dos ganhos crescentes em produtividade alcançados nas últimas décadas na produção mundial de soja, a cultura vem sofrendo perdas consideráveis quando da ocorrência de eventos de seca. Eventos esses que potencialmente poderão ocorrer com maior frequência nas próximas décadas tendo-se em vista cenários de mudanças climáticas globais. A soja é uma espécie sensível ao déficit hídrico, principalmente durante o período de emergência, floração e enchimento de vagens. Desta maneira, a ocorrência de déficit hídrico durante estas etapas pode levar a diminuições drásticas na produtividade da cultura [3]. Na safra 2004/2005, uma das mais afetadas pela estiagem nos últimos anos, o Rio Grande do Sul foi o estado mais atingido, aonde a queda na produção da cultura chegou a 78 % em relação à safra 2002/2003, quando não ocorreram problemas severos de seca [4]. Neste mesmo estado a estiagem levou a uma redução média de 22% da produtividade na safra 2007/2008, gerando prejuízos em torno de R\$ 1,6 bilhão [5].

Embora o déficit hídrico constitua uma grande adversidade para a planta, esta é capaz de desencadear uma série de mecanismos morfológicos e fisiológicos de aclimatação e tolerância ao déficit hídrico, visando prevenir a perda

de água para o ambiente ou mitigar os danos decorrentes desta perda [6]. Dentre estes mecanismos destacam-se, por exemplo: alterações na condutância estomática mediada pelo hormônio ABA [7]; ajuste osmótico, que consiste no acúmulo ativo de solutos na célula vegetal [8]; acúmulo de moléculas osmoprotetoras, que além de participarem do ajuste osmótico desempenham importante papel na proteção de estruturas celulares [9] e a ação de proteínas com atividade antioxidante [10].

Todavia, a tolerância ao déficit hídrico é uma característica multigênica, isto é, depende intrinsecamente da ação de diversos genes e da interação de diferentes mecanismos. E somente um controle preciso desta malha de diferentes rotas metabólicas de resposta é o que permite as plantas tolerarem períodos de déficit hídrico. Este controle é exercido por diferentes atores que funcionam desde a percepção do estresse até a transmissão desta informação aos diferentes mecanismos de defesa. Dentro deste processo os fatores de transcrição assumem um papel chave no controle das respostas à condição de déficit hídrico. Várias famílias de fatores de transcrição têm sido identificadas no processo de controle de respostas à seca onde duas vias principais podem ser identificadas, uma dependente de ABA e outra independente [11].

Na via independente de ABA, fatores de transcrição pertencentes à família AP2/EREBP ativam os elementos *cis*-atuantes DRE de genes estresse-induzidos [12]. Estes fatores de transcrição foram isolados e divididos em *CBF/DREB1* e *DREB2* [13]. Em uma análise *in silico*, WANG e seus colaboradores [14] identificaram 160 genes responsivos a estresses abióticos, alvo dos fatores de transcrição *DREB* em *Arabidopsis thaliana*. Os genes *DREB1* foram inicialmente identificados na resposta ao frio, sendo mais tarde também descritos na resposta à seca e alta salinidade, enquanto os genes *DREB2* têm sido descritos principalmente na resposta ao déficit hídrico e estresse osmótico [13]. A superexpressão dos genes *DREB* em diferentes espécies de plantas como girassol, arroz e milho, aumentou a tolerâncias das plantas transgênicas a estresses abióticos como altas temperaturas e déficit hídrico [15, 16, 17]. Em soja, análises de ESTs (*Expressed Sequence Tags*) indicam a presença de 290 membros da família AP2/EREBP [18], sendo que já foram isolados três genes homólogos de *DREB*: *GmDREBa*, *GmDREBb*, e *GmDREBc* [19]. Em tabaco, a superexpressão do gene endógeno de soja *GmDREB2* conferiu maior tolerância ao déficit hídrico e estresse salino nas plantas transgênicas do que o verificado em plantas não transformadas [20].

O uso de técnicas que possibilitam uma análise global da expressão gênica, como a construção de bibliotecas subtrativas de cDNA, associadas aos sequenciamentos de última geração, têm possibilitado uma visão ampla das respostas moleculares ao déficit hídrico, visto que possibilitam a identificação de diversos genes expressos em resposta ao estresse. A construção deste tipo de biblioteca visa eliminar os transcritos expressos tanto na situação de tratamento/estresse quanto na situação controle, permanecendo na biblioteca apenas os transcritos expressos diferencialmente pela imposição do tratamento [21].

A proposta do presente trabalho foi identificar genes envolvidos nos mecanismos de resposta ao déficit hídrico em duas cultivares de soja, contrastantes quanto à sensibilidade ao déficit hídrico, e, através de sua categorização funcional (*in silico*), compreender as alterações dos processos biológicos envolvidos neste tipo de resposta. Visto a grande importância dos fatores de transcrição no controle da expressão dos diversos genes estresse-responsivos, dentre os genes identificados focamos na análise deste grupo. Foram avaliados fatores de transcrição descritos na resposta ao déficit hídrico, principalmente da família AP2/EREBP, cuja expressão diferencial foi validada via PCR quantitativo. Em nossas bibliotecas identificamos também 12 fatores de transcrição ainda não relacionados ao déficit hídrico, e, através da função destes genes (descrita na literatura), formulamos hipóteses sobre seu papel neste tipo de estresse.

2 MATERIAL E MÉTODOS

2.1 DESENHO EXPERIMENTAL E MATERIAL BIOLÓGICO

Sementes da cultivar Embrapa 48, com características de tolerância ao déficit hídrico, e da cultivar BR 16, considerada sensível ao déficit hídrico [22], foram germinadas em papel filtro umedecido com água destilada, a 25 °C e umidade próxima a 100%, em câmara de germinação, durante 4 dias, até atingirem um comprimento de raiz de aproximadamente 10 cm.

Em seguida, 210 plântulas de cada genótipo foram fixadas em orifícios feitos em placas de isopor, com capacidade para 30 plantas cada. Essas placas foram acomodadas sobre caixas de plástico com capacidade para 36 L (litros) contendo solução nutritiva de Hoagland diluída a 50% [23]. O pH (pH 6,6) da solução

foi monitorado diariamente e sempre que os valores estavam abaixo de 6,0 a solução era trocada a fim de evitar alterações na absorção de nutrientes. Foram utilizados aeradores, ligados 24 h (horas) por dia, para disponibilizar oxigênio para as raízes. Todo experimento foi conduzido em casa de vegetação, com temperatura ambiente variando entre $30 \pm 5^\circ\text{C}$ e umidade relativa entre $60 \pm 20\%$. Os tratamentos foram iniciados quando as plantas atingiram o estágio de desenvolvimento V3 [24].

De acordo com protocolo estabelecido por Martins et al. [25], com algumas alterações, foram aplicados sete tempos de exposição das raízes ao déficit hídrico, sendo que essa condição foi caracterizada pelo tempo que a raiz permaneceu fora da solução nutritiva. Os tratamentos impostos foram: T0 (sem estresse), T25 (25 minutos de estresse), e T50 (50 minutos de estresse), T75 (75 minutos de estresse), T100 (100 minutos de estresse), T125 (125 minutos de estresse) e T150 (150 minutos de estresse). O experimento foi fatorial em Delineamento de Blocos Casualizados (DBC), com 3 repetições, cada uma formada por 10 plantas para cada tratamento. Durante todo o tratamento, as plantas permaneceram com as raízes no escuro e com condições monitoradas de temperatura ($30 \pm 1^\circ\text{C}$) e umidade relativa ($60 \pm 5\%$).

As amostras de tecido radicular coletadas em cada uma das três repetições consistiram de um “pool” formado pelas 10 plantas dos respectivos tratamentos. O material vegetal foi imediatamente congelado em nitrogênio líquido e armazenado a -80°C até a extração de RNA total.

2.2 ANÁLISES FISIOLÓGICAS

Para avaliar a indução do déficit hídrico, foram analisadas variáveis fisiológicas como taxa fotossintética, condutância estomática e taxa de transpiração, medidas pelo equipamento portátil “Photosynthesis System” (LICOR, Inc., model LI-6400), usando o folíolo mediano do segundo trifólio (do ápice em relação à base) completamente expandidos sob intensidade luminosa de aproximadamente $1.000 \mu\text{moles m}^{-2}\text{s}^{-1}$.

2.3 EXTRAÇÃO DE RNA E ISOLAMENTO DE MRNA

O RNA total foi isolado das raízes de cada tempo separadamente (T0-T150) utilizando o reagente *TRIZOL* (Invitrogen), de acordo com instruções do fabricante. Após extração o RNA total foi quantificado por espectrofotometria (NanoDrop-1000) e teve sua qualidade analisada segundo Ausubel et al. [26].

O mRNA foi obtido a partir de 500 µg de RNA total, por meio do kit *FastTrack MAG mRNA Isolation* (Invitrogen), que utiliza partículas magnéticas conjugadas com oligo(dT) para captura do mRNA Poli(A)⁺, segundo especificações do fabricante.

2.4 CONSTRUÇÃO DAS BIBLIOTECAS SUBTRATIVAS

A confecção da biblioteca Subtrativa Supressiva (SSH), foi realizada com auxílio do Kit PCR Select cDNA Subtraction (Clontech), segundo especificações do fabricante. Para isolar os transcritos induzidos pelo déficit hídrico as amostras de mRNA poli(A⁺) proveniente das plantas submetidas ao déficit hídrico foram utilizadas como *tester* enquanto as amostras provenientes das plantas controle (T0) foram utilizadas como *drivers* para a subtração. Antes da síntese do cDNA, quantidades equivalentes de mRNA Poli (A⁺) de cada tratamento de déficit hídrico (T25 ao T150) foram reunidas para formar 3 *bulks* de *tester*, da seguinte maneira:

- T25 + T50 = *Bulk 1*
- T75 + T100 = *Bulk 2*
- T125 + T150 = *Bulk 3*

Deste modo, 2 µg (microgramas) de mRNA Poli (A⁺) de cada “bulk” formado foi utilizado para a construção de uma biblioteca Subtrativa Supressiva, originando as bibliotecas R1 (T25-50 min sob déficit hídrico), R2 (T75-100 min sob déficit hídrico) e R3 (T125-150 min sob déficit hídrico), para ambas as cultivares (BR 16 e Embrapa 48), totalizando 6 bibliotecas. Paralelamente, Amaral [47] utilizou tecidos foliares provenientes do mesmo experimento para a construção de 6 bibliotecas Subtrativas Supressivas, sendo elas L1 (T25-50 min sob déficit hídrico),

L2 (T75-100 min sob déficit hídrico) e L3 (T125-150 min sob déficit hídrico), para ambas as cultivares (BR 16 e Embrapa 48).

O produto obtido da confecção das bibliotecas foi diretamente submetido ao processo de sequenciamento.

2.5 SEQUENCIAMENTO E BIOINFORMÁTICA

Os transcritos foram sequenciados no equipamento *Genome Analyzer Iie* (Illumina) utilizando a tecnologia de sequenciamento por síntese (Illumina). As bibliotecas da cultivar BR 16 foram sequenciadas na empresa GenXPro (Alemanha) e produziram *single reads* de 36 pb (pares de bases). No caso de Embrapa 48, as bibliotecas foram sequenciadas na empresa Fasteris S.A. (Suíça) com “single” reads de 76 pb. Os milhares de *reads* obtidos para cada biblioteca foram alinhados no genoma referência (*Glycine max* cultivar Williams 82) [27] utilizando o programa SOAP (Short Oligonucleotide Alignment Program) [28] e montados em *contigs* com auxílio do programa Edena (Exact DE Novo Assembler) [29] em colaboração com o Laboratório de Genômica e Expressão da Unicamp-LGE.

Posteriormente, as sequências foram submetidas a buscas por similaridade com sequências já depositadas nos bancos de dados de sequências não redundantes de proteínas do NCBI (*National Center for Biotechnology Information*- <http://www.ncbi.nlm.nih.gov>) utilizando-se o algoritmo BLASTX [30]. Adicionalmente, as sequências foram analisadas pelo software AutoFact [31], que realizou buscas por similaridade em diferentes bancos de dados, dentre eles UniRef90 e UniRef100 [32], NCBI nrdb (non-redundant database), KEGG [33], Pfam [34], Smart [35]. Buscas no banco do Gene Ontology (<http://www.geneontology.org>), por meio do programa Blast2GO (<http://www.blast2go.org/home>) [36], forneceram informações sobre a função molecular e processo biológico relacionado à possível proteína codificada pelo transcrito em estudo [37].

Através do banco de dados Phytozome (<http://www.phytozome.net/>), onde foi possível acessar o genoma de soja (cultivar Williams 82), obtivemos a sequência completa dos genes de interesse, incluindo as regiões codantes, 3' e 5' UTR (do inglês, *untranslated region*). Estas sequências foram utilizadas para

análises de filogenia realizadas através do programa Mega4 [38], que utiliza o algoritmo ClustalW [39] para realizar alinhamentos globais. As análises filogenéticas foram feitas a partir da construção de árvores pelo método da Máxima Parsimônia [40], com a aplicação do teste de Bootstrap, com 1000 repetições, para apreender a robustez das árvores obtidas [41].

2.6 VALIDAÇÃO DAS BIBLIOTECAS SUBTRATIVAS E ANÁLISE DE EXPRESSÃO GÊNICA VIA PCR QUANTITATIVO

Para a validação dos perfis transcricionais obtidos nas bibliotecas subtrativas, selecionamos 6 genes candidatos da família AP2/EREBP, diferencialmente expressos em nossas bibliotecas de raízes (R1, R2 e R3) e/ou em bibliotecas de folhas (L1, L2 e L3) [47]. Estes genes foram previamente avaliados quanto à similaridade com sequências do NCBI e quanto a suas relações filogenéticas com genes *DREB* de família Fabaceae. A validação foi feita através da análise da expressão dos genes via PCR quantitativo. Foram avaliadas folhas e raízes, de ambas as cultivares (Embrapa 48 e BR 16), analisadas pontualmente durante todo o período de estresse (0 a 150 min.). Para cada um dos tempos de estresse avaliados, foram analisadas 3 repetições biológicas e, para cada repetição, foram feitas 3 replicatas técnicas.

O cDNA empregado nestas análises foi obtido da transcrição reversa do RNA total (tratado com DNase -Invitrogen) por meio do kit *Superscript II First Strand Synthesis* (Invitrogen). Os genes endógenos foram selecionados de acordo com a literatura e avaliados em ensaios prévios de expressão relativa, nos quais foi examinada a estabilidade da expressão sob déficit hídrico por meio dos programas Genorm [42] e NormFinder [43]. Dentre os seis endógenos testados (*Elongation Factor 1- α* , *glyceraldehyde-3-phosphate dehydrogenase*, Beta actina, Alfa Tubulina, Beta tubulina, RNA ribossomal 18 S) os genes *Elongation Factor 1- α* e Beta actina apresentaram expressão mais estável quando utilizados em materiais dos diferentes tratamentos a serem avaliados, sendo por isso selecionados para a utilização como normalizadores no processo de quantificação da expressão relativa.

Os *primers* específicos para os genes alvo foram desenhados com o auxílio do programa *Primer Express 3.0* (Applied Biosystems). A sequência dos *primers* foi determinada próxima à região 3' do gene, sempre que esta estava

disponível, resultando em amplicons de 80 a 150 pares de bases (pb). Os primers foram previamente comparados com o genoma da soja utilizando-se a ferramenta BLAST no banco de dados Phytozome, para verificação da especificidade das sequências escolhidas. Os primers selecionados estão descritos na tabela 1.

Tabela 1 - Primers específicos para os genes alvo. Número do gene alvo e respectivo *genemodel* (gene predito em soja). Sequências de bases e temperatura de desnaturação (T_m) dos primers *Forward* e *Reverse*. Eficiência refere-se à eficiência das reações de amplificação dos pares de primers, calculada a partir da construção de curvas padrão.

Alvo	<i>Genemodel</i>	<i>Forward</i>		<i>Reverse</i>		Eficiência (%)
		Sequência	T_m	Sequência	T_m	
Alvo 1	Glyma10g2	AAAACCCATGCTTATCTTCTCCTT	59	CGGCAATTCCAATGTACAACA	58	92,14
Alvo 2	Glyma14g09	GATGATGATGCCTCGGAGTTG	59	CGGAAAAACAAGAAAAGGGATATATC	59	86,70
Alvo 3	Glyma04g1	CCTTGGGTCATAGCTGGTTTTT	58	AGATTACAGCGATGCACTTGAGAA	59	96,22
Alvo 4	Glyma14g3	ACCTTggTCAgCTgggTTTTT	59	gACAACACCTAGAAggCTAgAAgTCA	58	97,31
Alvo 5	Glyma13g01	TTGCTTATTGGCTATTCGATGGT	59	TCCATGGCCAAGCAAGAAA	59	89,12
Alvo 6	Glyma13g3	ACTCCTTTTGAACCTTCTCCACTCT	59	GAACCATCAAAGCAAAGGAAA	58	92,64

Depois de sintetizados, os *primers* foram utilizados em reações de quantificação absoluta via PCR quantitativo, nas quais a análise das curvas de dissociação também confirmou a especificidade dos *amplicons*, devido à presença de apenas um pico de dissociação. Foram também construídas curvas padrão, geradas a partir de diluições seriais (10^{-1}) de um pool de cDNA, para a estimativa da eficiência das reações de amplificação. As concentrações de *primers* foram ajustadas de maneira a obter eficiências maiores do que 85%. A expressão relativa dos genes alvo foi calculada de acordo com a metodologia $2^{-\Delta\Delta C_t}$ [44], onde os genes calibradores endógenos foram *Elongation Factor 1- α* e *Beta actina*, e o controle do tratamento foi o tempo 0 min (condição sem estresse).

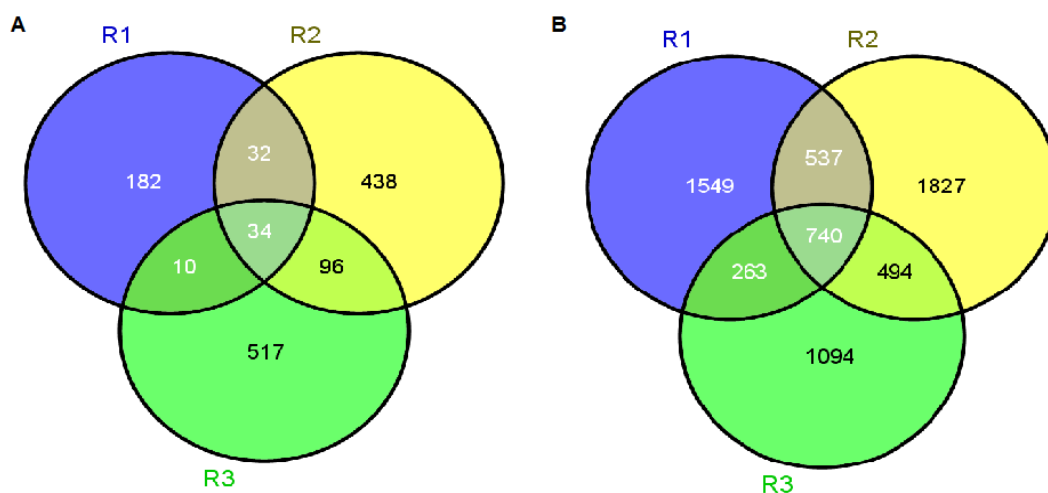
Para apreender a confiabilidade dos perfis transcricionais obtidos, além do cálculo pela metodologia $2^{-\Delta\Delta C_t}$, utilizamos o programa REST 2009 [45], que calcula a expressão gênica através da aleatorização dos C_t s das amostras controle e tratadas. Este programa também realiza análises estatísticas, utilizando randomizações para determinar se existem diferenças significativas entre a situação de tratamento e a controle.

3 RESULTADOS

3.1 BIBLIOTECAS SUBTRATIVAS DE CDNA

As bibliotecas da cultivar BR 16 (R1, R2 e R3), submetidas a 36 ciclos de sequenciamento, geraram 2.486.569, 2.458.847 e 2.428.923 de reads, respectivamente, enquanto que as bibliotecas de Embrapa 48 (76 ciclos de sequenciamento) produziram 5.144.645, 5.359.395 e 3.095.694 de *reads* nos respectivos tempos. A montagem dos *reads* destas bibliotecas gerou um total de 1.515 *contigs* de sequências diferencialmente expressas para BR 16 e 9.278 *contigs* diferencialmente expressos para Embrapa 48, distribuídos entre as bibliotecas como detalha a Figura 1.

Figura 1 - Diagrama de Venn mostrando a distribuição dos transcritos identificados nas bibliotecas subtrativas. Os números de genes exclusivos e comuns aos diferentes períodos do déficit hídrico são apresentados para as bibliotecas R1 (25-50 min sob estresse), R2 (75-100 min) e R3 (125-150 min). **(A)** BR 16 e **(B)** Embrapa 48.



3.2 ANÁLISES DE BIOINFORMÁTICA

Os *contigs* montados foram submetidos a análises de BLASTX para busca de sequências por similaridade. A maioria das sequências encontrou similaridade com proteínas descritas (*hits*), entretanto parte destas não encontrou proteínas similares (*no hits*) ou foi similar a sequências de proteínas cuja função

permanece desconhecida (*unknown, unnamed, hypothetical* ou *putative proteins*) (Tabela 2).

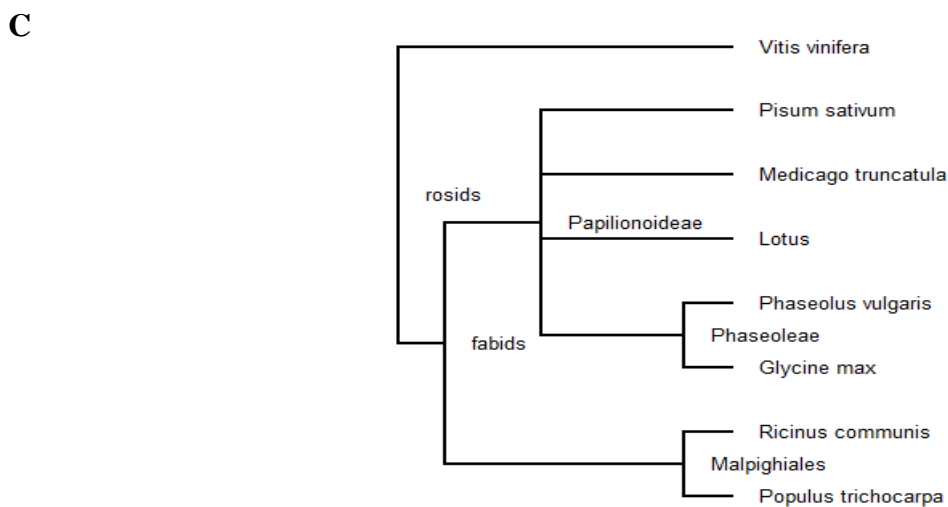
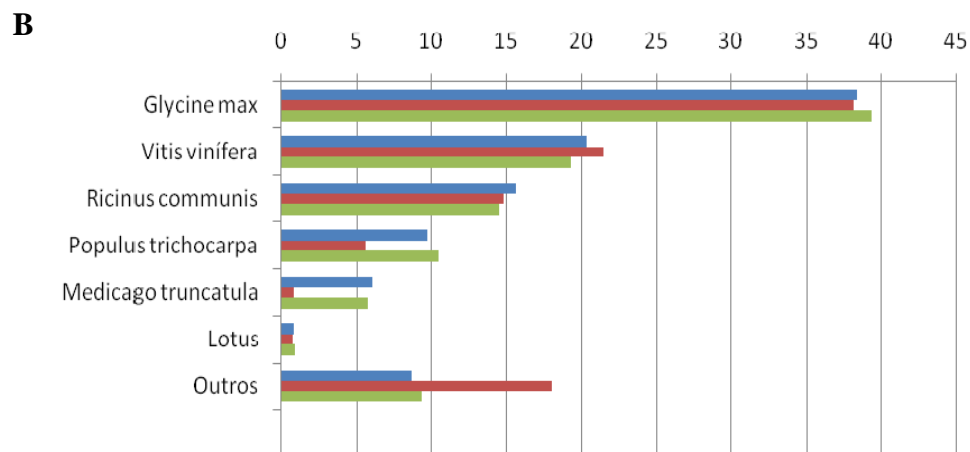
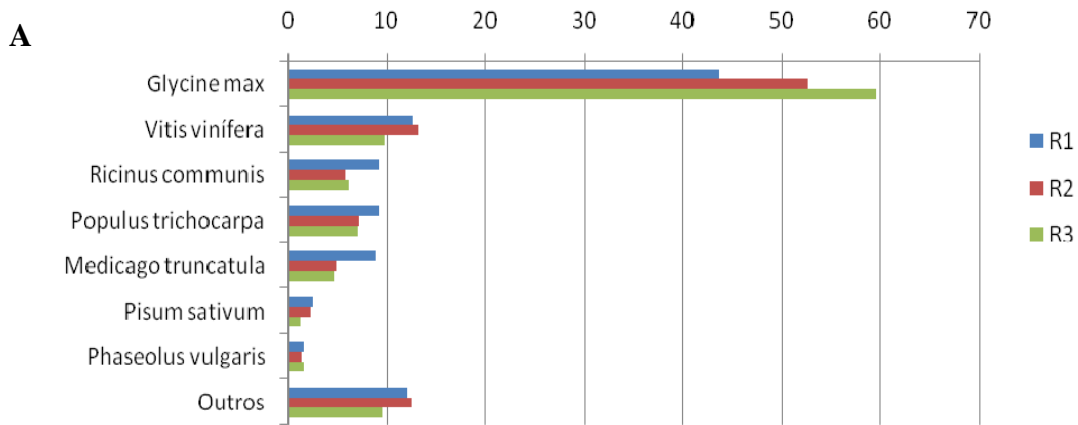
Tabela 2 - Análise dos *contigs* utilizando o BLASTX apresentando o número de *reads* geradas por biblioteca (R1, R2, R3) das cultivares BR 16 e Embrapa 48 e número de *contigs* montados a partir da sobreposição das *reads* (*Contigs*); sequências sem similaridade (*no hits*); sequências similares a proteínas descritas (*Hits*); *hits* com proteínas conhecidas (Sequências conhecidas) e desconhecidas (Sequências desconhecidas) .

Cultivar	Biblioteca	Reads	Contigs	no hits	Hits	Sequências conhecidas	Sequências desconhecidas
BR 16	R1	2.486.569	258	6	252	212	40
	R2	2.458.847	600	34	566	499	67
	R3	2.428.923	657	53	604	525	79
	TOTAL	7.374.339	1515	93	1422	1236	186
Embrapa 48	R1	5.144.645	3089	198	3445	1882	1563
	R2	5.359.395	3598	194	4409	2525	1884
	R3	3.095.694	2591	161	2948	1658	1290
	TOTAL	13.599.73	9278	553	10802	6065	4737

4

A maioria das sequências analisadas (38 a 59%) apresentaram similaridade com sequências de soja (*Glycine max*) depositadas no banco de dados NCBI, seguido por *Vitis vinifera*, *Ricinus comunis* e *Populus trichocarpa*, em ambas cultivares (Figura 2A e B). Uma árvore taxonômica foi construída com auxílio da ferramenta *Taxonomy Common Tree* (<http://www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwwcmt.cgi>) para mostrar as espécies mais observadas nas pesquisas de BLASTX das sequências diferencialmente expressas em nossas bibliotecas, o que possibilitou visualizar a proximidade das espécies em questão (Figura 2C).

Figura 2 - Distribuição dos Resultados de BLASTX por espécies. Porcentagem de sequências por espécies nas cultivares **(A)** BR 16 e **(B)** Embrapa 48, distribuídos nas bibliotecas R1 (T25-50 min sob estresse), R2 (T75-100 min) e R3 (T125-150 min); **(C)** árvore taxonômica com as espécies mais frequentes no resultado de BLASTX de ambas as cultivares.



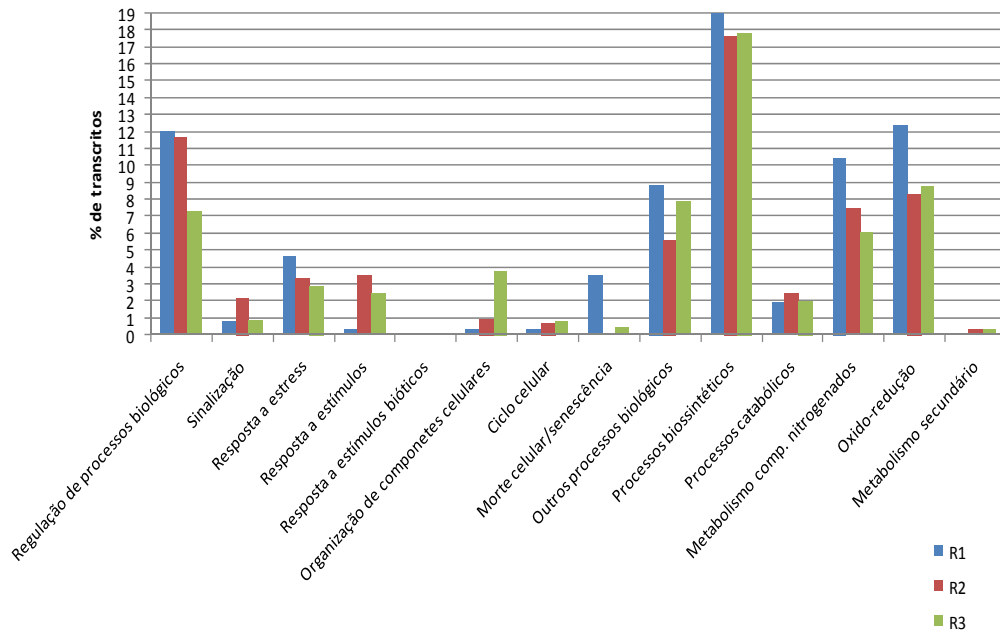
3.3 ALTERAÇÕES MOLECULARES EM DECORRÊNCIA DO DÉFICIT HÍDRICO

Para melhor compreensão das alterações moleculares decorrentes da aplicação do déficit hídrico, em ambas as cultivares estudadas, as sequências diferencialmente expressas foram analisadas quanto ao processo biológico relacionado à provável proteína codificada. Nesta análise, as sequências foram anotadas de acordo com os termos GO para Processos Biológicos (nível 3) segundo o banco de dados Gene Ontology (Figura 3).

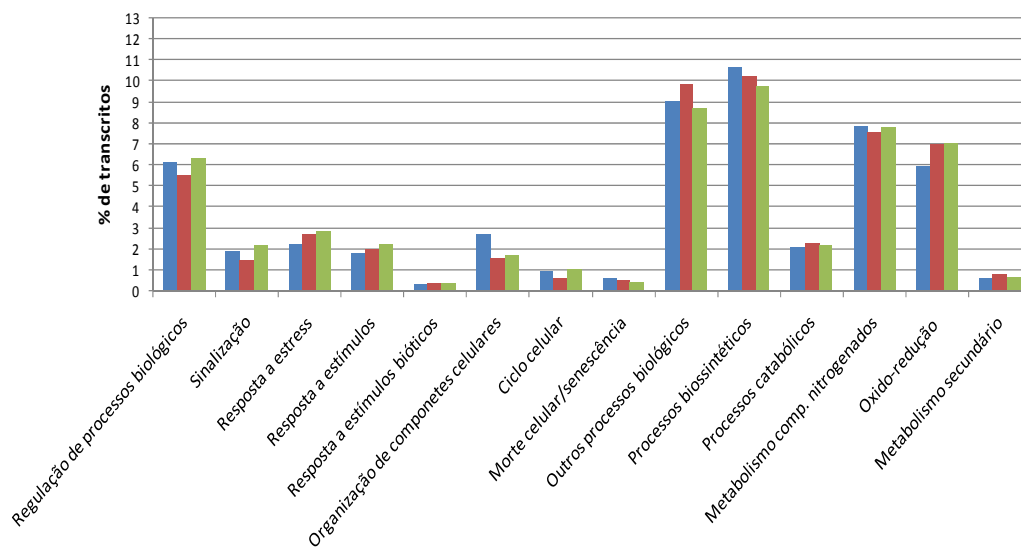
Nesta categorização, as categorias GO:0044237: *cellularmetabolicprocess* e GO:0044238: *primarymetabolicprocess* foram retiradas porque estes grupos são muito gerais, não sendo informativos [46]. Inicialmente foram obtidos de 29 a 54 termos GOs em cada biblioteca analisada. Estes termos GOs foram agrupados em grandes classes, de maneira que cada classe é composta por GOs relacionados a processos semelhantes, da seguinte maneira: Regulação de processos biológicos (GO:0065008; GO:0065009; GO:0050789), Sinalização (GO:0023033; GO:0023046; GO:0007154), Resposta a estresse (GO:0006950), Resposta a estímulos (GO:0042221; GO:0009607; GO:0009628; GO:0009719; GO:0009605; GO:0051606; GO:0051716; GO:0006955), Resposta a estímulos bióticos (GO:0009607), Organização de componentes celulares (GO:0071555; GO:0016044; GO:0071669; GO:0042546; GO:0070271; GO:0022607; GO:0022613), Ciclo celular (GO:0007017; GO:0022414; GO:0007049; GO:0051301; GO:0071103; GO:0030029), Morte celular/senescência (GO:0008219; GO:0007568), Outros processos biológicos (GO:0051234; GO:0045229; GO:0043933; GO:0061024; GO:0022411; GO:0006996; GO:0022406; GO:0048856; GO:0048869; GO:0006323; GO:0019725; GO:0033036; GO:0051641; GO:0034621; GO:0003008; GO:0008037), Processos biossintéticos (GO:0009058), Processos catabólicos (GO:0009056), Metabolismo de compostos. Nitrogenados (GO:0006807), Oxido-redução (GO:0055114) e Metabolismo secundário (GO:0019748). Devido ao fato das bibliotecas diferirem quanto ao número de genes diferencialmente expressos, cada classe apresentada foi normalizada pelo número total de genes de cada biblioteca, o que tornou possível identificar a porcentagem de sequências que cada classe representa em relação ao total de sequências obtidas (Figura 3).

Figura 3 - Categorização funcional dos transcritos. Porcentagem de sequências (%) atribuídas aos respectivos processos biológicos (Gene Ontology) para as cultivares (A) BR 16 e (B) Embrapa 48, distribuídas nas bibliotecas R1 (T25-50 min sob estresse), R2 (T75-100 min) e R3 (T125-150 min).

A



B



3.4 GENES RELACIONADOS AO DÉFICIT HÍDRICO

Dentre os transcritos identificados nas cultivares de soja podemos observar diversas sequências com similaridade aos genes descritos nas respostas das plantas ao déficit hídrico, como as sequências codificadoras de proteínas

chaperonas, proteínas relacionadas à transdução de sinais, aquaporinas, proteínas relacionadas à desidratação e ao estresse oxidativo, entre outras (Tabela 3).

Tabela 3 - Número de genes responsivos ao déficit hídrico nas bibliotecas R1 (T25-50 min sob estresse), R2 (T75-100 min) e R3 (T125-150 min), das cultivares (BR 16 e Embrapa 48).

	BR 16			Embrapa 48		
	R1	R2	R3	R1	R2	R3
Chaperonas						
<i>heat shock proteins</i>	2	5	7	28	48	40
<i>chaperonin</i>	2	0	0	2	3	3
<i>DNAJ heat shock</i>	0	2	3	4	10	7
<i>Calreticulin</i>	0	1	2	1	2	1
<i>seed maturation PM37</i>	0	1	1	4	3	5
Transdução de sinais						
<i>casein kinase</i>	3	3	0	3	4	3
<i>receptor-like kinase</i>	2	9	3	11	9	5
<i>stress-induced receptor-like kinase</i>	0	0	0	8	6	7
<i>calcium-dependent kinase</i>	3	3	1	9	11	8
<i>serine/threonine protein kinase</i>	7	4	4	20	30	18
<i>MapK</i>	0	2	1	4	9	6
<i>WD-40</i>	0	2	0	3	3	2
<i>GTP-binding</i>	0	4	7	13	26	20
<i>CBL</i>	0	1	2	3	1	0
<i>transducin</i>	0	0	0	0	0	3
<i>calmodulin</i>	0	8	0	3	4	5
Aquaporinas						
<i>aquaporin</i>	1	1	3	5	5	4
<i>PIP1</i>	0	3	6	1	3	0
<i>PIP2</i>	0	1	4	1	1	2
<i>nodulin 1, 26</i>	1	2	2	0	0	0
Desidratação						
<i>LEA</i>	0	3	0	2	1	2
<i>maturation-associated</i>	1	1	0	0	0	0
<i>Dehydrin</i>	0	2	4	0	1	1
<i>9-cis-epoxycarotenoid dioxygenase</i>	1	1	1	2	2	0
<i>dehydration-induced protein</i>	1	0	0	0	3	0
<i>early responsive to dehydration</i>	0	4	7	4	0	0
Estresse oxidativo						
<i>n-acetylglucosamine transferase</i>	1	0	0	0	0	0
<i>ubiquinone oxidoreductase</i>	0	0	1	0	0	0
<i>aconitase</i>	0	0	2	1	2	1
<i>membrane protein</i>	0	0	2	0	0	0
<i>monodehydroascorbate reductase</i>	1	1	0	0	0	0

<i>ATA15</i>	1	0	0	0	0	0
<i>superoxide dismutase</i>	0	1	4	0	1	0
<i>oxidase</i>	0	1	0	0	0	0
<i>catalase</i>	0	4	2	0	2	0
<i>peroxidase</i>	0	0	2	0	0	0
<i>purple acid phosphatase</i>	0	1	3	0	0	0
<i>pyridoxine biosynthesis protein</i>	0	1	0	0	0	0
<i>SPINDLY</i>	2	0	0	0	0	0
<i>ascorbate peroxidase</i>	0	0	0	1	0	0
<i>glutathione peroxidase</i>	0	0	0	0	2	2

3.5 FATORES DE TRANSCRIÇÃO

3.5.1 Fatores de Transcrição Relacionados ao Déficit Hídrico

Outra importante classe de proteínas relacionadas ao déficit hídrico em plantas são os Fatores de Transcrição. Esses fatores agem no início da cascata de resposta ao estresse, modulando a expressão de diversos genes estresse-responsivos. Dentre os principais fatores de transcrição descritos no déficit hídrico estão as proteínas NAC, MYB, WRKY e DREB. No presente trabalho, foram encontradas sequências relacionadas a todos estes fatores de transcrição, como pode ser observado na Tabela 4.

Tabela 4 - Principais fatores de transcrição relacionados ao déficit hídrico. Número de sequências (contigs únicos) identificadas nas bibliotecas (T25-50 min sob estresse), R2 (T75-100 min) e R3 (T125-150 min), das cultivares BR 16 e Embrapa 48.

	BR 16			Embrapa 48		
	R1	R2	R3	R1	R2	R3
DREB	2	1		2	3	2
WRKY		3	1	34	33	19
MYB	4	8	6	15	17	13
NAC			1	16	16	21

3.5.2 Fatores de transcrição AP2/EREBP

Visto a grande importância dos fatores de transcrição da família *AP2/EREB* para o déficit hídrico, onde os fatores *DREB* estão inseridos, possíveis candidatos foram selecionados para análises de filogenia, nas cultivares tolerante e

sensível. Além das sequências diferencialmente expressas em raiz, foram selecionadas também sequências diferencialmente expressas em folha, provenientes de bibliotecas subtrativas construídas com material biológico do mesmo experimento [47]. Para obter as sequências completas dos genes candidatos, os *contigs* foram submetidos ao BLAST no Phytozome, onde foram encontrados os *genemodels* ou “genes preditos” correspondentes, que posteriormente foram analisados por meio do BLASTX no NCBI (Tabela 5).

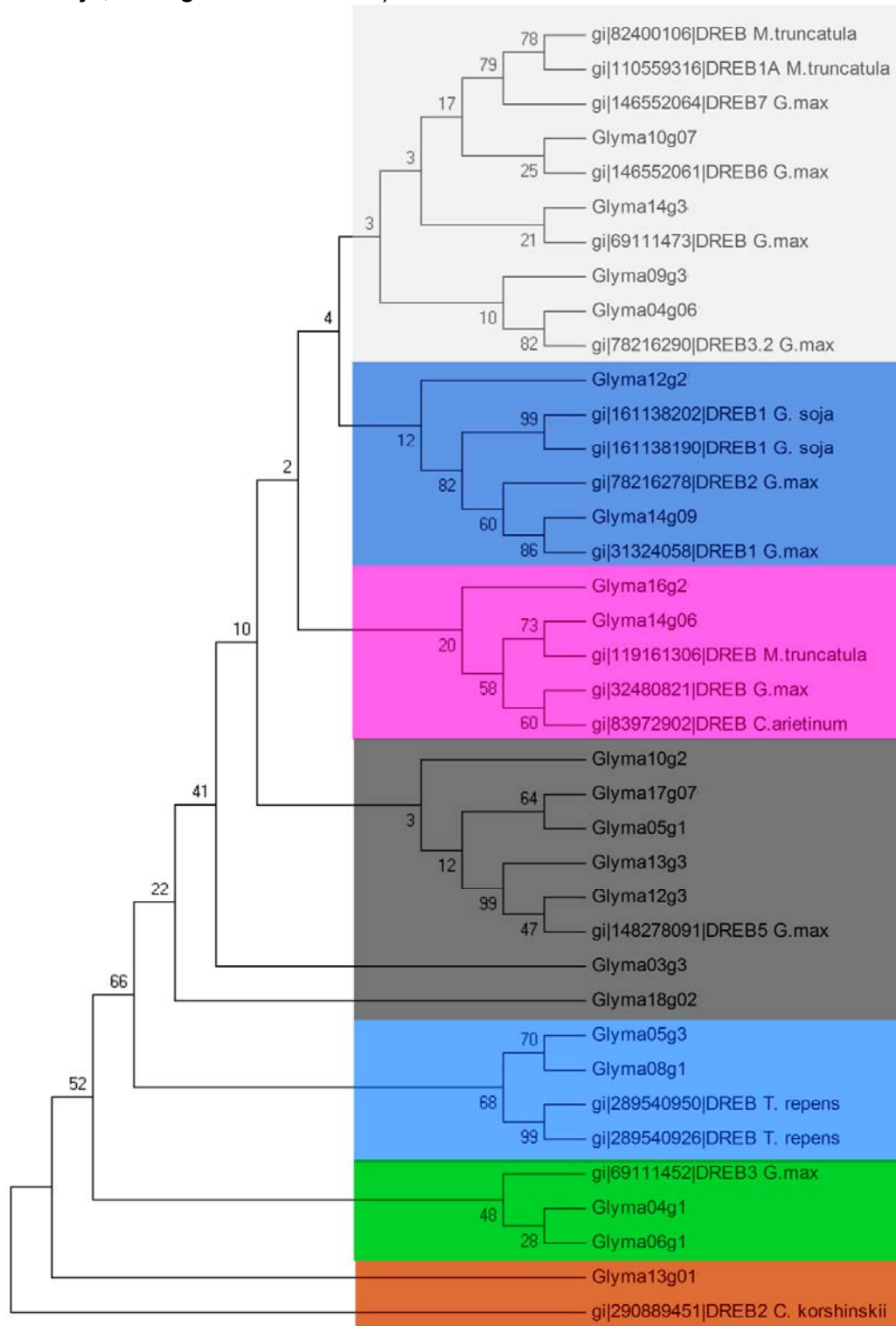
Tabela 5 - Transcritos expressos nas cultivares BR 16 e Embrapa 48 que apresentaram similaridade aos genes da família AP2/EREB. Os genes candidatos são representados pelo *genemodel* correspondente. As sequência com maior similaridade, o respectivo e-value e identidade (% de nucleotídeos idênticos) são referentes ao resultado do BLASTX contra o banco de dados NCBI. Distribuição dos genes por bibliotecas subtrativas, sendo R1, R2 e R3 referente às bibliotecas de raiz e L1, L2 e L3 referentes às bibliotecas de folha.

Gene candidato	Resultado de BLASTX no NCBI	e-value	Id (%)	Bibliotecas	
				BR 16	Emb. 48
Glyma03g3	gb ACU20075.1 APETALA2 and EREBP APETALA2 and EREBP (ethylene responsive element binding protein) [Glycine max]	3 E-115	100	L3	
Continua					
Gene candidato	Resultado de BLASTX no NCBI	e-value	Id%	Bibliotecas	
				BR 16	EmB. 48
Glyma04g06	gb ABB36646.1 ABB36646.1 dehydration-responsive element binding protein 3 [Glycine max]	4 E-29	71	L3	
Glyma04g1	gb AAZ03388.1 dehydration responsive element-binding protein 3 [Glycine max]	6 E-112	99	R1/R2/L3	R3
Glyma05g1	gb AAK14326.1 APETAL2-like protein [Pisum sativum]	1 E-151	75	L1	
Glyma05g3	gb AAQ57226.1 DREB2 [Glycine max]	8 E-109	84		L1
Glyma06g1	gb AAZ03388.1 dehydration responsive element-binding protein 3 [Glycine max]	4 E-94	80	R1/L2/L3	
Glyma08g1	gb AAQ57226.1 DREB2 [Glycine max]	3 E-125	100	L2	R1
Glyma09g3	GENE ID: 7496715 DREB55 AP2/ERF domain-containing transcription factor [Populus trichocarpa]	3 E-29	53	L3	
Glyma10g07	gb EEF33807.1 Dehydration-responsive element-binding protein 1B, putative [Ricinus communis]	3 E-42	52	L2/L3	R1/R2/R3
Glyma10g2	gb ACU19106.1 unknown [Glycine max]	9 E-60	71	L1	

	APETALA2 and EREBP				
Glyma12g2	gb AAM98171.1 ethylene responsive element binding factor 4 (AtERF4) [<i>Arabidopsis thaliana</i>]	4 E-13	100	L3	
Glyma12g3	GENE ID: 100101898 DREB5 dehydration-responsive element binding protein 5 [<i>Glycine max</i>]	2 E-145	93		L1
Glyma13g01	gb ADD69957.1 DREB2 [<i>Caragana korshinskii</i>]	8 E-62	68	L1/L2/L3	R1/R2
Glyma13g3	GENE ID: 100101898 DREB5 dehydration-responsive element binding protein 5 [<i>Glycine max</i>]	5 E-72	59	R2	L1
Glyma14g06	gb ACJ37441.1 AP2 domain-containing transcription factor 7 [<i>Glycine max</i>]	4 E-180	97		L2/L3
Glyma14g09	GENE ID: 547642 DREB1 dehydration responsive element binding protein [<i>Glycine max</i>]	7 E-97	100		R2
Glyma16g2	gb ACU24559.1 unknown [<i>Glycine max</i>] APETALA2 and EREBP	9 E-63	74	L1	
Glyma17g07	gb EEF07666.1 AP2 domain-containing transcription factor [<i>Populus trichocarpa</i>]	1 E-141	63	L3	
Glyma18g02	gb ACU17684.1 unknown [<i>Glycine max</i>] APETALA2 and EREBP	1 E-98	74		R2
Glyma14g3	dbj BAG50057.1 transcription factor AP2-EREBP [<i>Lotus japonicus</i>]	7 E-60	71		R1/R2/L1

O alinhamento global das sequências dos genes candidatos, traduzidas em aminoácidos, permitiu a montagem de uma árvore filogenética, que evidencia a relação das sequências candidatas com as sequências de fatores de transcrição da família AP2/EREB, mais especificamente da subfamília *DREB*, descritas em soja e outras plantas da família (*Fabaceae*), depositadas no banco de dados do NCBI (Figura 4).

Figura 4 - Árvore filogenética com proteínas codificadas pelos genes candidatos e os genes *DREB* descritos no banco de dados NCBI. Genes candidatos são representados pelos respectivos *genemodels* (Glyma) e as sequências de *DREBs* homólogos da família Fabaceae (*Glycine max*, *Medicago truncatula*, *Cypripedium arietinum*, *Trifolium repens*, *Glycine soja*, *Caragana korshinskii*).



A árvore foi construída utilizando algoritmo Clustal W no programa MEGA 4.0 com o teste *Bootstrap* (1000 repetições).

3.6 VALIDAÇÃO DAS BIBLIOTECAS SUBTRATIVAS E ANÁLISE DE EXPRESSÃO DE GENES DA FAMÍLIA AP2/EREBP

O aumento na expressão de genes da família AP2/EREB, encontrado nas bibliotecas subtrativas, pôde ser confirmado via PCR quantitativo, no qual todos os genes avaliados (Tabela 6) apresentaram-se induzidos em relação à amostra controle (0 min.- condição sem estresse), na maioria dos níveis de estresse aplicados (25 a 150 min.), tanto na cultivar BR 16 (Figura 5) quanto na cultivar Embrapa 48 (Figura6).

Tabela 6 - Genes avaliados por PCR quantitativo. Identificação do gene alvo, nome do gene predito em soja segundo Phytozome e Descrição do gene, segundo análises de similaridade com sequências do NCBI. A última coluna refere-se ao tecido onde o gene foi isolado (por meio das bibliotecas subtrativas), nos quais os genes serão validados via PCR quantitativo.

Alvo	Gene	Descrição do gene	Tecido
1	Glyma10g2	Similar a APETALA2 e EREBP desconhecido de soja	Folha
2	Glyma14g09	Similar a DREB1 de soja	Raiz
3	Glyma04g1	Similar a DREB3 de soja	Raiz
4	Glyma14g3	Similar a fator de transcrição AP2-EREBP de <i>Lotus japonicus</i>	Raiz
5	Glyma13g01	Similar a DREB2 de <i>Caragana korshinskii</i>	Folha e raiz
6	Glyma13g3	Similar a DREB5 de soja	Folha e raiz

Figura 5 - PCR quantitativo de genes de família AP2/EREB. Expressão relativa dos genes na cultivar BR 16, durante os seis períodos de estresse (25 a 150 min.de déficit hídrico- T25 a T150) em relação ao controle (0 min.- condição sem déficit hídrico). **(A)**alvo 1 (Glyma10g2); **(B)**alvo 2 (Glyma14g09); **(C)**alvo 3 (Glyma04g1); **(D)**alvo 4 (Glyma14g3); **(E)**alvo 5 (Glyma13g01) avaliado em folha; **(F)**alvo 5 (Glyma13g01) avaliado em raiz; **(G)**alvo 6 (Glyma13g3) avaliado em folha; **(H)**alvo 6 (Glyma13g3) avaliado em raiz. Barras representam erro padrão da média

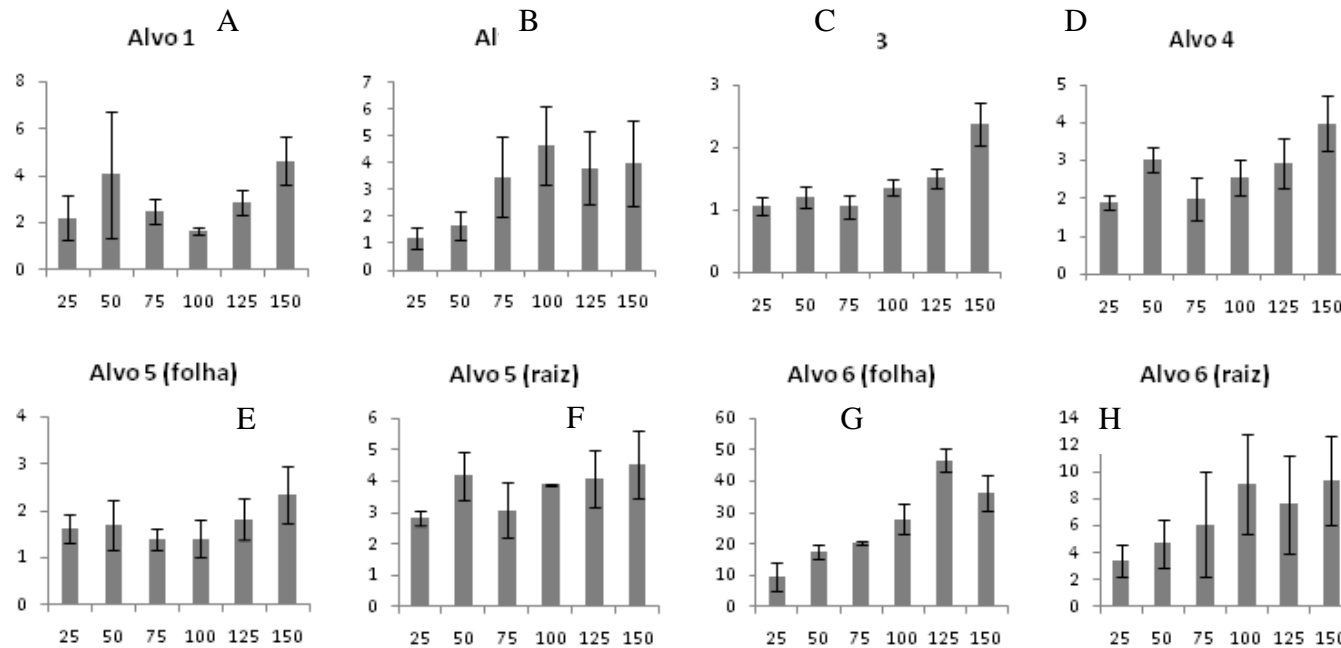
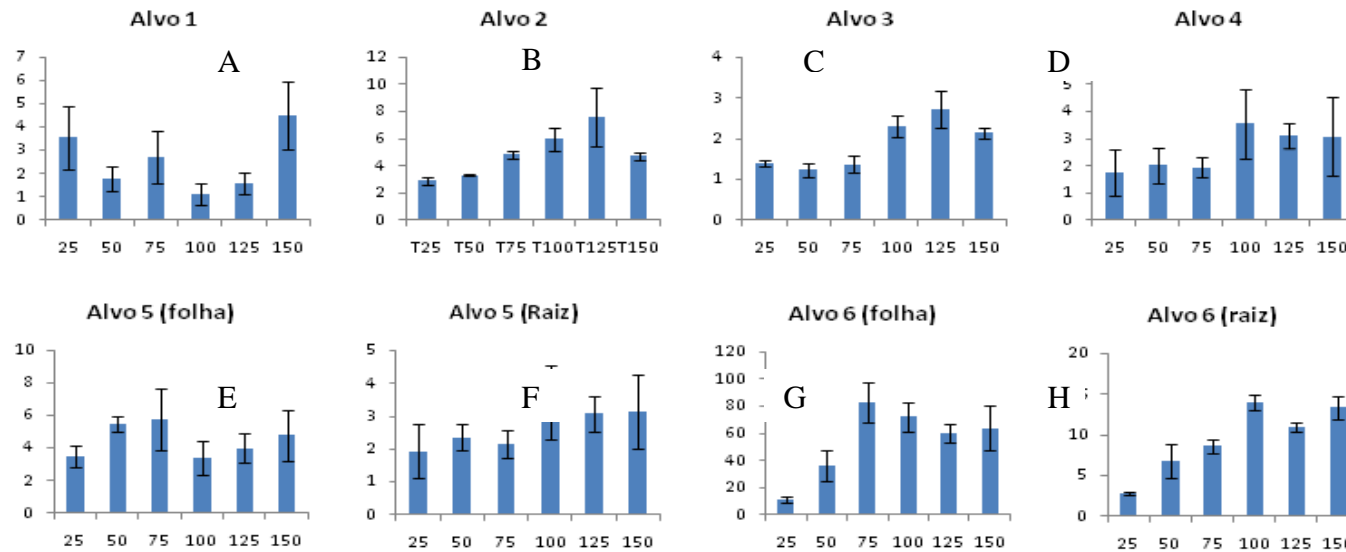


Figura 6 - PCR quantitativo de genes de família AP2/EREB. Expressão relativa dos genes na cultivar Embrapa 48, durante os seis períodos de estresse (25 a 150 min. de déficit hídrico- T25 a T150) em relação ao controle (0 min.- condição sem déficit hídrico). **(A)** alvo 1 (Glyma10g2); **(B)** alvo 2 (Glyma14g09); **(C)** alvo 3 (Glyma04g1); **(D)** alvo 4 (Glyma14g3); **(E)** alvo 5 (Glyma13g01) avaliado em folha; **(F)** alvo 5 (Glyma13g01) avaliado em raiz; **(G)** alvo 6 (Glyma13g3) avaliado em folha; **(H)** alvo 6 (Glyma13g3) avaliado em raiz. Barras representam erro padrão da média.



Através do programa REST 2009 [45], foram realizados testes estatísticos para determinar se existem diferenças significativas entre a situação de tratamento e a controle, como mostra a Tabela 7. Como pode ser observado, todos os genes foram significativamente induzidos pela maioria dos tempos de tratamento aplicados (Tabela 7).

Tabela 7 - Teste de hipóteses baseado no programa Rest 2009. Genes alvo: alvo 1 (Glyma10g2); alvo 2 (Glyma14g09); alvo 3 (Glyma04g1); alvo 4 (Glyma14g3); alvo 5 (Glyma13g01); alvo 6 (Glyma13g3). P(H1) refere-se à probabilidade do resultado de expressão das plantas tratadas (25 a 150 min.de déficit hídrico) ser diferente das plantas controle (0 min.- condição sem déficit hídrico) devido ao acaso. Result. demonstra o resultado da expressão relativa, onde considera $P < 0,05$ para genes significativamente induzidos (UP).

Gene alvo	Tratamento	BR 16		Embrapa 48		Gene alvo	Tratamento	BR 16		Embrapa 48	
		P(H1)	Result.	P(H1)	Result.			P(H1)	Result.	P(H1)	Result.
Alvo 1	T25	0,012	UP	0,001	UP	Alvo 5 folha	T25	0,000	UP	0,000	UP
	T50	0,005	UP	0,075			T50	0,000	UP	0,000	UP
	T75	0	UP	0,003	UP		T75	0,000	UP	0,000	UP
	T100	0,013	UP	0,513			T100	0,000	UP	0,000	UP
	T125	0,001	UP	0,174			T125	0,000	UP	0,000	UP
	T150	0	UP	0	UP		T150	0,000	UP	0,000	UP
Alvo 2	T25	0,963		0,000	UP	Alvo 5 raiz	T25	0,000	UP	0,000	UP
	T50	0,246		0,000	UP		T50	0,000	UP	0,000	UP
	T75	0,008	UP	0,000	UP		T75	0,000	UP	0,000	UP
	T100	0,001	UP	0,000	UP		T100	0,000	UP	0,000	UP
	T125	0,003	UP	0,000	UP		T125	0,000	UP	0,000	UP
	T150	0,005	UP	0,000	UP		T150	0,000	UP	0,000	UP
Alvo 3	T25	0,248		0,001	UP	Alvo 6 folha	T25	0,000	UP	0,000	UP
	T50	0,026	UP	0,096			T50	0,000	UP	0,000	UP
	T75	0,889		0,044	UP		T75	0,000	UP	0,000	UP
	T100	0,000	UP	0,000	UP		T100	0,000	UP	0,000	UP
	T125	0,000	UP	0,000	UP		T125	0,000	UP	0,000	UP
	T150	0,000	UP	0,000	UP		T150	0,000	UP	0,000	UP
Alvo 4	T25	0,000	UP	0,196		Alvo 6 raiz	T25	0,000	UP	0,000	UP
	T50	0,000	UP	0,000	UP		T50	0,000	UP	0,000	UP
	T75	0,004	UP	0,004	UP		T75	0,000	UP	0,000	UP
	T100	0,000	UP	0,000	UP		T100	0,000	UP	0,000	UP
	T125	0,000	UP	0,000	UP		T125	0,000	UP	0,000	UP
	T150	0,000	UP	0,000	UP		T150	0,000	UP	0,000	UP

3.7 NOVOS FATORES DE TRANSCRIÇÃO NA RESPOSTA AO DÉFICIT HÍDRICO

Dentre os diversos transcritos diferencialmente expressos em nossas bibliotecas subtrativas, obtivemos um grande número de sequências para fatores de transcrição. Dentre estes, identificamos 12 genes para fatores de transcrição, diferencialmente expressos nas bibliotecas de raízes da cultivar tolerante (Embrapa 48), descritos na literatura em diferentes processos biológicos, como: desenvolvimento, metabolismo, nodulação, silenciamento gênico, entre outros (Tabela 8). Estes fatores de transcrição, abaixo relacionados, não foram descritos, até o presente momento, na resposta ao déficit hídrico.

Tabela 8 - Novos Fatores de transcrição relacionados ao déficit hídrico em soja. Nome dos genes e organismo que apresentam maior similaridade (segundo análises por BLASTX e AutoFact) com os transcritos obtidos nas bibliotecas de raízes submetidas ao déficit hídrico. Função do fator de transcrição identificado segundo literatura e respectivas referências.

Fator de transcrição	Organismo	Função	Referências
AGL15	<i>Glycine max</i>	desenvolvimento embrionário	[48]
		expressão de chaperonas em resposta à baixas temperaturas	[49]
BASIC PENTACYSSTEINE 1	<i>Glycine max</i>	desenvolvimento ovariano	[50]
		ativação de chaperonas	[51]
HD-Zip III	<i>Populus trichocarpa</i> ; <i>Medicago truncatula</i>	desenvolvimento do meristema apical, tecido vascular e transporte de auxina	[52]
GATA transcription factor 3	<i>Arabidopsis thaliana</i> ; <i>Ricinus communis</i>	metabolismo de carbono e nitrogênio; transporte de açúcares	[53]
Transcription factor CPP	<i>Lotus japonicus</i>	nodulação	[54]
NIN-like protein 2	<i>Lotus japonicus</i>	nodulação	[55]
		controle estomático	[56]
Transcription factor DP1	<i>Populus tremula</i>	controle do ciclo celular, proliferação celular e apoptose	[57]
Apoptosis antagonizing transcription factor	<i>Oryza sativa</i>	anti-apoptose	[58]
BAH domain-containing protein	<i>Solanum lycopersicum</i>	silenciamento gênico	[59]
Transcription factor B3-EAR motif	<i>Arabidopsis thaliana</i>	silenciamento gênico	[60]
BEL1 homeotic protein	<i>Solanum</i> ; <i>Trifolium pratense</i> ; <i>Ricinus</i>	resposta a patógenos	[61]
		resposta à injúrias e desenvolvimento de tubérculos	[62]
DDT Homeo domain-related	<i>Medicago truncatula</i>	desconhecida	[63]

4 DISCUSSÃO

4.1 SEQUENCIAMENTO E ANÁLISES DE BIOINFORMÁTICA

A obtenção de perfis transcricionais é uma importante abordagem para entender o funcionamento de células, órgãos e do organismo como um todo. Embora neste tipo de estudo não seja possível prever precisamente o conjunto de proteínas que efetivamente serão expressas, assim como sua precisa atividade nas células, a natureza e abundância de transcritos possibilita identificar o *status* de sinalização e a maquinaria de resposta molecular, que são um indicativo das possíveis proteínas presentes no tecido sob a condição estudada [64], neste caso, o déficit hídrico.

O sequenciamento de clones de cDNA pelo método de Sanger não é capaz de catalogar todo o transcriptoma de eucariotos. Isto ocorre porque em um típico sequenciamento pelo método de Sanger, aproximadamente 20–40% dos transcritos não são devidamente sequenciados ou não são sequenciados, principalmente devido as etapas de clonagem requeridas. Isso porque estes

transcritos podem incluir aqueles em baixa abundância e aqueles que interferem no metabolismo do organismo utilizado para a clonagem, como, por exemplo, sequências para fatores de transcrição [65].

Por outro lado, os sequenciamentos de última geração (NGS, do inglês Next Generation Sequencing) são capazes de aproximar-se de uma cobertura mais abrangente dos cDNAs expressos em um determinado tecido/situação, devido ao fato de não requererem clonagem *in vivo* e serem capazes de sequenciar mais acuradamente a amostra de cDNA. O sequenciador *Genome Analyzer IIe*, utilizado no presente trabalho para o sequenciamento das bibliotecas subtrativas, é capaz de produzir milhões de *reads*, que variam em tamanho de 35 a 150 nucleotídeos [66].

Embora o sequenciamento por síntese (Illumina) seja capaz de sequenciar um maior número de cDNAs, a sequência completa do transcrito precisa ser montada a partir dos milhares de *reads* gerados. Devido ao pequeno tamanho das *reads*, a ancoragem das sequências em um genoma referência é de suma importância neste tipo de sequenciamento [28]. Esta ancoragem no genoma referência da soja possibilitou não só realizar a montagem dos *contigs* (Figura 2), como também relacionar as sequências obtidas com os genes preditos de soja (<http://www.phytozome.net/>), que forneceram informações sobre as regiões codante, 3'UTR (do inglês, *untranslated region*) e 5' UTR, imprescindíveis para as posteriores análises filogenéticas e a construção dos *primers* utilizados para quantificação da expressão por PCR quantitativo.

A utilização da ferramenta BLASTX permitiu identificar as prováveis proteínas codificadas pelos genes diferencialmente expressos, através da similaridade com sequências depositadas nos bancos de dados. Uma parte dos transcritos diferencialmente expressos não apresentou similaridade com proteínas depositadas no banco de dados (Tabela 2). Uma das explicações para este fato pode estar relacionada ao pequeno tamanho do fragmento gerado ou ainda ao fato das sequências corresponderem à região 3' UTR, que apresentam uma menor conservação entre os genes, reduzindo a probabilidade de encontrar similaridades. Entretanto, parte destes *no hits* pode corresponder a genes ainda não identificados, visto que a técnica de hibridização subtrativa é capaz de identificar genes raros, dificilmente detectados em bibliotecas comuns de cDNA, graças a etapas de normalização e amplificação das sequências [21, 67].

A maioria das sequências que apresentaram *hits* possui similaridade com sequências de soja (Figura 2A e B), visto que este foi o organismo estudado e cujo genoma já foi sequenciado, disponibilizando diversas sequências no banco de dados. Um grande número de sequências foi similar a outros organismos (Figura 2A e B), o que pode indicar que estas sequências ainda não foram descritas em soja ou ainda não foram depositadas no banco de dados utilizado. Sabe-se que a frequência de espécies nos resultados de BLAST depende do número de sequências que está depositado no banco de dados assim como com o grau de proximidade com o organismo estudado. A construção de uma árvore taxonômica com as espécies mais comuns do resultado de BLASTX com as sequências diferencialmente expressas nas bibliotecas de BR 16 e Embrapa 48 (Figura 2C), revela que a espécie mais próxima de soja é *Phaseolus vulgaris*, entretanto, a maioria dos resultados de BLASTX pertence a *Vitis vinifera* e *Ricinus comunis*. Isto ocorre provavelmente devido ao fato de no banco de proteínas do NCBI até a data de 16/09/2010 existirem apenas 1.813 sequências de *Phaseolus vulgaris* depositadas, contra 76.180 de *Vitis vinifera* e 62.996 de *Ricinus communis*.

4.2 DINÂMICA DOS PROCESSOS BIOLÓGICOS DURANTE O DÉFICIT HÍDRICO EM BR 16 E EMBRAPA 48

A categorização é uma importante ferramenta para a compreensão das alterações moleculares decorrentes da aplicação do déficit hídrico [68]. Como pode ser observado na Figura 3A, a cultivar BR 16 apresentou maior porcentagem de genes relacionados à regulação de processos biológicos nos períodos inicial (R1) e intermediário de estresse (R2), entretanto, observa-se declínio na porcentagem de genes desta categoria no período mais severo do estresse (R3). A categoria Regulação de processos biológicos, segundo o Gene Ontology, inclui processos que modulam a frequência ou taxa de determinado processo biológico, incluindo, por exemplo, o controle da expressão gênica, a modificação de proteínas ou interações proteína-substrato. Já a cultivar Embrapa 48 apresentou menor porcentagem de genes nesta categoria, entretanto, não houve diminuição neste tipo de processo biológico (Figura 3B), o qual apresentou pouca variação no decorrer da aplicação do estresse. Em BR 16 também se observa uma menor porcentagem de sequências relacionadas aos processos de sinalização e respostas a estímulos (Figura 3A), que aumenta apenas no tempo intermediário de estresse (R2), mas volta a cair no tempo

de estresse mais severo (R3). De maneira contrastante, em Embrapa 48 (Figura 3B), os processos de sinalização apresentam-se com maior porcentagem desde o período inicial do estresse, que se mantém no decorrer de todo o estresse; os genes da categoria de resposta a estímulos também aumentam nesta cultivar à medida que há aumento da severidade do estresse. Os processos de sinalização são essenciais para a resposta ao déficit hídrico visto que constituem uma cascata de eventos moleculares para transmitir a informação através de complexas vias, culminando na alteração da expressão de diversos genes [69].

Embora todas as plantas tenham permanecido em casa de vegetação com condições controladas, observou-se na cultivar Embrapa 48 o aparecimento de genes relacionados à Resposta a estímulos bióticos (Figura 3B). Tal fato pode ser explicado pela resposta cruzada entre diferentes tipos de estresses [70]. Membros da cascata de MAPKs, por exemplo, podem ser ativados por mais de um tipo de estresse, biótico ou abiótico, tais como frio, salinidade, injúria, ácido salicílico, peróxido de hidrogênio e ataque de patógenos [71].

Em Embrapa 48, observa-se também um aumento da classe de processos de Resposta a estresses na medida em que o déficit hídrico tornou-se mais severo (Figura 3B), enquanto em BR 16 observou-se justamente o contrário, ou seja, quanto mais severo o nível de estresse, menor o número de genes em resposta a esta condição (Figura 3A). Nesta classe estão inclusos genes relacionados a mudanças nas atividades celulares (secreção, produção de enzimas, expressão gênica, etc) em decorrência da perturbação do estado de homeostase por um estímulo exógeno, neste caso, principalmente o déficit hídrico.

Outro ponto interessante nesta análise diz respeito aos genes relacionados a processos de morte celular e senescência, que na cultivar BR 16 aparecem em alta porcentagem logo no tempo inicial de estresse (Figura 3A), o que não ocorre em Embrapa 48 (Figura 3B). Essa alta porcentagem de genes envolvidos nos processos de morte celular e senescência pode estar relacionado ao acúmulo de EROs (Espécies Reativas de Oxigênio) em consequência do déficit hídrico, que, em última instância, induzem à morte celular [72].

Com relação aos processos metabólicos, vemos em BR 16 um declínio nos processos de metabolismo de compostos nitrogenados (Figura 3A). Nesta classe incluem-se principalmente os processos de assimilação de compostos nitrogenados (fotossíntese), essenciais para os processos bioenergéticos. Sabe-se

que a fixação biológica de nitrogênio (N_2) nas raízes é um processo extremamente vulnerável ao déficit hídrico [73]. Com o declínio da fixação de N_2 há uma redução no suprimento de N_2 da planta para síntese protéica dificultando a manutenção do desenvolvimento de células e tecidos. Simultaneamente com o declínio nos processos de metabolismo de compostos nitrogenados em BR 16, observa-se a diminuição dos mecanismos de oxirredução, que abrangem os processos de transporte de elétrons nas cadeias fotossintética e de respiração celular. Em Embrapa 48 o declínio do metabolismo de compostos nitrogenados não é acentuado (Figura 3B) e os processos de oxido-redução não são afetados como em BR 16.

De maneira interessante, na condição de déficit hídrico, a categoria Processos biossintéticos abrangeu a maior porcentagem de transcritos em ambas as cultivares (Figura 3A e B). Esta pode ser uma evidência das necessidades celulares de biossintetizar proteínas e moléculas em resposta ao déficit hídrico. Isto porque o déficit hídrico induz a expressão de diversos genes, o que, na maioria das vezes, culminará na biossíntese das respectivas proteínas, que serão efetoras de algum tipo de resposta ao estresse, a exemplo de chaperonas, aquaporinas, proteínas LEA, proteínas de atividade antioxidante e substâncias que participam do ajuste osmótico, as quais abordaremos mais à frente.

4.3 GENES DIFERENCIALMENTE EXPRESSOS E SUA RELAÇÃO COM O DÉFICIT HÍDRICO

Em uma visão mais pontual, os genes estresse-responsivos encontrados em nossas bibliotecas podem ser divididos em dois grandes grupos. O primeiro consiste em genes relacionados aos processos de sinalização, como genes para proteínas quinases e fatores de transcrição. O segundo grupo é formado por genes *downstream* relacionados à tolerância ao déficit hídrico, como aqueles correspondentes a proteínas chaperonas, enzimas relacionadas ao estresse oxidativo, aquaporinas e outras ligadas a processos de desidratação como proteínas LEA, por exemplo [74].

Como visto anteriormente, os processos de sinalização são essenciais para o desencadeamento das respostas ao déficit hídrico [69]. No presente trabalho foram identificadas diversas sequências correspondentes a proteínas de sinalização (Tabela 3), que já foram relacionadas ao déficit hídrico em diversos trabalhos, sendo as principais: proteínas quinases, dentre elas

MAPKinases, CDPKs, CBLs e calmodulinas [75], proteínas GTP-binding [76] e WD40 repeat-containing proteins [77].

Com relação aos genes *downstream* na resposta ao estresse, encontrou-se um grande número de sequências codificadoras de proteínas chaperonas, envolvidas na proteção contra a desnaturação protéica [78], dentre elas proteínas de choque térmico (Heat shocks proteins), DNAJ e calreticulinas (Tabela 3).

A importância deste tipo de informação pode ser avaliado em trabalhos onde a superexpressão de genes que codificam proteínas de choque térmico, DNAJ e calreticulinas aumentam a tolerância de plantas a estresses abióticos [79, 80, 81]. Uma das consequências do déficit hídrico é o estresse oxidativo causado pela geração de EROs. Diversas proteínas com atividade antioxidante têm sido descritas na remoção de EROs como Superóxido dismutase, Peroxidases e Catalases [10]. Visto que as EROs podem oxidar múltiplos componentes celulares, como membranas, organelas e ácidos nucleicos, podendo levar, em última instância à morte celular, a identificação dos transcritos que codificam para estas proteínas nas bibliotecas subtrativas (Tabela 3) confirma a presença dos mecanismos de defesa contra o estresse oxidativo nas cultivares BR 16 e Embrapa 48 submetidas ao déficit hídrico, como observado por Khanna-Chopra e Selote [72].

Outra importante classe de proteínas relacionadas ao déficit hídrico são as aquaporinas, proteínas essenciais para o transporte de água em plantas. Embora a diminuição da expressão de genes para estas proteínas tenha sido relacionada ao déficit hídrico [82], estudos relatam a ativação destes genes em resposta a este tipo de estresse [83, 84]. Coccozza et al. [84] propuseram duas estratégias contrastantes para a expressão de genes codificadores de aquaporinas durante o déficit hídrico. A primeira é a regulação negativa da expressão destes genes, o que aumentaria a conservação da água e, logo, o turgor celular. A segunda estratégia envolve o aumento da expressão dos genes, que resultaria no aumento da permeabilidade da membrana celular, facilitando o transporte entre células. Segundo os autores, se expressos no sistema vascular de raízes, genes de aquaporinas poderiam otimizar o transporte de água para os tecidos da parte aérea da planta, melhorando sua condição hídrica frente ao estresse [84]. No presente trabalho, esta segunda estratégia parece ser mais atuante, uma vez que diversos

genes relacionados à aquaporinas mostraram-se diferencialmente expressos (*up-regulated*) nas plantas BR 16 e Embrapa 48 em resposta a aplicação do déficit hídrico (Tabela 3).

Os genes codificadores de proteínas relacionadas à desidratação também são amplamente descritos na resposta ao déficit hídrico (Tabela 3). Genes relativos às proteínas LEA, por exemplo, foram inicialmente descritos na proteção durante a maturação de embriões em sementes de angiospermas, sendo posteriormente relacionados à aquisição de maior tolerância ao déficit hídrico [85, 86]. Os genes *ERD* (early responsive to dehydration) também desempenham um importante papel nos estádios iniciais da aclimatação de planta à condição de déficit hídrico. Em 2007, LI e seus colaboradores [87] também utilizaram a construção de bibliotecas subtrativas para identificar genes relacionados ao déficit hídrico em milho, e identificaram o gene *ERD4*. Dois anos mais tarde o mesmo grupo demonstrou que a superexpressão deste gene em *Arabidopsis* aumentou a tolerância ao déficit hídrico e salino [88]. O acúmulo de desidrinas, também diferencialmente expressas no presente trabalho, tem sido positivamente correlacionado a tolerância à desidratação em plantas [89, 90].

Em decorrência do déficit hídrico, as plantas acumulam ABA e essa mudança na concentração do ABA celular desencadeia a ativação de numerosas mudanças na transcrição de diversos genes e no controle do fechamento estomático [11]. O papel da enzima *9-cis-epoxycarotenoid dioxygenase*, codificada pelo gene *NCED*, foi descrito como essencial na biossíntese de ABA durante o déficit hídrico [91]. Neste estudo a diminuição da condutância estomática revelada pelas medidas fisiológicas realizadas (Anexo 2B), evidencia o fechamento estomático, enquanto sequências para o gene *NCED* aparecem diferencialmente expressa (Tabela 3), demonstrando concordância entre os dados fisiológicos e moleculares.

4.4 FATORES DE TRANSCRIÇÃO DESCRITOS NA RESPOSTA AO DÉFICIT HÍDRICO

A expressão de genes estresse-responsivos é regulada por fatores de transcrição específicos [11]. A análise das bibliotecas subtrativas possibilitou identificar fatores de transcrição previamente relacionados ao déficit hídrico como as proteínas *NAC* [92], *MYB* [93], *WRKY* [94] e *DREB* [95] (Tabela 4). Visto que estes fatores estão no topo da cascata de resposta ao estresse, estudos mais

aprofundados sobre seu perfil de expressão têm se tornado cada vez mais importantes.

Os fatores de transcrição da subfamília *DREB* pertencem a grande família de proteínas AP2/EREBP (*APETALA2/ethylene responsive element binding protein*). Os fatores reguladores da transcrição codificados pelos genes da família AP2/EREBP compartilham um ou dois domínio(s) AP2/EREBP e desempenham várias funções envolvidas no desenvolvimento e nos processos fisiológicos das respostas das plantas ao estresse [96, 97]. Sendo assim, estudos no sentido de descobrir novos genes desta família em resposta ao déficit hídrico são de grande interesse.

Neste trabalho foram detectadas 20 sequências de genes candidatos a codificadores de proteínas da família AP2/EREB, similares a sequências de soja, *Pisum sativum*, *Populus trichocarpa*, *Ricinus communis*, *Arabidopsis thaliana* e *Caragana korshinskii* (Tabela 5). Análises dessas sequências por BLAST de nucleotídeos no banco de dados Phytozome permitiu identificar os *genemodels* ou “genes preditos” correspondentes. O alinhamento e construção de árvore filogenética destes genes preditos (Figura 4) levou a formação de subgrupos relacionados à sequências de *DREB* descritos em soja, e outras plantas na mesma família (*Fabaceae*) o que pode indicar homologia de função entre essas sequências, como sugerido por OH et al. [16] em análises filogenéticas de genes *AP2* em arroz. Sendo estes genes fortes candidatos a genes AP2/EREB, alguns ainda não descritos em soja, a etapa seguinte foi a confirmação da expressão diferencial destes genes em resposta ao déficit hídrico, via PCR quantitativo.

4.5 VALIDAÇÃO DAS BIBLIOTECAS SUBTRATIVAS E AVALIAÇÃO DA EXPRESSÃO DE GENES DA FAMÍLIA AP2/EREBP

Para validação dos resultados das bibliotecas subtrativas e melhor caracterização da resposta dos genes candidatos da família AP2/EREB aos diferentes níveis de estresse, nas cultivares BR 16 e Embrapa 48, foram realizadas reações de PCR quantitativo com *primers* específicos para estes genes.

Dentre os 6 genes candidatos selecionados para esta análise (Tabela 6), todos tiveram sua indução durante o déficit hídrico confirmada, em ambas as cultivares (Figura 5 e 6). O teste estatístico realizado pelo programa REST

permitiu-nos confirmar que as diferenças entre a expressão gênica das plantas tratadas em relação às plantas controle eram significativas e não devido ao acaso. Este teste é feito através da aleatorização, ou seja, da troca entre valores de expressão de ambos os grupos (tratado e controle), para verificar se há diferenças na expressão em comparação com o grupo inicial. O teste de hipóteses realiza 10.000 realocações entre os grupos “tratado” e “controle”, e avalia se a expressão relativa é alterada [45]. Como pode ser visto na Tabela 7, todos os genes foram induzidos durante o déficit hídrico e, na maioria dos tempos avaliados, a chance deste resultado ser devido ao acaso é menor do que 5%, o que demonstra a confiabilidade dos dados obtidos.

O gene alvo 1 (Glyma10g2) foi identificado em bibliotecas de tecidos foliares da cultivar BR 16 [47], e, por este motivo, realizamos análises de PCR quantitativo para este gene especificamente neste tecido. Nossas análises revelaram que este gene foi induzido durante todos os períodos de tratamento de déficit hídrico (0 a 150 min.) na cultivar BR 16 (Figura 5), e nos tempos 25, 75 e 100 min. na cultivar Embrapa 48 (Figura 6). Este gene codifica para uma proteína desconhecida de soja (Tabela 5), que, segundo seus domínios, pertence à subfamília DREB ou ERF. Estas duas famílias tem sido amplamente descritas na resposta ao déficit hídrico [11, 98], o que corrobora com nossos resultados, que evidenciam a participação deste gene na resposta ao déficit hídrico nestas duas cultivares (BR 16 e Embrapa 48).

O gene alvo 2 (Glyma14g09), identificado em nossas bibliotecas de tecido radicular, mostrou-se induzido durante todo o tratamento de déficit hídrico na cultivar mais tolerante (Embrapa 48) (figura 6, Tabela 7), enquanto em BR 16 sua expressão foi significativa apenas a partir do tempo 75 min. (figura 5, Tabela 7). Segundo análises de similaridade (Tabela 5) e filogenia (figura 4), este gene codifica para uma proteína similar a *DREB1* de soja. Os fatores de transcrição *DREB1* foram amplamente descritos nas respostas ao déficit hídrico. Diversos genes alvo destes fatores de transcrição já foram identificados, sendo que a maioria possui em seus domínios o motivo DRE, com a sequência conservada (A/G)CCGACNT, e, segundo Shinozaki e Yamaguchi-Shinozaki [11], seus produtos devem estar envolvidos nos mecanismos de tolerância ao déficit hídrico. Visto a grande importância deste fator de transcrição na resposta ao déficit hídrico, a indução mais tardia deste gene em BR 16 pode estar envolvida na maior sensibilidade desta cultivar ao déficit hídrico.

Ainda com relação aos genes analisados em raízes, o gene alvo 3 (Glyma04g1), que codifica para uma proteína similar a *DREB3* de soja (figura 4, Tabela 5) mostrou-se induzido durante a maioria dos tempos de estresse avaliados (Figura 5 e 6). Este fator de transcrição foi identificado recentemente neste organismo, e sua expressão foi induzida em resposta à baixas temperaturas [99]. No trabalho anteriormente citado, os autores reportam que a superexpressão deste gene em plantas transgênicas é capaz de aumentar sua tolerância ao déficit hídrico, entretanto, análises da expressão deste gene em plantas não transformadas não foram capazes de comprovar sua indução em resposta ao déficit hídrico. Nossos resultados, de maneira complementar, comprovam a indução deste fator de transcrição em resposta ao déficit hídrico, nas cultivares de soja BR 16 (Figura 5) e Embrapa 48 (Figura 6).

Os genes alvo 5 (Glyma13g01) e 6 (Glyma13g3) foram identificados em diferentes bibliotecas de folhas e raízes de ambas as cultivares (BR 16 e Embrapa 48), tornando interessante a análise de sua expressão nestes dois tecidos. As análises de similaridade (Tabela 5) e filogenia (figura 4) indicam que o gene alvo 5 codifica para uma proteína similar a um fator de transcrição *DREB2* descrito em *Caragana korshinskii* (*CkDBF*), uma planta adaptada à condições de baixa disponibilidade de água, tipicamente encontrada em regiões desérticas da China [100]. De maneira interessante, em tecidos foliares, houve uma maior expressão deste gene na cultivar mais tolerante (Embrapa 48) (figura 6) em relação à cultivar menos tolerante (BR 16) (figura 5). Esta diferença ocorreu principalmente em dois momentos: no tempo 50 min., onde a expressão relativa deste gene em BR 16 e Embrapa 48 foi de 1,7 e 5,5, respectivamente, e no tempo de estresse 75 min., no qual a expressão relativa de BR 16 e Embrapa 48 foi de 1,4 e 5,7, respectivamente (figura 5 e 6). A diferença na expressão deste gene pode ser parte de algum dos mecanismos responsáveis pela diferença da tolerância ao déficit hídrico entre estas duas cultivares, uma vez que, em tabaco, a superexpressão do gene *CkDBF* foi capaz de ativar a expressão de diversos genes estresse-responsivos, resultando no aumento da tolerância ao déficit hídrico [100].

O gene alvo 6 codifica para *DREB5* de soja. Este gene está depositado no banco de dados do NCBI, mas maiores informações sobre sua função ainda não foram publicadas pelos autores (*Unpublished data*). Em nossos estudos, a análise da expressão deste gene demonstra um grande contraste entre os tecidos

avaliados, sendo que em folhas a expressão foi significativamente maior do que em raízes, para ambas as cultivares (figura 5 e 6). Este é um indicativo da diferença da resposta deste fator de transcrição nos diferentes tecidos sob déficit hídrico, semelhante ao descrito por WANG e seus colaboradores [100], que observaram um grande aumento no nível de transcritos para um gene *DREB* (*CkDBF*) em folhas, após 4 horas de exposição das plantas à desidratação, enquanto em raízes o nível de transcritos, após o mesmo período, foi moderado.

4.6 NOVOS FATORES DE TRANSCRIÇÃO NA RESPOSTA AO DÉFICIT HÍDRICO

O genoma da soja codifica 5.671 possíveis fatores de transcrição, distribuídos em 63 diferentes famílias. Este número representa 12,2% dos 46.430 genes preditos, correspondendo a cerca de 8–10% de cada cromossomo [27]. Embora diversos destes fatores de transcrição já tenham sido relacionados ao déficit hídrico, como por exemplo, os fatores *NAC* [92], *MYB* [93], *WRKY* [94] e *DREB* [95], a complexidade da resposta ao déficit hídrico sugere a participação de outros fatores reguladores da transcrição, ainda não descritos.

O fator de transcrição AGL15 (Agamous-like 15) participa do desenvolvimento de embriões somáticos e da germinação em plantas (Tabela 8). Em *Arabidopsis thaliana*, AGL15 promove a expressão do gene *At1g02400*, que codifica uma GA 2-oxidase responsável pelo catabolismo de giberilina, e viabiliza o desenvolvimento embrionário e a germinação [48]. Este controle da biossíntese e degradação de hormônios pode ter alguma função na resposta ao déficit hídrico, à exemplo do mecanismo de ação do ABA. Adicionalmente, um trabalho mais recente demonstra que AGL15 liga-se ao promotor de genes que codificam chaperonas "Cold shock domain proteins" (CSPs) em *Arabidopsis* em resposta à baixa temperatura [49].

O fator de transcrição *BASIC PENTACYSTEINE1* é uma proteína de ligação ao motivo GA ("GAGA-binding protein"), promovendo o remodelamento da cromatina, alterando a expressão de genes (Tabela 8), dentre eles genes de *Arabidopsis thaliana* responsáveis pelo controle do desenvolvimento ovariano. O grande número de genes no genoma das plantas com promotores ricos em motivos GAGA sugere que estas proteínas podem afetar a expressão de uma grande variedade de processos [50], alguns deles talvez implicados na resposta ao déficit hídrico em plantas. Em animais, semelhantemente a AGL15, proteínas de ligação ao

GAGA ativam chaperonas, mais especificamente Hsp70 (70 kDa heat shock proteins) [51]. As proteínas Hsp70 são altamente conservadas; por esse motivo, talvez este tipo de fator de transcrição possua a mesma função em plantas. Como vimos anteriormente, as chaperonas desempenham um importante papel na estabilização de proteínas durante o déficit hídrico [78]. Em nossas bibliotecas subtrativas identificamos diversas chaperonas diferencialmente expressas durante o déficit hídrico (Tabela 3), o que pode ser um indício da participação destes fatores de transcrição na ativação da sua expressão.

De maneira geral, os membros subfamília III dos fatores de transcrição HD-Zip, estão bem caracterizados funcionalmente como reguladores do desenvolvimento do meristema apical, transporte de auxina e reguladores do desenvolvimento vascular (Tabela 8). Diferenciam-se dos membros das subfamílias I e II (cujos membros estão envolvidos na resposta ao déficit hídrico) principalmente por possuírem, além dos domínios comuns homeodomínio e zíper de leucina, os domínios MEKHLA, SAD e START [52]. Além de compartilharem os principais domínios das proteínas das subfamílias I e II (homeodomínio e zíper de leucina), outra evidência da possível participação destes fatores de transcrição no déficit hídrico é a implicação de alguns de seus membros no desenvolvimento de tecidos vasculares [52], que pode desempenhar uma importante função nos processos de aclimação às condições de déficit hídrico.

Por sua vez, os fatores de transcrição da família GATA são descritos no metabolismo de carbono (C), controle de genes responsáveis pelo transporte de açúcares, e controle de genes relacionados ao metabolismo de dissacarídeos e polissacarídeos (Tabela 8). Estão envolvidos também na sinalização via glicose e no metabolismo de nitrogênio. Como visto anteriormente, uma grande parte dos transcritos identificados nas bibliotecas subtrativas inserem-se na categoria Processos biossintéticos (Figura 3), possivelmente devido às necessidades celulares de biossintetizar proteínas e moléculas em resposta ao déficit hídrico. Neste contexto, o metabolismo de carbono e nitrogênio pode desempenhar um importante papel na célula sob déficit hídrico, visto a demanda energética necessária para os processos de resposta ao estresse. Por outro lado, o transporte de açúcares também pode ser implicado em processos como o ajuste osmótico, que consiste no acúmulo ativo de solutos, dentre eles açúcares como a trehalose [8], e tem por

finalidade diminuir o potencial osmótico da célula, contribuindo para manutenção do turgor celular.

Os fatores de transcrição semelhantes à proteína NIN2 e CPP foram descritos nos processos de associação das raízes com organismos simbiontes fixadores de nitrogênio (Tabela 8). Em *Medicago truncatula*, uma proteína NIN participa de transdução de sinais para a nodulação ativada via proteína quinase dependente de Cálcio/Calmodulina (CCaMK) [55]. Como já abordamos neste trabalho, as quinases dependentes de cálcio (CDPKs) também participam na transdução de sinais e desencadeamento de respostas no déficit hídrico, sendo que, nas bibliotecas subtrativas identificamos um grande número de genes para CDPKs diferencialmente expressos (Tabela 3). Logo, o transcrito correspondente ao fator de transcrição *NIN* identificado no presente trabalho pode estar implicado nesta via de transdução de sinais. Mais recentemente, CASTAINGS e seus colaboradores [56] identificaram um tipo diferente de proteína semelhante à NIN em *Arabidopsis thaliana*, denominada *proteína semelhante à NIN7* (NLP7). Segundo os autores, este fator de transcrição é capaz do controle da abertura estomática, sendo relacionado à tolerância/sensibilidade ao déficit hídrico. Com relação aos genes da Família CCP, embora se saiba que eles codificam proteínas amplamente distribuídas em plantas, e tenham sido descritos como moduladores do processo de nodulação em *Lotus*, desconhece-se o mecanismo preciso de ação deste tipo de gene e, por conseguinte, sua implicação no déficit hídrico [54].

Membros da família E2F, os fatores de transcrição *DP1* participam principalmente dos processos de controle do ciclo celular (Tabela 8). Em humanos, formam heterodímeros que regulam o ciclo celular durante a transição da fase G1 para S, via regulação de genes alvo [101]. Em plantas, foram descritos em diversos processos ligados ao ciclo celular, dentre eles a apoptose em resposta a danos no DNA [57]. Como abordamos anteriormente, algumas sequências diferencialmente expressas nas bibliotecas subtrativas relacionam-se aos processos de senescência/morte celular (Figura 3), provavelmente porque durante o déficit hídrico a geração de espécies reativas de oxigênio pode levar a oxidação de diversas moléculas e componentes da célula, incluindo o DNA. Por conseguinte, estes danos seriam passíveis de desencadear respostas apoptóticas, como a mediada pelo fator de transcrição *DP1*.

Em um mecanismo contrário a esta resposta, identificamos transcritos relativos a um fator de transcrição antagonista à apoptose (*AATF*- do inglês *Apoptosis-antagonizing transcription factor*) (Tabela 8). Em animais, a super-expressão de *AATF* resultou em um aumento significativo da atividade anti-apoptótica, enquanto o *knockout* por meio de pequenos RNAs interferentes causou resposta exacerbada de morte celular após aplicação de estresse por hipoxia [102]. Embora o papel de *AATF* em plantas não tenha sido bem esclarecido, uma proteína antagonista de apoptose (*API5*) age como fator-antiapoptótico em humanos, e estudos revelam que ela possui diversos domínios comuns com proteínas preditas em plantas [58], o que sugere que estas também possuam mecanismos similares de defesa ou controle da apoptose, que pode ser uma das consequências do déficit hídrico.

Durante o déficit hídrico, enquanto alguns genes de defesa são induzidos, outros tem sua expressão silenciada ou reduzida. Em nossas bibliotecas subtrativas, identificamos dois fatores de transcrição ligados aos processos de silenciamento da expressão: *BAH* e *B3-EAR* (Tabela 8). Os fatores de transcrição *BAH* são proteínas que possuem domínios adjacentes de bromo (*BAH*), que interagem com lisinas acetiladas das histonas, alterando a expressão gênica por remodelagem da cromatina. Em humanos, há evidências de que estes fatores de transcrição agem como silenciadores, recrutando para os genes alvo uma série de proteínas que coordenam a formação de heterocromatina [59]. Outra evidência da função de silenciamento destes fatores é a presença de domínios adjacentes de bromo (*BAH*) em *METHYLTRANSFERASE 1* de *A. thaliana*. [103]. Adicionalmente em *Saccharomyces cerevisiae* mutações no domínio *BAH* de proteínas de silenciamento (*Sir3*) levaram à perda da função repressora [104]. De maneira similar, os fatores de transcrição *B3-EAR* também têm sido descritos no silenciamento gênico. Pertencem à família *VP1/ABI3* e, segundo estudos em *Arabidopsis*, são requeridos na repressão de genes embrionários, necessária para um desenvolvimento normal da planta durante a germinação [60].

Os fatores de transcrição que codificam proteínas homeóticas *BEL1/BELL1* (Tabela 8) são amplamente descritos no desenvolvimento de flores, frutos e tubérculos. Todavia, além desta função de desenvolvimento, a expressão do gene *BELL* de arroz em tabaco, provocou a ativação de genes de resposta a patógenos [61]. Concordantemente, em batata (*Solanum tuberosum*), um gene

BELL1 foi induzido em resposta à injúrias impostas a ramos e folhas [62]. Sabe-se que estresses bióticos e abióticos compartilham mecanismos de resposta. Em uma simples comparação, durante os estresses bióticos, como a herbivoria, danos e alterações da membrana celular agem como sinalizadores do estresse, ativando diversos genes. Durante o déficit hídrico, a diminuição do turgor celular também é percebida via alterações na membrana, que podem levar a ativação de genes comuns aos dois tipos de estresse, em um mecanismo denominado *cross-talk* [70].

5 CONCLUSÃO

Através da construção de bibliotecas subtrativas e do emprego de análises de bioinformática, identificamos diversos genes induzidos pelo déficit hídrico. Verificamos que ambas as cultivares estudadas (BR 16 e Embrapa 48) expressaram transcritos relacionados à resposta ao déficit hídrico, tanto na via de sinalização (transcritos para MAPKinases, CDPKs, CBLs e calmodulinas), proteínas GTP- binding e WD40) quanto na via efetora, como proteção contra a desnaturação protéica (Heat shocks proteins, DNAJ e calreticulinas), proteínas de atividade antioxidante (Peroxidases e Catalases) e aquaporinas. A categorização de todos os genes identificados, segundo os processos biológicos envolvidos, possibilitou uma visão geral das alterações celulares em resposta a este tipo de estresse. Através desta visão, constatamos que as cultivares avaliadas, BR 16 e Embrapa 48, apresentaram diferença nos perfis transcricionais de raízes submetidas ao déficit hídrico. Identificamos diversos fatores de transcrição pertencentes à família AP2/EREBP, que demonstraram proximidade filogenética com genes *DREB* de Fabaceae. A avaliação da expressão de 6 candidatos confirmou sua indução pelo déficit hídrico, validando o resultado obtido nas bibliotecas subtrativas. Entretanto, estudos mais aprofundados a cerca dos domínios das proteínas codificadas são requeridos, para que se possa afirmar que estes genes se tratam de novos fatores de transcrição *DREB*. Identificamos também diferentes classes de fatores de transcrição diferencialmente expressos em resposta ao déficit hídrico. Uma vez que diversos destes fatores de transcrição não foram relacionados à resposta ao déficit hídrico até o presente momento, sua identificação e a especulação sobre seu possível papel neste tipo de estresse foi uma importante contribuição deste trabalho. Todavia, faz-se necessária a confirmação e quantificação da expressão diferencial

destes genes neste tipo de estresse, pelo emprego de PCR quantitativo, por exemplo. Adicionalmente, a superexpressão ou inibição destes fatores de transcrição em plantas geneticamente modificadas poderá demonstrar sua real função em resposta à desidratação celular. Além dos fatores de transcrição mencionados, também identificamos diversos transcritos similares à proteínas desconhecidas (*unknown proteins*), algumas delas assinaladas com a função de regulação da transcrição. Estes fatores de transcrição merecem atenção em estudos posteriores para melhor caracterização.

6 REFERÊNCIAS BIBLIOGRÁFICAS

- [1] FAO, 2008. Disponível em: <<https://www.fao.org.br/default.asp>> Acesso em Dez 2008.
- [2] CONAB 2010. Acompanhamento de safra brasileira: grãos, sexto levantamento, março 2010. Disponível em <http://www.conab.gov.br/conabweb/download/safra/6graos_09.03.10.pdf> Acesso em 31/03/2010, 15:40h, 2010.
- [3] PEDERSON, P. **Soybean growth and development**. Publ. PM-1945. Iowa State University Extension, 2004.
- [4] FARIAS, J.R.B.; NEUMAIER, N., NEPOMUCENO, A.L., 2009. **Seca: soja em carência de água**. Disponível em: < <http://bioinfo.cnpso.embrapa.br/seca/index.php/ecofisiologia/exigencias-climaticas>> Acesso em: 01/02/2011.
- [5] EMBRAPA, 2008. Disponível em: <<http://www.embrapa.br/imprensa/noticias/2008/junho/2a-semana/levantamento-quantifica-perdas-na-soja-no-rs/>> Acesso em: 17/08/2009.
- [6] CATTIVELLI, L. et al. Drought tolerance improvement in crop plants: An integrated view from breeding to genomics. **Field Crops Research**, v.105, p.1-14, 2008.
- [7] MELCHER, K. et al. Agate–latch–lock mechanism for hormone signalling by abscisic acid receptors. **Nature**, v.462, Dez 2009.
- [8] ALMEIDA, A.M. et al. Trehalose and its applications in plant biotechnology. **In Vitro Cellular & Developmental Biology – Plant**,v.43, p.167-177,2007.
- [9] MARTÍNEZ, J.P. et al. Is osmotic adjustment required for water stress resistance in the Mediterranean shrub *Atriplex halimus* L.? **Journal of Plant Physiology**, v.161, p.1041-1051, 2004.
- [10] CARVALHO, M.H.C. Drought stress and reactive oxygen species: Production, scavenging and signaling. **Plant Signaling & Behavior**, v.3, n.3, p.156-165, 2008.
- [11] SHINOZAKI K., YAMAGUCHI-SHINOZAKI, K. Gene networks involved in drought stress response and tolerance. **Journal of Experimental Botany**, v.58, p.221-227, 2007.
- [12] CHINNUSAMY, V., SCHUMAKER, K., ZHU, J. Molecular genetic perspectives on cross-talk and specificity in abiotic stress signaling in plants. **Journal of Experimental Botany**, v.55, p.225-236, 2004.
- [13] YAMAGUCHI-SHINOZAKI, K., SHINOZAKI, K. Organization of cis-acting regulatory elements in osmotic- and cold-stress responsive promoters. **Trends Plant Science**, v.10, p.88-94, 2005.
- [14] WANG, S. et al. An *in silico* strategy identified the target gene candidates regulated by dehydration responsive element binding proteins (DREBs) in *Arabidopsis* genome. **Plant Molecular Biology**,v.69, p.167-178, 2009.
- [15] ALMOGUERA, et al. The HaDREB2 transcription factor enhances basal thermotolerance and longevity of seeds through functional interaction with HaHSFA9. **BMC Plant Biology**, v.9, p.75, 2009.
- [16] OH, S.J. et al. Overexpression of the Transcription Factor AP37 in rice improves grain yield under drought conditions. **Plant Physiology**, v.150, p.1368-1379, 2009.
- [17] QIN, F.Y. et al. Cloning and functional analysis of a novel DREB1/CBF transcription factor involved in cold-responsive gene expression in *Zea mays* L. **Plant Cell Physiology**,v.45, p.1042-1052, 2004.

- [18] TIAN, A.G. et al. Characterization of soybean genomic features by analysis of its expressed sequence tags. **Theoretical and Applied Genetics**, v.108, p.903-913, 2004.
- [19] LI, X.P. et al. Soybean DRE-binding transcription factors that are responsive to abiotic stress. **Theoretical Applied Genetics**, v.110, p.1355-1362, 2005.
- [20] CHEN, M. et al. GmDREB2, a soybean DRE-binding transcription factor, conferred drought and high-salt tolerance in transgenic plants. **Biochemical and Biophysical Research Communications**, v.353, p.299-305, 2007.
- [21] DIATCHENKO, L. et al. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. **Proceedings Of The National Academy Of Sciences Of The U.S.A.**, v.93, p.6025-6030, 1996.
- [22] OYA, T. et al. The drought tolerance characteristics of the brazilian soybean cultivars. **Plant Production Science- Japão**, v.7, n.2, p.129-137, 2004.
- [23] HEWITT, E.J. **Nutrition of plants in culture media**. Ed. F. C. Steward. Academic Press, Vol III. New York. p.97-133, 1963.
- [24] FEHR, W.R., CAVINESS, C.E. **Stages of soybean development**. Ames: Iowa State University, 1977.
- [25] MARTINS, P.K. et al. Differential gene expression and mitotic cell analysis of the drought tolerant soybean (*Glycine max* L. Merrill Fabales, Fabaceae) cultivar MG/BR46 (Conquista) under two water deficit induction systems. **Genetics and Molecular Biology**, v.31, 2008.
- [26] AUSUBEL, F.M. et al. **Short Protocols in Molecular Biology**. New York: John Wiley & Sons, 1995.
- [27] SCHMUTZ, J. et al. Genome sequence of the palaeopolyploid soybean. **Nature**, v.463, 2010.
- [28] LI, R. et al. SOAP: short oligonucleotide alignment program. **Bioinformatics**, v.24, p.713-714, 2008.
- [29] HERNANDEZ, D. et al. De novo bacterial genome sequencing: Millions of very short reads assembled on a desktop computer. **Genome Research**, published online March 10, 2008.
- [30] ALTSCHUL, S.F. et al. Basic local alignment search tool. **Journal of Molecular Biology**, v.215, p.403-410, 1990.
- [31] KOSKI, L.B. et al. AutoFACT: An Automatic Functional Annotation and Classification Tool. **BMC Bioinformatics**, v.16, p.151, 2005.
- [32] WEIZHONG, L., JAROSZEWSKI, L., GODZIK, A. Clustering of highly homologous sequences to reduce the size of large protein databases. **Bioinformatics**, v.17, p.282-283, 2001.
- [33] OGATA, H. et al. KEGG: Kyoto Encyclopedia of Genes and Genomes. **Nucleic Acids Research**, v.27, p.29-34, 1999.
- [34] FINN, R.D. et al. The Pfam protein families database. **Nucleic Acids Research**, v.38, p.211-222, 2010.
- [35] SCHULTZ, J. et al. SMART: A Web-based tool for the study of genetically mobile domains. **Nucleic Acids Research**, v.28, p.231-234, 2000.
- [36] GÖTZ, S. et al. High-throughput functional annotation and data mining with the Blast2GO suite. **Nucleic Acids Research**, v.36, p.3420-3435, 2008.
- [37] CARBON, S. et al. AmiGO: online access to ontology and annotation data. **Bioinformatics**, v.25, p.288, 2009.
- [38] KUMAR, S. et al. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. **Briefings in Bioinformatics**, v.9, p.299-306, 2008.
- [39] THOMPSON, J.D., HIGGINS, D.G., GIBSON, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. **Nucleic Acids Research**, v.22, p.4673-4680, 1994.
- [40] NEI, M., KUMAR, S. **Molecular evolution and phylogenetics**. Oxford University Press, 2000.
- [41] EFRON, B., TIBSHIRANI, R. **An introduction to the bootstrap**. New York: Chapman & Hall, 1993.
- [42] VANDESOMPELE, J. et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. **Genome biology**, v.3, 2002.
- [43] ANDERSEN, C.L., JENSEN, J.L., ORNTOFT, T.F. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. **Cancer research**, v.64, p.5245-5250, 2004.
- [44] LIVAK, K.J., SCHMITTGEN, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. **Methods**, v.25, p.402-408, 2001.

- [45] PFAFFL, M.W., HORGAN, G.W., DEMPFLER, L. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in realtime PCR. **Nucleic Acids Research**, v.30, 2002.
- [46] KIM, C., LEMKE, C., PATERSON, A.H. Functional dissection of drought-responsive gene expression patterns in *Cynodon dactylon* L. **Plant Molecular Biology**, v.70, p.1-16.
- [47] AMARAL, L.C. **Análise de respostas fisiológicas e expressão diferencial de genes em soja submetida a déficit hídrico**. Dissertação (Mestrado em Genética e Biologia Molecular) – Universidade Estadual de Londrina, Londrina, 2010.
- [48] THAKARE, D. et al. The MADS-Domain transcriptional regulator AGAMOUS-LIKE15 promotes somatic embryo development in *Arabidopsis* and soybean. **Plant Physiology**, v.146, n.4, p.1663-1672, 2008.
- [49] NAKAMINAMI, K. et al. *Arabidopsis* cold shock domain proteins: relationships to floral and silique development. **Journal of Experimental Botany**, v.60, n.3, p.1047-1062, 2009.
- [50] KOOIKER, M. et al. BASIC PENTACYSSTEINE1, a GA binding protein that induces conformational changes in the regulatory region of the homeotic *Arabidopsis* gene *SEEDSTICK*. **The Plant Cell**, v.17, p.722-729, 2005.
- [51] TURTURICI, G. et al. Hsp70 is required for optimal cell proliferation in mouse A6 mesoangioblast stem cells. **Biochemical Journal**, v.421, p.193-200, 2009.
- [52] ARIEL, F.D. et al. The true story of the HD-Zip family. **Trends in Plant Science**, v.12, n.9, p.419-426, 2007.
- [53] BI, Y.M. et al. Genetic analysis of *Arabidopsis* GATA transcription factor gene family reveals a nitrate-inducible member important for chlorophyll synthesis and glucose sensitivity. **The Plant Journal**, v.44, p.680-692, 2005.
- [54] ASAMIZU, E. et al. A positive regulatory role for LjERF1 in the nodulation process is revealed by systematic analysis of nodule-associated transcription factors of *Lotus japonicas*. **Plant Physiology**, v.147, p.2030-2040, 2008.
- [55] MARSH, J.F. et al. *Medicago truncatula* NIN is essential for rhizobial-independent nodule organogenesis induced by autoactive calcium/calmodulin-dependent protein kinase1. **Plant Physiology**, v.144, p.324-335, 2007.
- [56] CASTANGS, L. et al. The nodule inception-like protein 7 modulates nitrate sensing and metabolism in *Arabidopsis*. **The Plant Journal**, v.57, p.426-435, 2009.
- [57] LAMMENS, T. et al. Atypical E2Fs: new players in the E2F DP1 transcription factor family. **Trends in Cell Biology**, v.19, n.3, p.111-118, 2009.
- [58] NEDELCO, A.M. Comparative genomics of phylogenetically diverse unicellular eukaryotes provide new insights into the genetic basis for the evolution of the programmed cell death machinery. **Journal of Molecular Evolution**, v.68, p.256-268, 2009.
- [59] BIERNEA, H. et al. Human BAHD1 promotes heterochromatic gene silencing. **Proceedings Of The National Academy Of Sciences Of The U.S.A.**, v.106, n.33, p.13826-13831, 2009.
- [60] SUZUKI, M., WANG, H.H.Y., MCCARTY, D.R. Repression of the LEAFY COTYLEDON 1/B3 Regulatory Network in Plant Embryo Development by VP1/ABSCISIC ACID INSENSITIVE 3-LIKE B3 Genes. **Plant Physiology**, v.143, p.902-911, 2007.
- [61] LUO, H. et al. Up-regulation of OsBIHD1, a rice gene encoding BELL homeodomain transcriptional factor, in disease resistance responses. **Plant Biology**, v.7, p.459-468, 2005.
- [62] CHATTERJEE, M., BANERJEE, A.K., HANNAPPEL, D.J. A BELL1-Like gene of potato is light activated and wound inducible. **Plant Physiology**, v.145, p.1435-1443, 2007.
- [63] MUKHERJEE, K., BROCCIERI, L., BÜRGLIN, T.R. A Comprehensive classification and evolutionary analysis of plant homeobox genes. **Molecular Biology and Evolution**, v.26, p.2775-2794, 2009.
- [64] POROYKO, V. Comparing regional transcript profiles from maize primary roots under well-watered and low water potential conditions. **Journal of Experimental Botany**, v.58, p.279-289, 2007.
- [65] BRENT, M.R. Steady progress and recent breakthroughs in the accuracy of automated genome annotation. **Nature Reviews Genetics**, v.9, p.62-73, 2008.
- [66] ILLUMINA 2010. Disponível em:
<http://www.illumina.com/technology/solexa_technology.ilmn> Acesso em: 23/09/2010.
- [67] FOUCCART, C. et al. Transcript profiling of a xylem vs phloem cDNA subtractive library identifies new genes expressed during xylogenesis in *Eucalyptus*. **New Phytologist**, v.170, p.739-752, 2006.

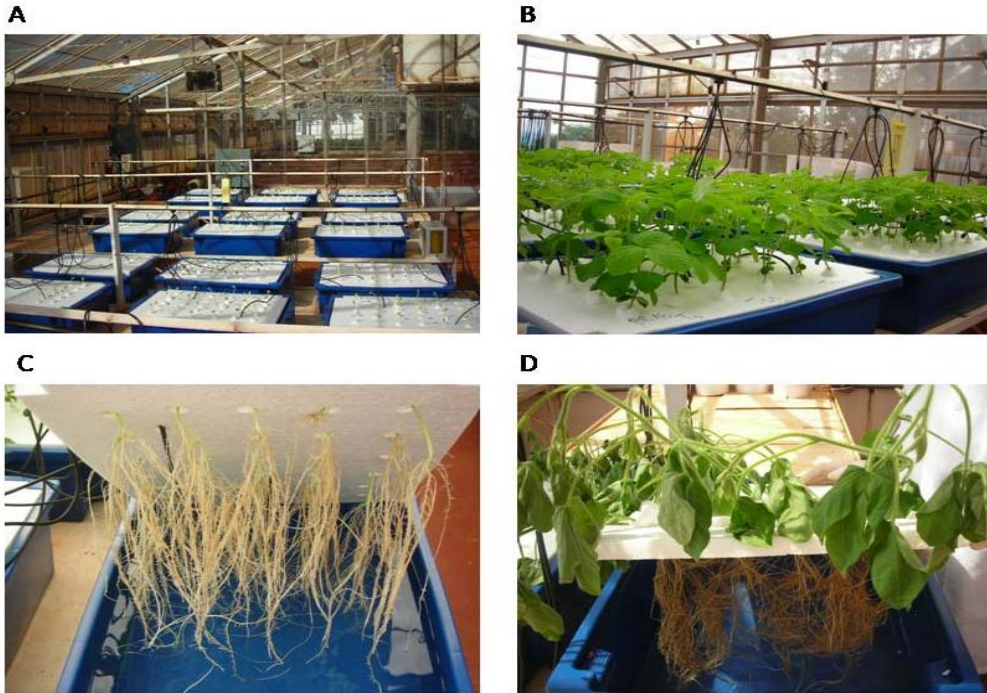
- [68] HARB, A. et al. molecular and physiological analysis of drought stress in *Arabidopsis* reveals early responses leading to acclimation in plant growth. **Plant Physiology**, v.1, 2010.
- [69] SEKI, M. et al. Regulatory metabolic networks in drought stress responses. **Current Opinion in Plant Biology**, v.10, p.296–302, 2007.
- [70] YOSHIOKA, K., SHINOZAKI, K. **Signal crosstalk in plant stress responses**. Wiley-Blackwell, p.23-32, 2009.
- [71] WANG, M. et al. A novel MAP kinase gene in cotton (*Gossypium hirsutum* L.), GhMAPK, is involved in response to diverse environmental stresses. **Journal of Biochemistry and Molecular Biology**, v.40, p.325-332, 2007.
- [72] KHANNA-CHOPRA, R., SELOTE, D.S. Acclimation to drought stress generates oxidative stress tolerance in drought-resistant than -susceptible wheat cultivar under field conditions. **Environmental and Experimental Botany**, v.60, n.2, 2007.
- [73] SINCLAIR, R.T. et al. Drought tolerance and yield increase of soybean resulting from improved symbiotic N₂ fixation. **Field Crops Research**, v.101, p.68–71, 2007.
- [74] VÁSQUEZ-ROBINET, C. et al. Differential expression of heat shock protein genes in preconditioning for photosynthetic acclimation in water-stressed loblolly pine. **Plant Physiology and Biochemistry**, v.48, p.256-264, 2010.
- [75] BATISTIČ, O., KUDLA, J. Plant calcineurin B-like proteins and their interacting protein kinases. **Biochimica et Biophysica Acta**, v.1793, p.985–992, 2009.
- [76] AGARWAL, P.K. et al. Constitutive overexpression of a stress-inducible small GTP-binding protein PgRab7 from *Pennisetum glaucum* enhances abiotic stress tolerance in transgenic tobacco. **Plant Cell Reports**, v.27, p.105–115, 2008.
- [77] LEE, M.S. et al. Detection of two alternative bcr/abl mRNA junctions and minimal residual disease in Philadelphia chromosome positive chronic myelogenous leukemia by polymerase chain reaction. **Blood**, v.73, p.2165–2170, 1989.
- [78] PEREIRA, J. et al. Crystal structures of a group II chaperonin reveal the open and closed states associated with the protein folding cycle. **The Journal of Biological Chemistry**, v.285, p.27958-27966, 2010.
- [79] MONTERO-BARRIENTOS, M. et al. Transgenic expression of the *Trichoderma harzianum* hsp70 gene increases *Arabidopsis* resistance to heat and other abiotic stresses. **Journal of Plant Physiology**, v.167, p.659–665, 2010.
- [80] ZHICHANG, Z. et al. Over-expression of *Arabidopsis* DnaJ (Hsp40) contributes to NaCl-stress tolerance. **African Journal of Biotechnology**, v.9, p.972-978, 2010.
- [81] JIA, X.Y. et al. Molecular cloning and characterization of wheat calreticulin (CRT) gene involved in drought-stressed responses. **Journal of Experimental Botany**, v.59, p.739-751, 2007.
- [82] SECCHI, F., LOVISOLO, C., SCHUBERT, A. Expression of OePIP2.1 aquaporin gene and water relations of *Olea europaea* twigs during drought stress and recovery. **The Annals of Applied Biology**, v.150, p.163–167, 2007.
- [83] MONTALVO-HERNÁNDEZ, L. et al. Differential accumulation of mRNAs in drought-tolerant and susceptible common bean cultivars in response to water deficit. **New Phytologist**, v.177, p.102–113, 2008.
- [84] COCOZZA, C. et al. Early effects of water deficit on two parental clones of *Populus nigra* grown under different environmental conditions. **Functional Plant Biology**, v.37, p.244–254, 2010.
- [85] XIAO, B. et al. Over-expression of a *LEA* gene in rice improves drought resistance under the field conditions. **Theoretical and Applied Genetics**, v.115, n.1, p.35-46, 2007.
- [86] DALAL, M. et al. Abiotic stress and ABA-inducible Group 4 *LEA* from *Brassica napus* plays a key role in salt and drought tolerance. **Journal of Biotechnology**, v.139, n.2, p.137-145, 2009.
- [87] LI, H.Y. et al. Isolation and characterization of induced genes under drought stress at the flowering stage in maize (*Zea mays*). **DNA Sequence**, v.18, p.445–460, 2007.
- [88] LIU, Y. et al. A maize early responsive to dehydration gene, ZmERD4, provides enhanced drought and salt tolerance in *Arabidopsis*. **Plant Molecular Biology Reports**, v.27, p.542–548, 2009b.
- [89] MOHAMMADKHANI, N., HEIDARI, R. Effects of drought stress on soluble proteins in two maize varieties. **Turk Journal of Biology**, v.32, p.23-30, 2008.
- [90] HU, L. et al. Differential accumulation of dehydrins in response to water stress for hybrid and common bermudagrass genotypes differing in drought tolerance. **Journal of Plant Physiology**, v.167, p.103–109, 2010.
- [91] ENDO, A. et al. Drought induction of *Arabidopsis* 9-cis-Epoxycarotenoid Dioxygenase occurs in vascular parenchyma cells. **Plant Physiology**, v.147, p.1984–1993, 2008.

- [92] ZHENG, X. et al. Overexpression of a NAC transcription factor enhances rice drought and salt tolerance. **Biochemical and Biophysical Research Communications**, v.379, p.985–989, 2009.
- [93] DAI, X. et al. Overexpression of an R1R2R3 MYB Gene, *OsMYB3R-2*, increases tolerance to freezing, drought, and salt stress in transgenic *Arabidopsis*. **Plant Physiology**, v.143, p.1739–1751, 2007.
- [94] GIACOMELLI, J.I. et al. Expression analyses indicate the involvement of sunflower WRKY transcription factors in stress responses, and phylogenetic reconstructions reveal the existence of a novel clade in the Asteraceae. **Plant Science**, v.178, p.398–410, 2010.
- [95] SAKUMA, Y. et al. Functional analysis of an *Arabidopsis* transcription factor, DREB2A, involved in drought-responsive gene expression. **The Plant Cell**, v.18, p.1292–1309, 2006.
- [96] CERNAC, A., BENNING, C. Wrinkled1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in *Arabidopsis*. **Plant Journal**, v.40, p.575–585, 2004.
- [97] LIN, R. et al. Molecular cloning and characterization of a rice gene encoding AP2/EREBP-type transcription factor and its expression in response to infection with blast fungus and abiotic stresses. **Physiological and Molecular Plant Pathology**, v.70, p.60–68, 2007.
- [98] ZHANG, G. et al. Overexpression of the soybean GmERF3 gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. **Journal of Experimental Botany**, v.60, p.3781–3796, 2009.
- [99] CHEN, M. et al. Cold-induced modulation and functional analyses of the DRE-binding transcription factor gene, GmDREB3, in soybean (*Glycine max* L.). **Journal of Experimental Botany**, v.60, n.1, p.121–135, 2009.
- [100] WANG, X. et al. A novel dehydration-responsive element-binding protein from *Caragana korshinskii* is involved in the response to multiple abiotic stresses and enhances stress tolerance in transgenic tobacco. **Plant Molecular Biology Reporter**, v.28, p.664–675, 2010.
- [101] MILTON, A. et al. A functionally distinct member of the DP family of E2F subunits. **Oncogene**, v.25, p.3212–3218, 2006.
- [102] XIE, J., GUO, Q. Apoptosis antagonizing transcription factor protects renal tubule cells against oxidative damage and apoptosis induced by ischemia-reperfusion. **Journal of American Society of Nephrology**, n.17, p.3336–3346, 2006.
- [103] HENDERSON, I., JACOBSEN, S.E. Epigenetic inheritance in plants. **Nature**, v.447, p.418–424, 2007.
- [104] CONNELLY, J.J. et al. Structure and function of the *Saccharomyces cerevisiae* Sir3 BAH domain. **Molecular and Cellular Biology**, v.26, n.8, p.3256–3265, 2006.

ANEXOS

ANEXO A

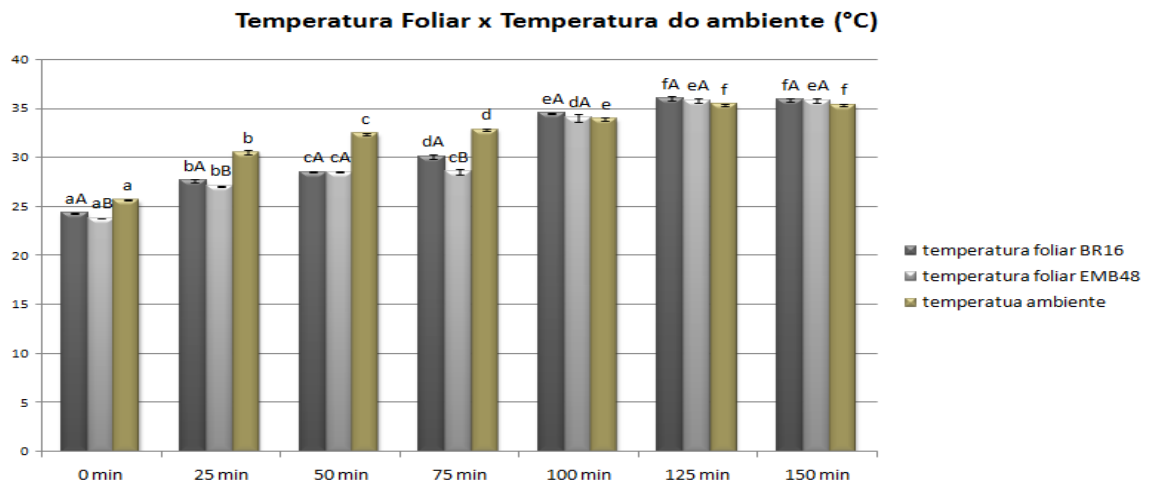
Experimento de déficit hídrico



Anexo A:(A) Plantas durante crescimento em hidroponia na casa de vegetação; (B) plantas em estágio V3 antes da aplicação do estresse. (C) retirada das solução hidropônica para aplicação do tratamento de déficit hídrico. (D) Plantas após aplicação do tratamento de déficit hídrico (150 min sob estresse).

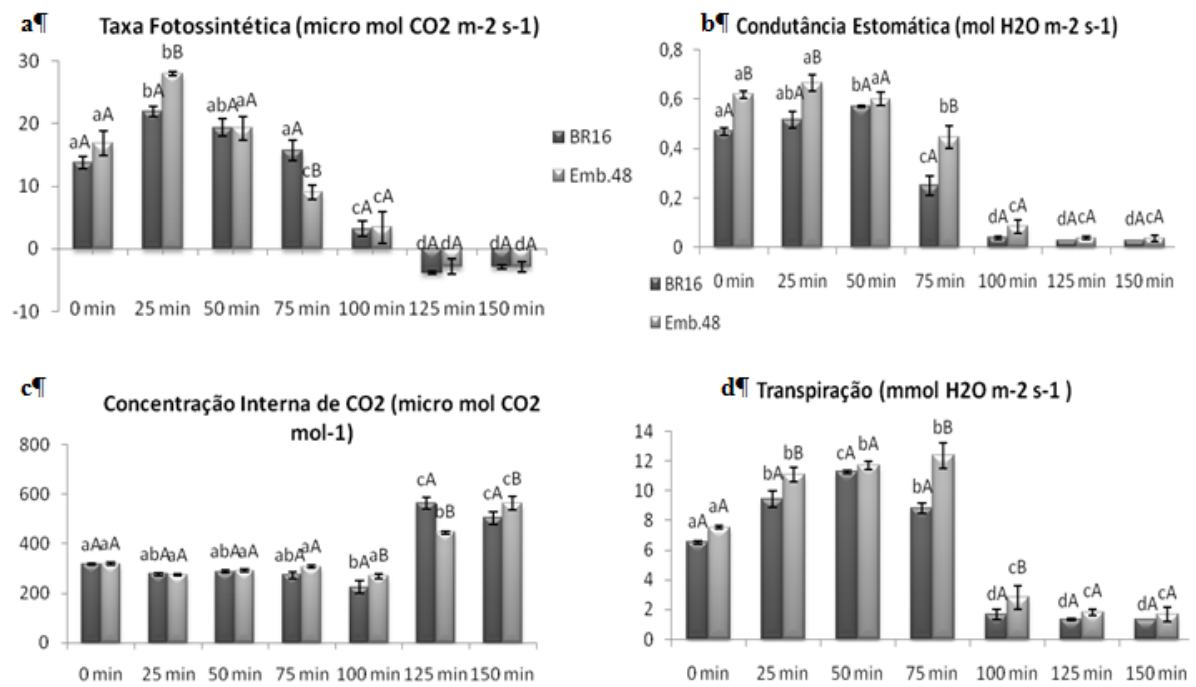
ANEXO B

Análises fisiológicas



Anexo B: A. Temperatura ambiente x Temperatura Foliar (°C) das Cultivares BR 16 e Embrapa48 submetidas ao déficit hídrico. Os valores apresentados são valores de teste de médias para tempo de aplicação de estresse e cultivar (Teste de Tukey com $p \leq 0,05$). Traços verticais entre as barras: erro padrão; letras maiúsculas: diferenças entre os cultivares dentro de cada tempo; letras minúsculas: diferenças entre os tempos dentro de cada cultivar; letras iguais não apresentam diferença estatística observadas pelo teste de Tukey

ANEXO B2



Anexo B 2: Parâmetros Fisiológicos das Cultivares BR 16 e Embrapa 48 submetidos ao déficit hídrico, os valores testes de médias para tempo de aplicação de estresse e cultivar (Teste de Tukey com $p \leq 0,05$). **(a)** taxa fotossintética (micro mol CO₂ m⁻² s⁻¹); **(b)** condutância estomática (mol H₂O m⁻² s⁻¹), **(c)** concentração interna de CO₂ (micro mol CO₂ mol⁻¹); **(d)** taxa de transpiração (mmol H₂O m⁻² s⁻¹). Traços verticais entre as barras: erro padrão; letras maiúsculas: diferenças entre os cultivares dentro de cada tempo; letras minúsculas: diferenças entre os tempos dentro de cada cultivar; letras iguais não apresentam diferença estatística observadas pelo teste de Tukey.