



UNIVERSIDADE  
ESTADUAL DE LONDRINA

---

SILVANA REGINA ROCKENBACH MARIN

**“IDENTIFICAÇÃO DE REGIÕES PROMOTORAS DE GENES  
EXPRESSOS EM SOJA SOB DÉFICIT HÍDRICO”**

---

Londrina  
2014

SILVANA REGINA ROCKENBACH MARIN

**“IDENTIFICAÇÃO DE REGIÕES PROMOTORAS DE GENES  
EXPRESSOS EM SOJA SOB 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. Renata Fuganti-Pagliarini*

Londrina  
2014

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

**Dados Internacionais de Catalogação-na-Publicação (CIP)**

M337i Marin, Silvana Regina Rockenbach.

Identificação de regiões promotoras de genes expressos em soja sob déficit hídrico / Silvana Regina Rockenbach Marin. – Londrina, 2014.  
137 f. : il.

Orientador: Alexandre Lima Nepomuceno.

Coorientador: Renata Fuganti-Pagliarini.

Dissertação (Mestrado em Genética e Biologia Molecular) – Universidade Estadual de Londrina, Centro de Ciências Biológicas, Programa de Pós-Graduação em Genética e Biologia Molecular, 2014.

Inclui bibliografia.

1. Soja – Genética molecular – Teses. 2. Soja – Condições hídricas – Teses. 3. Regiões promotoras (Genética) – Teses. 4. Engenharia de genética vegetal – Teses. I. Nepomuceno, Alexandre Lima. II. Fuganti-Pagliarini, Renata. III. Universidade Estadual de Londrina. Centro de Ciências Biológicas. Programa de Pós-Graduação em Genética e Biologia Molecular. IV. Instituto Agrônomo do Paraná. V. EMBRAPA. VI. Título.

CDU 631.52:633.34

SILVANA REGINA ROCKENBACH MARIN

**“IDENTIFICAÇÃO DE REGIÕES PROMOTORAS DE GENES  
EXPRESSOS EM SOJA SOB 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**

---

Orientador: Prof. Dr. Alexandre Lima  
Nepomuceno  
Empresa Brasileira de Pesquisa Agropecuária  
– Embrapa

---

Prof. Dr. Hugo Bruno Correa Molinari  
Empresa Brasileira de Pesquisa Agropecuária  
– Embrapa

---

Prof. Dr. Douglas Silva Domingues  
Instituto Agrônômico do Paraná – IAPAR

Londrina, 26 de Fevereiro de 2014.

### **Dedico e Ofereço**

Aos amados de minha vida: Maurício, Daniel,  
Lucas e Matheus... que tornam os meus dias  
mais felizes.

## **AGRADECIMENTOS**

Agradeço acima de tudo a Deus, autor e consumidor de tudo; por ter me criado e capacitado, e em sua soberania preenchido minha vida com pessoas e oportunidades maravilhosas... Minha alma anseia por ti!

À minha família amada pelo constante incentivo e carinho;

À Universidade Estadual de Londrina, pela oportunidade de aprendizado, e convivência com competentes docentes.

À Coordenadora do Programa de Pós Graduação em Genética e Biologia Molecular Dra. Ana Lúcia Dias, e à querida Secretária Sueli, pela atenção, dedicação, empenho e carinho que sempre demonstraram.

À Embrapa-Soja, local em que amo trabalhar, pela infra-estrutura laboratorial, suporte imprescindível para a realização deste trabalho.

A JICA/JST e ao laboratório da Dra. Kazuko Yamaguchi Shinozaki no instituto Jircas pela experiência única de conhecer e desenvolver parte deste trabalho em Tsukuba –Japão.

Ao meu admirável orientador Dr. Alexandre Nepomuceno, um exemplo de profissionalismo e competência, um cientista com visão além do alcance... pela orientação, ensinamentos, compreensão e incentivo em todos os momentos.

À co-orientadora e amiga Dra. Renata Fuganti Pagliarini por toda ajuda, paciência, conversas, conselhos, risadas e momentos de crescimento pessoal e profissional.

Aos pesquisadores da Embrapa Soja Dr. Ricardo Abdelnoor, Dra. Francismar Marcelino e Dra. Liliane Henning pela confiança e agradável convivência;

Aos amigos; César, André, Lizandra, Mayla, Juliane pela inestimável amizade, carinho e disposição em ajudar sempre.

À Dra. Fabiana Rodrigues e à Valéria Lopes Caitar pela preciosa colaboração para desenvolvimento desse trabalho.

A todos do Laboratório de Biotecnologia Vegetal: Verinha, Vieira, Márcia, Danielle, Renan, Talitta, Rafaela, Andriessa, Adriana, Patricia, Luana, Suellen, Aguida, Allan, Bruna, Katia, Thiago, Juliana, Michelle, Anelise, Cibelle, Alessandra, Luciano e todos os demais que passaram pelo Lab. São os bons momentos de convívio, aprendizado e amizade que fazem diferença na vida...

**MUITO OBRIGADA!**

*“Todas as coisas foram feitas por Ele, e sem Ele nada do que  
foi feito se fez”.*  
*João 1:3*

MARIN, Silvana Regina Rockenbach. **Identificação de regiões promotoras de genes expressos em soja sob déficit hídrico**. 2014. 137f. Dissertação (Mestrado em Genética e Biologia Molecular) – Universidade Estadual de Londrina, 2014.

## RESUMO

A regulação da transcrição gênica é complexa, envolvendo vários mecanismos, desde os elementos presentes no *core* promoter, elementos proximais e distais até as marcas epigenéticas e a regulação baseada em arranjos da cromatina. Os avanços nas tecnologias de transcriptomas têm permitido a análise em larga escala de transcritos e o estudo de suas regiões promotoras, e seus componentes de regulação para confirmação de sua funcionalidade como sequências reguladoras. Estudos de promotores responsáveis pela regulação gênica em condições específicas fornecem informações importantes sobre a regulação de genes e podem aumentar as opções de promotores para uso em engenharia genética vegetal. Neste contexto, o objetivo do presente estudo foi caracterizar as regiões promotoras de genes diferencialmente expressos em plantas de soja sob condições de déficit hídrico. A cultivar brasileira de soja BR 16 foi avaliada em duas condições de déficit hídrico: progressivo (oito dias) e de curto período (25 a 150 min). Os transcritos diferencialmente expressos identificados nos tratamentos de déficit hídrico foram comparados e 134 transcritos com mesma regulação foram encontrados em comum. A análise de 1kb a jusante da região promotora desses transcritos revelou a presença de 85 *cis* elementos comumente encontrados em regiões promotoras de plantas, sendo que 35 estavam presentes em 50% dos TDEs (Transcritos Diferencialmente Expressos). Entre esses *cis* elementos encontram-se TATA Box, CCAAT Box e GC Box e INR necessários para o acoplamento da maquinaria de transcrição e *cis* elementos especificamente induzidos por fatores abióticos tais como ABRE, DREB, HEAT e SALT. Todos os modelos de promotores apresentaram combinações binárias dos *cis* elementos GTBX e AHBP podendo haver uma relação entre esse modelo e a regulação transcricional em condição de déficit hídrico.

**Palavras-chaves:** Déficit hídrico. Promotores. Cis elementos. *Glycine max*

MARIN, Silvana Regina Rockenbach. **Identificação de regiões promotoras de genes expressos em soja sob déficit hídrico**. 2014. 137p. Dissertação (Mestrado em Genética e Biologia Molecular) – Universidade Estadual de Londrina, 2014.

## ABSTRACT

Transcription regulation is complex, involving several mechanisms, like elements in the core promoter, distal and proximal elements, to epigenetic traits and also the regulation based on chromatin rearrangements. Advances in transcriptome technology have allowed the analysis in large-scale of transcripts, the study of their promoter regions and their regulation components to confirm its functionality as regulatory sequences. Studies of promoters responsible for gene regulation under specific conditions provide important information on the regulation of genes and may increase the options of promoters to be used in plant genetic engineering. In this context, the objective of the present study was to characterize the promoter region of transcripts differentially expressed (TDEs) in soybean plants under water deficit conditions. Brazilian soybean cultivar BR 16 was evaluated under two water deficits: progressive (08 days) and short period (25 to 150 min). The transcripts differentially expressed identified in the treatments were compared and 134 transcripts with the same regulation were identified in common. The analysis of 1kb upstream of the promoter region in these transcripts revealed the presence of 85 cis elements generally found in the promoter region of plants, 35 were present in at least 50% of TDEs. Among these cis elements were *TATA Box*, *CCAAT Box*, *GC Box* and *INR* required for transcription machinery coupling, and cis elements specific induced by abiotic factors such as *ABRE*, *DREB*, *HEAT* and *SALT*. All promoter models presented binary combinations with *GTBX* and *AHBP* cis elements, which may have a relationship with the transcriptional regulation in water stress conditions.

**Keywords:** Soybean. Water requirements. Promoters (Genetics). Plant genetic engineering.

## LISTA DE FIGURAS

### CAPÍTULO 1

- Figura 1** – Via geral de resposta das plantas ao estresse. O sinal de estresse é percebido de modo extracelular pelos receptores de membrana, em seguida, é ativada uma grande e complexa cascata de sinalização intracelular, incluindo a síntese de moléculas de sinais secundários. Como resultado da cascata de sinalização ocorre a indução da expressão de múltiplos genes responsivos ao estresse, cujos produtos podem direta ou indiretamente levar a tolerância ao estresse. (Fonte: Huang et al., 2012) .....21
- Figura 2** – Motivos conservados do core promoter. Diagrama com sequência dos elementos que podem contribuir para a transcrição basal a partir de um core promoter, o qual pode conter todos, alguns ou nenhum desses elementos. TATA Box pode funcionar na ausência de elementos BRE, Inr, DPE. Em contraste, o motivo DPE requer a presença de um Inr. O BRE está localizado imediatamente a montante de TATA Box. (Fonte: Smale e Kadonaga, 2003).....31
- Figura 3** – Elementos regulatórios de um gene eucariótico. Os elementos cis são sítios de ligação presentes no DNA reconhecidos pelas proteínas regulatórias (elementos trans). (Fonte: Alberts et al., 2002).....35
- Figura 4** – O complexo transcricional consiste de RNA polimerase II, em associação com os reguladores gerais de transcrição, incluindo a proteína de ligação TATA Box. A maquinaria de transcrição é controlada por meio da ligação do FT aos cis elementos reguladores específicos no promotor. O elemento de regulação de ABA (ABRE) é contatado por FTs bZIP diméricos. O elemento de acoplamento (CE) é alvo de FTs do tipo AP2, que aumentam a transcrição ABA-regulado. As proteínas HD-ZIP se ligam a sequências pseudopalindrômicas ricas em AT. RNA Pol II é regulada pela fosforilação de seu domínio carboxilo terminal (CTD). A desfosforilação de uma fosfatase de proteína específica-CTD (CTD ADPG), diminui a expressão dos genes induzíveis por ABA. (Fonte: Himmelbach et al., 2003).....38

## CAPÍTULO 2

- Figura 5** – Sistema utilizado para o cultivo das plantas de soja cultivar BR 16 e aplicação do déficit hídrico progressivo.....79
- Figura 6** – **A)** Sistema hidropônico utilizado para o cultivo das plantas de soja cultivar BR 16 e **B)** aplicação do déficit hídrico de curto período .....84
- Figura 7** – Aspecto observado nas plantas de soja cultivar BR 16 submetidas ao déficit hídrico progressivo e seus respectivos controles. Início do experimento (T0); oito dias condição controle T(8)C e oito dias sob déficit hídrico T(8)DH.....89
- Figura 8** – Umidade relativa do solo durante aplicação do déficit hídrico progressivo no tratamento e controle. Início do experimento (T0); oito dias condição controle T(8)C e oito dias sob déficit hídrico T(8)DH .....90
- Figura 9** – *Northern Blot do gene LEA-8 nas amostras das plantas de soja submetidas aos tratamentos de oito dias sob déficit hídrico progressivo T(8)DH e plantas controle T(8)C e no início do experimento T(0)C. A hibridização da sonda ocorreu somente no período de déficit hídrico progressivo.....91*
- Figura 10** – Classificação por categoria funcional dos transcritos diferencialmente expressos identificados durante o déficit hídrico progressivo em plantas de soja cultivar BR 16. A frequência de cada classe foi normalizada pelo total de genes organizados nas diferentes categorias (n = 8.945). Aproximadamente 10,6% dos transcritos não possui função atribuída.....91
- Figura 11** – Classificação por categoria funcional dos transcritos diferencialmente expressos identificados nos três períodos analisados do déficit hídrico de curto período em plantas de soja cultivar BR 16. A frequência de genes observados em cada categoria foi normalizada pelo total de genes identificados nos três períodos (n = 10.562).....93

<b>Figura 12</b> – Diagrama de Venn. Comparação entre os transcritos diferencialmente expressos identificados no déficit hídrico progressivo e de curto período. Observa-se o número de transcritos exclusivos e comuns entre os déficits hídricos progressivo de curto período. A quantidade de transcritos up e down regulados entre os transcritos comuns em cada comparação foi T25-50 min / Progressivo 264 up e 83 down regulados, T75-100 min / Progressivo 764 up e 205 down regulados e T125-150 / Progressivo 1042 up e 812 down regulados .....	96
<b>Figura 13</b> – Regulação dos transcritos diferencialmente expressos comuns ao déficit hídrico progressivo e nos diferentes tempos de déficit hídrico de curto período .....	97
<b>Figura 14</b> – Diagrama de Venn, comparação entre os transcritos diferencialmente expressos comuns em cada um dos tempos de déficit hídrico de curto período e o déficit hídrico progressivo.....	97
<b>Figura 15</b> – Classificação por categoria funcional dos transcritos diferencialmente expressos comuns ao déficit hídrico progressivo e de curto período. A frequência de TDEs observados foi normalizada pela somatória dos genes categorizados comuns e com mesma regulação da expressão (n = 134) .....	98
<b>Figura 16</b> – Análise de enriquecimento dos transcritos diferencialmente expressos comuns ao déficit hídrico progressivo e de curto período e que apresentaram mesma regulação da expressão .....	99
<b>Figura 17</b> – Diagrama hierárquico de categoria funcional dos transcritos diferencialmente expressos comuns ao déficit hídrico progressivo e de curto período e com mesma regulação da expressão. Os quadros coloridos indicam termos com significância ( $p \leq 0,05$ ), e os termos não significativos são mostrados em quadros brancos. O grau de saturação da cor dos quadros é positivamente correlacionado com o nível de enriquecimento do termo .....	100
<b>Figura 18</b> – Frequência de cis-elementos putativos encontrados nos 134 transcritos diferencialmente expressos identificados em comum a ambos os experimentos de déficit hídrico progressivo e de curto período .....	102

**Figura 19** – Novos putativos cis-elementos encontrados pelo módulo CoreSearch do software Genomatix nos transcritos diferencialmente expressos *down* regulados comuns ao déficit hídrico progressivo e de curto período .....104

## LISTA DE TABELAS

- Tabela 1** – Conjuntos de primers testados para amplificação do gene alvo LEA e para os genes normalizadores (GmB-actina e 18S). O conjunto de primers (alvo e normalizador) com melhor relação entre as eficiências foi utilizado para validação do déficit hídrico progressivo ..... 82
- Tabela 2** – Processos observados na classificação funcional dos transcritos diferencialmente expressos durante os déficits hídricos. Número de transcritos up e down regulados identificados nos três tempos de déficit hídrico de curto período e no déficit hídrico progressivo.....94
- Tabela 3** – Principais fatores de transcrição relacionados a estresses abióticos. Número de transcritos up e down regulados identificados nos três tempos de déficit hídrico de curto período e no déficit hídrico progressivo.....95
- Tabela 4** – Elementos encontrados nos modelos de promotores obtidos pelo módulo FramWorker do software Genomatix nos transcritos diferencialmente expressos comuns ao déficit hídrico progressivo e de curto período ..... 103

## MATERIAL SUPLEMENTAR

- Tabela 1** – Transcritos diferencialmente expressos comuns ao déficit hídrico progressivo e os três tempos do déficit hídrico de curto período .....127
- Tabela 2** – Identificação dos *cis* elementos putativos encontrados nos 134 transcritos *up* e *down* identificados em comum a ambos os experimentos de déficit hídrico progressivo e de curto período: sua abundância, respectivas sequências, identificação de família e matriz ..... 132

## SUMÁRIO

<b>CAPÍTULO 1</b> .....	15
<b>1 Introdução</b> .....	15
<b>2 Revisão Bibliográfica</b> .....	17
<b>2.1 A Soja</b> .....	17
<b>2.2 Mecanismos moleculares de resposta ao déficit hídrico</b> .....	20
<b>2.3 Genes envolvidos nas respostas das plantas a estresses abióticos</b> ....	22
<b>2.4 Estudos Transcricionais de Larga Escala</b> .....	25
<b>2.5 Engenharia Genética de Plantas</b> .....	26
<b>2.6 Promotores em eucariontes</b> .....	28
<b>2.6.1 Elementos cis-Regulatórios: Core Promoter</b> .....	30
<b>2.6.2 Elementos Regulatórios: trans e cis elementos</b> .....	34
<b>2.7 Elementos Regulatórios envolvidos em estresses abióticos</b> .....	36
<b>2.8 Promotores constitutivos e estresse induzidos</b> .....	45
<b>2.9 Uso de ferramentas computacionais para identificação de genes de interesse biotecnológico</b> .....	47
<b>3 Objetivos</b> .....	50
<b>3.1 Objetivo Geral</b> .....	50
<b>3.2 Objetivos Específicos</b> .....	50
<b>4 Referências</b> .....	51
<b>CAPÍTULO 2</b> .....	75
<b>5 Artigo</b> .....	75
<b>5.1 Resumo</b> .....	76
<b>5.2 Introdução</b> .....	77
<b>5.3 Material e Métodos</b> .....	79
<b>5.3.1 Experimento de déficit hídrico progressivo</b> .....	79
<b>5.3.1.1 Obtenção do material biológico</b> .....	79
<b>5.3.1.2 Extração e Purificação de RNA Total</b> .....	80
<b>5.3.1.3 Síntese de cDNA</b> .....	81
<b>5.3.1.4 Análise via PCR em tempo real</b> .....	81
<b>5.3.1.5 Análise da expressão gênica via Northern Blot</b> .....	82
<b>5.3.1.6 Microarranjos de DNA</b> .....	83

5.3.2	Experimento de déficit hídrico de curto período.....	84
5.3.2.1	Cultivo das plantas e tratamento experimental .....	84
5.3.2.2	Construção das bibliotecas e sequenciamento .....	85
5.3.2.3	Alinhamento dos reads contra os transcritos de soja.....	85
5.3.2.4	Identificação dos transcritos diferencialmente expressos (TDEs).....	86
5.3.3	Classificação funcional dos transcritos diferencialmente expressos .....	86
5.3.4	Seleção dos transcritos comuns .....	86
5.3.5	Análise de enriquecimento dos transcritos comuns .....	87
5.3.6	Identificação dos cis-elementos putativos nas regiões promotoras dos transcritos selecionados.....	87
5.4	Resultados .....	89
5.4.1	Validação do déficit hídrico progressivo aplicado nas plantas de soja .....	89
5.4.2	Identificação de transcritos diferencialmente expressos no déficit hídrico progressivo .....	91
5.4.3	Identificação de transcritos diferencialmente expressos no déficit hídrico de curto período .....	92
5.4.4	Transcritos identificados no déficit hídrico progressivo e de curto período .....	95
5.4.5	Categorização Funcional.....	98
5.4.6	Análise de cis elementos putativos.....	101
5.5	Discussão.....	105
5.6	Considerações finais.....	114
5.7	Referências .....	116
5.8	Material Suplementar .....	127

## 1. Introdução

Os avanços científicos e a disponibilização de tecnologias ao setor produtivo permitiram nas últimas três décadas, um crescimento significativo da cultura da soja, proporcionando a mesma um lugar de destaque na economia brasileira e mundial. Entretanto, diversos fatores bióticos e abióticos podem interferir diretamente na produtividade, provocando prejuízos significativos para os produtores e para a economia do país.

A reduzida disponibilidade de água é um dos fatores mais importantes que afetam a regulação do crescimento e desenvolvimento das plantas e uma das principais causas para a redução do rendimento nas colheitas, e que afetam a maioria das zonas agronomicamente cultivadas por todo o mundo.

Tal problema tende a se agravar, pois previsões climáticas para as próximas décadas apontam para um aumento da temperatura média global, com consequente alteração de regime de chuvas, de evaporação da água, e aumento na frequência e intensidade dos eventos de seca (Burke et al., 2006). Sendo assim, investimentos em melhoramento genético de plantas, para características específicas como tolerância à seca, é um componente importante no esforço global para se adaptar os sistemas de cultivo aos constantes cenários de mudanças climáticas.

Plantas geneticamente modificadas (PGMs) que tolerem períodos mais prolongados de déficit hídrico, com capacidade de no momento de uma nova precipitação, restaurar uma condição metabólica ideal para dar continuidade ao seu desenvolvimento, são o alvo para a agricultura moderna. Dessa forma, a terceira geração de PGMs, tem por objetivo principalmente, a transferência de genes caracterizados em plantas modelos para outras espécies de importância agrônômica, como a soja, conferindo-lhes ganho de funções como a melhor adaptabilidade a ambientes e condições específicas adversas.

O entendimento da modulação da atividade transcricional tem ampla aplicabilidade para a produção de PGMs, onde é necessário, além da sequência que codifica a proteína a ser expressa, uma sequência promotora que module a expressão do transgene inserido (Visarada et al., 2009). No processo de modulação da atividade transcricional, genes são ativados ou reprimidos para controlar/minimizar o dano causado pelo estresse.

(Mazzucotelli et al., 2008). Dessa maneira, as regiões promotoras dos genes de plantas, têm atraído cada vez mais atenção da pesquisa básica, devido ao importante papel que desempenham na modulação da expressão espacial e/ou temporal de genes por meio da interação com fatores de transcrição (FTs) desencadeando respostas adaptativas. Entre outros aspectos, o conhecimento dos elementos *cis*-regulatórios presentes nos promotores é essencial para o entendimento da regulação de um determinado gene (Wang et al, 2009).

De modo geral, para expressão dos genes de interesse em plantas transgênicas são utilizados promotores constitutivos, que promovem a expressão do gene ao longo de todo o ciclo. Entretanto, a expressão de genes de forma constitutiva direciona o metabolismo da planta para a síntese dos genes exógenos, prejudicando características relacionadas à produtividade agrônômica. Promotores induzidos apenas em condições de estresse, promotores tecido-específicos, ou ainda, aqueles induzidos somente em determinados momentos do desenvolvimento vegetal constituem uma ferramenta chave a fim de garantir que a expressão de um gene de interesse seja efetiva e regulada no momento desejado.

No contexto da geração de PGMs visando cultivares mais tolerantes especificamente durante situações de déficit hídrico, o uso de estratégias de engenharia genética que utilizem promotores ativados ou desativados em soja, permitirá a regulação precisa de genes que promovam a proteção das estruturas celulares somente durante a desidratação celular, repercutindo positivamente na manutenção de atividades metabólicas e fisiológicas importantes na composição de características agrônômicas finais. Mesmo pequenas melhoras na tolerância à seca ou no uso mais eficiente da água em plantas economicamente importantes, poderiam resultar em considerável impacto econômico na agricultura mundial.

Diante do exposto e, dentro do contexto científico atual que busca novas alternativas para amenizar os prejuízos decorrentes da seca, foi objetivo deste trabalho caracterizar as regiões promotoras de transcritos diferencialmente expressos em plantas de soja sob condições de déficit hídrico progressivo e de período curto.

## 2. Revisão Bibliográfica

### 2.1 A Soja

Componente essencial na fabricação de rações animais e um elemento importante na dieta de grande parte da população mundial, a soja (*Glycine max* (L.) Merrill), é uma leguminosa muito versátil, pois o seu grão dá origem a vários produtos e subprodutos com ampla utilização na agroindústria, na indústria química e de alimentos, e também nos últimos anos para a produção de bicomustível e nutracêuticos (Hartman et al., 2011; Dwevedi e Kayastha, 2011). Sua ampla utilização deve-se a composição do grão, fonte de óleo e proteína vegetal com alto valor nutritivo; além de, macronutrientes, minerais e outros metabólitos secundários como isoflavonas (Sakai e Kogiso, 2008) e fitoestrógenos (Ososki e Kenelly, 2003), benéficos para a saúde.

Originária da Ásia Oriental, a soja foi introduzida no Brasil em 1882; contudo, devido às condições edafoclimáticas somente a partir do final da década de 1960 tornou-se cultivada com objetivos comerciais (Miyasaka e Medina, 1981). Nos últimos anos, os investimentos em pesquisa, desenvolvimento e inovação revolucionaram a produção de soja no Brasil. O melhoramento genético permitiu sua adaptação às diferentes regiões edafoclimáticas, além de minimizar os efeitos dos diferentes estresses bióticos e abióticos que interferem na produtividade da cultura (Popp et al., 2003; EMBRAPA, 2010; Shurtleff e Aoyagi, 2009). Os avanços nessa área possibilitaram também o incremento da produtividade, atingindo os maiores índices mundiais (MAPA 2012).

Atualmente a sojicultura corresponde a 49% da área cultivada em grãos no país, abastecendo o mercado interno e enviando o excedente ao mercado externo e movimentando negociações anuais que ultrapassam US\$ 20 bilhões. Essa *commodity* é o principal produto agrícola na pauta de exportações brasileiras e o maior responsável pelo aumento da colheita nacional de grãos, posicionando nosso país como segundo maior produtor e exportador, com perspectivas de alcançar e se consolidar na primeira posição. Na safra 2012/13, em razão da quebra na safra americana em decorrência da forte estiagem, o Brasil superou os EUA e ocupou a primeira colocação com uma produção de 81 milhões de toneladas. Anualmente cerca de 30,7 milhões de

toneladas de soja são processadas pela indústria nacional. A previsão da taxa de crescimento anual de produção de soja é de 2,3% até 2022, quando a produção nacional deve representar 40% do comércio mundial do grão e 73% do óleo de soja (MAPA 2012).

Entretanto, apesar de números tão positivos e expressivos, a produtividade e a distribuição espacial de importantes culturas como a soja, podem ser severamente limitadas pelas restrições ambientais, impostas por fatores bióticos e abióticos aos quais as plantas estão constantemente expostas. A seca, a salinidade e as temperaturas extremas são as causas primárias de quebras de produtividade de safra na agricultura mundial, sendo o déficit hídrico o principal responsável por reduzir os rendimentos médios da maioria das culturas (Salekdeh et al., 2009) levando a prejuízos financeiros significativos.

Em decorrência da estiagem prolongada ocorrida no Brasil na safra 2011/2012 a redução na produção de soja chegou a 43,8% no Rio Grande do Sul (RS), estado mais afetado. Também em outros países produtores, como os Estados Unidos, a severa estiagem ocorrida neste mesmo ano afetou 70% da área plantada com perdas de aproximadamente 14% na produção de soja, causando um grande impacto na economia mundial (USDA, 2012).

De acordo com Salekdeh e colaboradores (2009), dois terços da produção agrícola mundial ocorre sob estresse hídrico. Prevê-se ainda, que a ocorrência de secas em muitas regiões produtoras de alimentos, irá aumentar significativamente em resposta às mudanças climáticas (Collins et al, 2008; Reynolds e Ortiz, 2010). A duração, a severidade e a fase de desenvolvimento da planta em que o estresse é imposto, influenciam no padrão de resposta e na dimensão das perdas de produtividade que ocorrem (Pinheiro & Chaves, 2011, Bertolli et al., 2012; Saint Pierre et al., 2012).

Dessa forma, apesar do aumento expressivo da produção de alimentos no último meio século, um dos maiores desafios para as próximas décadas será a segurança alimentar. Considerando-se a previsão de que até 2050 a população mundial deverá atingir nove bilhões de pessoas, será necessário um acréscimo de 70% na produção de alimentos (Godfray et al., 2010). Ao cenário de crescimento demográfico e à alta demanda por alimentos, adicionam-se os efeitos adversos para a agricultura decorrentes das mudanças climáticas

globais; impondo restrições às áreas aptas para o cultivo devido aos estresses bióticos e abióticos que serão mais severos, interferindo diretamente na produção.

Portanto, novas alternativas/soluções para minimizar/mitigar os problemas ocasionados por eventos de seca são necessárias, porém esta não é uma missão simples. Várias estratégias que vão desde o manejo adequado da cultura visando maior eficiência do uso da água disponível, até o uso de estratégias de engenharia genética que permitam aumentar a eficiência do uso de água pelas plantas (Cominelli et al., 2013) podem e devem ser implementadas.

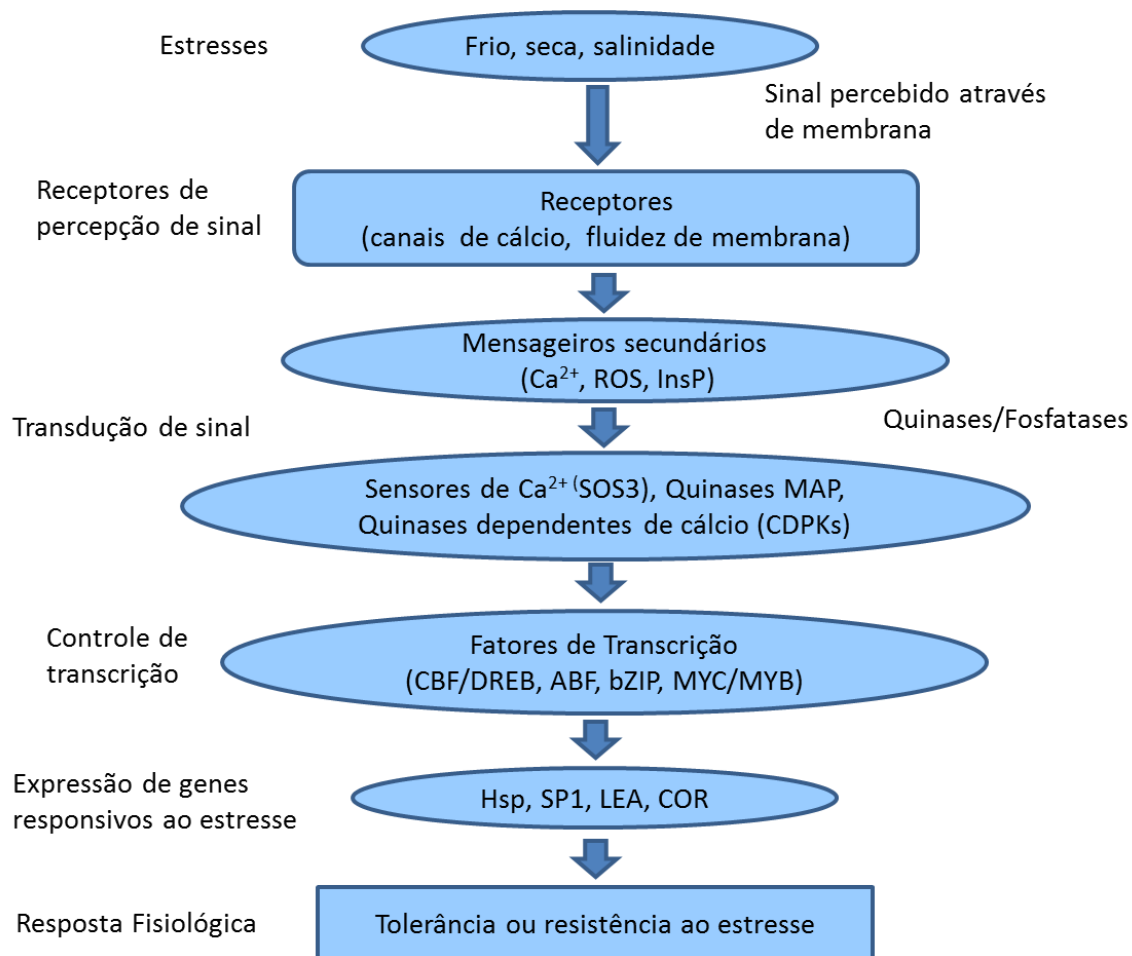
Neste contexto, os avanços científicos na área biotecnológica vêm ampliando as oportunidades na agricultura. Técnicas de melhoramento molecular e engenharia genética permitem, atualmente, a manipulação de genes de interesse possibilitando o desenvolvimento de plantas geneticamente modificadas (PGMs) com características mais adaptadas às diferentes condições edafoclimáticas (Ainsworth et al., 2012). Ainda nesse panorama, o sequenciamento do genoma da soja, ampliou as informações sobre esse organismo, tornando-se uma ferramenta importante para estudos que visam à compreensão da complexa rede de interações gênicas.

O genoma da soja consiste em 1,15 bilhões de bases, com 46.430 genes preditos como codificadores de proteína, sendo que 12,2% foram identificados como fatores de transcrição (FTs) (Schmutz et al., 2010). Está organizado em 20 cromossomos, de acordo com os 20 grupos de ligação da soja, com a presença de sequências altamente repetitivas, heterocromáticas com baixa recombinação (57% do genoma), e sequências em regiões eucromáticas, pouco repetitivas e com alta taxa de recombinação (43% do genoma). Algumas famílias de FTs são mais esparsas e outras mais abundantes no genoma da soja em relação ao genoma de *Arabidopsis* indicando que pode haver diferenças nas vias regulatórias entre essas espécies (Schmutz et al., 2010). A anotação destas sequências identificou 5.671 genes de soja como putativos fatores de transcrição.

## **2.2 Mecanismos moleculares de resposta ao déficit hídrico**

Os estados de deficiência hídrica na planta podem ser causados por seca, salinidade e temperaturas extremas, e geralmente estão inter-relacionados, promovendo danos celulares semelhantes, como por exemplo, estresse oxidativo e perturbações do equilíbrio osmótico e iônico das células (Reynolds e Ortiz, 2010). Para sobreviver a estas adversidades, os vegetais desenvolveram ao longo da evolução mecanismos de adaptação aos estresses a níveis celulares, moleculares, fisiológicos e bioquímicos (Shinozaki; Yamaguchi-Shinozaki, 2007, Huang et al., 2012). Estes mecanismos são dinamicamente regulados em resposta ao tipo, severidade e duração de um ou mais estresses combinados (Atkinson e Urwin, 2012). Apesar das respostas moleculares a esses estresses apresentarem uma grande complexidade, os mecanismos de defesa envolvidos no processo de tolerância à seca parecem ser muito similares em todo o reino vegetal, apresentando etapas comuns nas rotas de sinalização moleculares (Kasuga et al., 2004; Reynolds e Ortiz, 2010, Huang et al., 2012).

Alguns avanços na compreensão desses mecanismos revelaram que as respostas às perturbações da homeostase celular são geralmente acompanhadas de grandes mudanças no transcriptoma, proteoma e metaboloma (Chae et al., 2009, Shulaev et al., 2008). A expressão de uma variedade de genes e seus produtos induzida por essas alterações ambientais não somente agem diretamente na tolerância ao estresse abiótico, como também regulam os mecanismos de percepção do estresse e ativam a transdução de sinais que ao final ativam a expressão dos genes diretamente envolvidos na atividade biológica de proteger as estruturas celulares durante a desidratação celular (Figura 1) (Xiong et al., 2002; Shinozaki et al., 2003; Bartels Sunkar, 2005, Shinozaki e Yamaguchi-Shinozaki, 2007; Huang et al., 2012).



**Figura 1.** Via geral de resposta das plantas ao estresse. O sinal de estresse é percebido de modo extracelular pelos receptores de membrana, em seguida, é ativada uma grande e complexa cascata de sinalização intracelular, incluindo a síntese de moléculas de sinais secundários. Como resultado da cascata de sinalização ocorre a indução da expressão de múltiplos genes responsivos ao estresse, cujos produtos podem direta ou indiretamente levar a tolerância ao estresse. (Fonte: Huang et al., 2012).

Assim, os sinais do estresse desencadeiam processos de sinalização e de controle da transcrição gênica que, por sua vez, ativam os mecanismos de resposta ao estresse para restabelecer a homeostase e proteger as estruturas celulares danificadas. Dessa forma, a resposta inadequada em um ou mais passos da sinalização e ativação gênica pode resultar em mudanças irreversíveis da homeostase celular e na destruição de proteínas funcionais, estruturais e de membranas, levando à morte das células (Wang et al., 2003; Huang et al., 2012).

Sendo assim, no contexto molecular, a percepção do estresse pelas células vegetais representa uma etapa crucial, pois resulta na ativação de uma cascata de eventos levando a respostas fisiológicas, metabólicas e de

desenvolvimento em vários níveis. As vias de transdução de sinal podem em última instância regular a expressão de genes responsivos à seca e a outros estresses através de diversos fatores de transcrição (Balderas-Hernandez et al., 2013), que se ligam a *cis* elementos específicos na região promotora dos genes alvo, ativando ou reprimindo sua expressão.

As vias de transdução de sinais para resposta ao déficit hídrico, salinidade e temperaturas extremas podem ser divididas em três grandes tipos de sinalização: 1) a sinalização aos estresses osmóticos e oxidativos que usam os módulos MAPK (*Mitogen Activated Protein Kinase*) e envolvem a geração de enzimas de busca a espécies reativas de oxigênio (ROS – *Reactive Oxygen Species*) além de compostos anti-oxidantes e osmólitos; 2) a sinalização dependente de  $Ca^{2+}$  que leva a ativação dos genes LEA (*Late Embryogenesis Abundant* – Embriogenese tardia abundante) e também envolve a síntese de diversas proteínas responsivas ao estresse; e 3) a sinalização que envolve a via SOS (*Salt Overlay Sensitive*), específica para estresse iônico e que age na regulação da homeostase celular (Rodriguez et al., 2005).

### **2.3 Genes envolvidos nas respostas das plantas a estresses abióticos**

A complexa resposta das plantas a estresses abióticos envolve a ativação de muitos genes que desencadeiam mecanismos fisiológicos, bioquímicos e moleculares de resposta; logo, a identificação dos genes chaves nessa via de sinalização é importante para entender as bases moleculares da tolerância ao estresse (Venu et al., 2013). Alguns genes respondem rapidamente ao déficit hídrico, enquanto outros são induzidos ou reprimidos lentamente. Geralmente os genes de resposta rápida são os fatores de transcrição que, quando expressos, ativam os genes de resposta mais tardia, que corresponde à maioria dos genes responsivos ao estresse atuando fisicamente na defesa/manutenção das estruturas celulares durante a desidratação (Cheong et al., 2002; Zhu, 2002).

Alguns genes estrategicamente localizados na complexa rede molecular, e que desempenham papéis centrais de regulação nas respostas das plantas a estresses bióticos e abióticos têm sido relatados como reguladores mestres, ou seja, não devem estar sob a influência regulatória de qualquer outro gene.

Esses reguladores controlam positiva ou negativamente a transcrição de uma variedade de genes envolvidos no mecanismo de adaptação e sobrevivência da planta, associando-se diretamente na região promotora dos genes-alvos ou indiretamente pela ativação de fatores de transcrição ou repressores gerais. Os principais reguladores descritos para estresses bióticos e abióticos são NPR (*Non-expressor of Pathogenesis Related Protein*), SnRK1 (*SNF1-Related Kinase 1*) e TOR (*Target of Rapamycin*), que agem como controladores para garantir que os recursos celulares sejam otimizados (Balderas-Hernandez et al., 2013).

O gene NPR1 de *Arabidopsis* foi introduzido em arroz e atuou como um regulador negativo da transcrição de vários genes de arroz como *rab21* (uma desidrina de arroz), *salT* (codifica a proteína induzida por estresse salino) e *dip1* (codifica proteína induzida por estresse de desidratação). Plantas transgênicas de arroz expressando *AtNPR1* foram hipersensíveis ao estresses salino e de seca (Quilis et al., 2008). *AtNPR1* é provavelmente um componente chave no aumento da tolerância ao calor e ao estresse salino mediada por brassinosteróides (grupo de fitormônio esteróides) (Divi et al., 2010).

As SnRK são classificadas em três subgrupos, SnRK1, SnRK2 e SnRK3 baseados na similaridade de sequências e estrutura dos domínios. Os membros da família SnRK1 controlam a sinalização mediada por hormônios como o ABA, a auxina e a citocinina (Hubbard et al., 2010). SnRK2 são reguladores positivos (dependente ou independente de ABA) na resposta a estresses como déficit hídrico, salinidade, baixa temperatura, cádmio e estresse oxidativo (Kulik et al., 2011, Kulik et al., 2012).

De maneira geral, entre os genes de resposta ao déficit hídrico encontram-se os genes envolvidos na percepção do sinal como as fosfolipases, e possíveis mensageiros secundários, como as quinases dependentes de  $Ca^{+2}$ , CDPKs; e as ROS. Também, os genes relacionados com os mecanismos de resposta ao sinal, com a função putativa de proteção osmótica como a prolina, glicina betaína, trealose, e açúcares, além das chaperonas moleculares (Hsp, SP1, LEA, COR) e os genes relacionados ao transporte iônico e de água como as aquaporinas. Há ainda, os genes de controle transcricional, como os fatores de transcrição que ativam uma cascata de genes que agem em conjunto aumentando a tolerância frente à múltiplos estresses. Neste grupo destacam-

se as famílias de fatores de transcrição AP2/ERF (*APETALA2/ Ethylene Response Factor*), bZIP (*Basic leucine Zipper domain*), NAC (*NAM, ATAF1, 2, e CUC2*), MYB (*Myelo Blastosis oncogene*), MYC (*Myelo Cytomatosis oncogene*), Cis2His2 ZF (*Cistein2 Histidin 2Zinc-Finger*) e WRKY (*amino acids W, R, K, Y*) (Bhatnagar-Mathur et al., 2008, Huang et al., 2012).

Alguns desses genes foram identificados utilizando análises em larga escala de transcriptoma como significativamente *down* regulados em *Arabidopsis* especificamente em resposta à desidratação, alta salinidade e tratamento com ABA. Entre os genes identificados estão os pertencentes a classe das proteínas LEA como *RD29B*, *AIL1*, *RAB18*, *EM1* e *EM6*; genes regulados por ABA como *RD20*, *LTP4* e *SAG29*; genes de fatores de transcrição MYB como *AtMYB74*, *AtMYB79*, *AtMYB102* e *AtMYB121*; e genes da fosfatase grupo-A tipo 2 (*Group-A type-2C Protein Phosphatase -PP2C*) como *AHG1*, *AHG3*, *HAI1*, *HAI2* e *HAI3* (Yoshida et al., 2010).

Ainda, estudos com PGMs têm sido realizados para validar as respostas de genes relatados como responsivos a estresses abióticos. Amara e colaboradores (2013) trabalharam com milho GM superexpressando o gene LEA *Rab28* e constataram que sob estresse osmótico, as mudas transgênicas mostraram aumento de área foliar e de raiz, teor relativo de água superior, reduzida perda de clorofila e menor produção de malondialdeído quando comparadas as plantas não GMs, evidenciando o potencial do gene LEA *Rab28* para melhorar a tolerância ao déficit hídrico. Além disso, em condições de falta de água, as sementes transgênicas exibiram taxas de germinação mais altas do que as sementes do tipo selvagem.

O gene *P5CS* que codifica para uma enzima chave na síntese de prolina foi inserido via engenharia genética em petúnia (Yamada et al., 2005) e tabaco (Gubis et al., 2007). Em outro estudo, o gene *P5CR* que codifica para uma enzima chave na degradação de prolina foi inserido em soja (Ronde et al., 2004) Todas as linhas transgênicas destas espécies vegetais apresentaram maior acúmulo de prolina e tolerância à seca. Plantas transgênicas de tabaco super expressando a enzima CMO, envolvida na síntese de glicina betaína, foram obtidas e apresentaram alto acúmulo deste osmólito sob condições de deficiência hídrica e aumento da tolerância à seca (Shen et al., 2002; Zhang et al., 2008).

## 2.4 Estudos Transcricionais de Larga Escala

O estudo das respostas vegetais aos estresses abióticos, em todos os níveis do transcriptoma, ou seja, o conjunto de transcritos expressos em uma célula em um determinado estágio de desenvolvimento ou condição fisiológica revela a natureza complexa das respostas multigênicas em plantas e tem demonstrado ser uma importante ferramenta no entendimento dos processos biológicos, contribuindo para caracterização da resistência ou tolerância de plantas a uma condição específica (Payton et al., 2009).

Nas duas últimas décadas ocorreu um avanço na compreensão de aspectos estruturais e funcionais de genomas de plantas. Estes progressos foram facilitados pela adoção de plantas modelo para a maioria das famílias vegetais economicamente importantes, bem como, o desenvolvimento de recursos da genômica, transcriptômica e metaboloma. Os projetos de sequenciamento de grandes culturas têm melhorado ainda mais a compreensão da estrutura genômica, da evolução e das funções dos genes (Augustine et al., 2013; Das e Parida, 2013).

Os avanços obtidos nos perfis de transcriptoma foram alcançados pelo sequenciamento em larga escala de transcritos, pelo emprego de tecnologia de Microarranjo de DNA e ainda por RNA-Seq. O emprego de técnicas como o sequenciamento de bibliotecas de transcritos completos (*full length*), de etiquetas de sequências expressas (ESTs - *Expressed Sequence Tags*), Super Sage, e de bibliotecas subtrativas têm proporcionado o acesso cada vez mais amplo ao genoma das espécies, possibilitando a análise de milhares de genes simultaneamente, e permitindo a identificação de genes e rotas metabólicas relacionadas aos mais variados tipos de estresses, fases de desenvolvimento e condições de tratamento (Matsumura et al., 2010).

Entretanto, cada estratégia de construção de bibliotecas possui limitações específicas. Enquanto as bibliotecas subtrativas podem privilegiar genes diferencialmente expressos de alta expressão, em detrimento a genes de baixo nível de expressão como fatores de transcrição, as bibliotecas para produção de ESTs produzem abundante informação, mas muitas vezes com alta redundância.

A capacidade de medir simultaneamente a expressão de milhares de genes é um sistema analítico poderoso, e a disponibilidade de novas tecnologias para esse fim tem gerado muitas novas estratégias de estudo da resposta gênica e, mais importante ainda, permitido a utilização deste conhecimento para o melhoramento de culturas economicamente importantes (Augustine et al., 2013), levando a um novo estado da arte na biologia molecular, a genômica translacional. Este termo é a combinação das palavras translacional que se refere à expressão de um conceito em diferentes linguagens, e genômica a qual trata do entendimento da natureza e função do genoma, ou seja, da aplicação prática da linguagem genômica (Augustine et al., 2013).

Vários trabalhos, utilizando a técnica de Microarranjo de DNA, NGS (*Next Generation Sequencing* – Sequenciamento de nova geração) ou a associação de microarranjo e técnicas de NGS têm sido realizados visando avaliar o nível de expressão diferencial de plantas sob estresses abióticos (Yoshida et al., 2010; Lorenz et al., 2011; Liu et al., 2013a). Ainda, especificamente em resposta à seca, análises da expressão de múltiplos genes têm contribuído significativamente para o entendimento dos tipos e da quantidade de genes que são regulados pela deficiência hídrica, além da identificação dos *cis* elementos regulatórios presentes em seus promotores (Mochida et al., 2009, Yamamoto et al., 2011; Maruyama et al., 2012; Rodrigues et al., 2012).

## **2.5 Engenharia Genética de Plantas**

A engenharia genética de plantas tem possibilitado estratégias para o melhoramento de culturas para uso generalizado na agricultura, e também usos específicos potenciais para fins industriais e farmacêuticos (Potenza et al., 2004). O desenvolvimento de plantas transgênicas já é relatado em muitas espécies vegetais economicamente importantes, no entanto, a comercialização desta tecnologia ainda é limitada a algumas culturas como o algodão, milho, soja e canola (Visarada et al., 2009). Mais de 170 milhões de hectares de plantas geneticamente modificadas foram cultivadas no ano de 2012 (James, 2012). No entanto, o aumento no potencial de utilização desta tecnologia

implica no desenvolvimento de métodos de controle da expressão do transgene. Dessa forma, a disponibilidade de promotores que diferem quanto à sua capacidade para regular os padrões de expressão espacial e temporal do transgene pode aumentar a aplicação e o sucesso da tecnologia de transgenia (Potenza et al., 2004).

A transgenia consiste na inserção de um ou mais genes capaz de conferir ao organismo receptor uma característica desejável (Visarada et al., 2009). O termo transgene refere-se às sequências de nucleótidos basicamente constituídas por uma região promotora, uma sequência codificadora do gene de interesse, uma sequência *enhancer*, e um terminador da codificação. O promotor deve ser escolhido com base na expressão espacial e temporal desejada para o transgene (Visarada et al., 2009) e pode ser oriundo de organismos semelhantes ou filogeneticamente distantes do hospedeiro (Singh et al., 2006). As plantas resultantes desta metodologia são chamadas de geneticamente modificadas (GM) (Key et al., 2008).

A regulação da expressão do transgene na planta é feita, em maior parte, pelo promotor e suas sequências regulatórias, uma vez que a transcrição é o primeiro processo de regulação gênica. Os promotores utilizados para regular a expressão de transgenes em plantas são classificados como constitutivos, órgão/tecido/célula específicos, induzíveis e promotores sintéticos (Hernandez-Garcia e Finer, 2013). A escolha do promotor a ser utilizado depende do objetivo final da transformação genética (Potenza et al., 2004).

Apesar da utilização de culturas GMs gerar potenciais benefícios para a agricultura, um fator importante a ser considerado são as questões relacionadas à segurança ecológica e toxicológica dessas culturas (Singh et al., 2006), e neste contexto, a utilização de promotores que limitem a expressão dos transgenes a determinado órgão ou período pode diminuir estas preocupações (Potenza et al., 2004).

Outro aspecto importante refere-se à variação na expressão do transgene entre as plantas geradas sob as mesmas condições, que ocorre em razão de combinação de fatores como número de cópias inseridas, local de inserção no genoma, a ocorrência de silenciamento e o emprego de sequências regulatórias. A escolha de um promotor adequado para regular a expressão do transgene também pode contribuir para diminuir essa

variabilidade de expressão e aumentar a eficiência da técnica (Butaye et al., 2005; Hernandez-Garcia e Finer, 2013).

Ainda, considerando-se que a expressão de genes de defesa de forma constitutiva também pode comprometer o metabolismo da planta e prejudicar características relacionadas ao desenvolvimento e produtividade, nos últimos anos muitos trabalhos têm sido desenvolvidos no sentido de isolar e avaliar promotores induzidos específicos para uma variável ou um tecido (Xu et al., 2009; Tittarelli et al., 2009; Kato et al., 2010; Yi et al., 2010; Wang et al., 2013; Noh et al., 2012).

## **2.6 Promotores em eucariontes**

O promotor desempenha um papel central na regulação da expressão gênica, pois contém os sítios de ligação para os complexos protéicos que em sincronia com as RNA polimerases atuam na transcrição dos genes (Griffiths et al., 2000). Em plantas além das RNA polimerase I, II e III existem ainda as RNA polimerases IV e V que participam da síntese de pequenos RNAs envolvidos no silenciamento gênico (Haag e Pikaard, 2011). O processo de iniciação transcricional em eucarioto é complexo e a visão de como o aparato de transcrição é montado no promotor tem mudado na última década (Zhou et al., 2009).

De acordo com o modo de iniciação da transcrição, os promotores podem ser classificados em dois tipos: de iniciação focada, quando a transcrição inicia-se sempre em um único nucleotídeo ou em um grupo deles; e de iniciação dispersa, quando a transcrição ocorre a partir de vários sítios de iniciação fracos dentro de uma região de 50 a 100 nucleotídeos. Em geral os genes regulados espacialmente e temporalmente apresentam promotores com iniciação focada e os genes constitutivos apresentam iniciação dispersa, entretanto, os promotores também podem apresentar os dois modos de iniciação, contendo múltiplos sítios de iniciação fracos ou um único sítio forte (Juven-Gershon e Kadonaga, 2010).

Quanto a forma de transcrição, organismos eucariotos em sua maioria apresentam transcritos monocistrônicos, onde a expressão dos genes é regulada individualmente, por uma região promotora a montante e um sítio de

terminação da transcrição a jusante. No entanto, alguns genes são transcritos de forma policistrônica, como descrito por Field e colaboradores (2011) que identificaram em *Arabidopsis thaliana* dois grupos de genes similares a operons, necessários para a síntese de dois tipos diferentes de triterpenos.

De modo geral, os promotores eucarióticos são capazes de iniciar a transcrição em duas direções, entretanto, na maioria dos casos o alongamento ocorre no sentido 5'. Portanto, devem existir mecanismos para regular a orientação, como sinais de sequências específicas presentes no promotor ou na região de codificação. Sugere-se ainda que a composição de nucleótidos em torno do promotor afete a sua bidirecionalidade, ou que exista um mecanismo associado a modificações na cromatina, e dessa forma rodadas anteriores de transcrição marcariam a orientação favorecida nas rodadas subsequentes, além de um mecanismo espacial, onde a estrutura 3D da transcrição poderia favorecer a formação de um *looping* e a ligação ao promotor (Wei et al., 2011).

Os promotores bidirecionais são definidos como regiões reguladoras compartilhadas que influenciam a expressão de dois genes opostamente orientados. Pares de genes bidirecionais com funções similares parecem ter forte correlação na expressão (Wang et al., 2009). Um grande número de regiões não codificantes (ncRNA) próximo a promotores de genes que codificam proteínas tem sido relatado, e sugere-se ainda que pode haver o envolvimento desses ncRNAs em circuitos de regulação (Wei et al., 2011).

Em geral, a posição do sítio inicial de transcrição TSS (*Transcription Site Start*) é um dos principais mecanismos para gerar variação genética em eucariotos. Análise de entropia relativa na região de TSS em arroz e *Arabidopsis* revelou que ambos possuíam em média dois ou mais TSSs por loco, e os sinais de nucleótidos próximos aos TSSs foram semelhantes nas duas espécies, entretanto, foram diferentes em relação a posição. Os TSSs localizados mais a montante apresentavam elementos *cis* canônicos, enquanto TSSs mais a jusante apresentaram características de nucleotídeos atípicos. Padrões de expressão se distinguiram entre TSSs a montante e a jusante, indicando que TSSs de plantas foram geralmente diversificados nas regiões a jusante, o que resultou no desenvolvimento de novos padrões de expressão de genes (Tanaka et al., 2009).

Apesar da sua importância, as características básicas de um promotor de plantas ainda não são bem compreendidas. Liu e colaboradores (2013b) analisaram sequências de promotores vegetais para determinar a diversidade de sequências e as características estruturais gerais nas mesmas, e observaram um teor médio de GC menor em dicotiledôneas do que em monocotiledôneas, o que pode sugerir diferentes pressões evolutivas para as sequências promotoras. Aproximadamente 3,3% dos promotores de plantas abrigaram sequências minissatélites, e 15,4% possuía sequências microsatélites. Muito poucos elementos transponíveis foram detectados dentro dos promotores vegetais. O motivo de fator de transcrição mais comum foi AGAGAGAGA, seguido por TTAGGGTTT e GCCGCC. Transcritos com as regiões promotoras de genes que contêm esses motivos estão comumente envolvidos em processos metabólicos, regulação biológica, estímulo e resposta em plantas.

### **2.6.1 Elementos *cis*-Regulatórios: *Core Promoter***

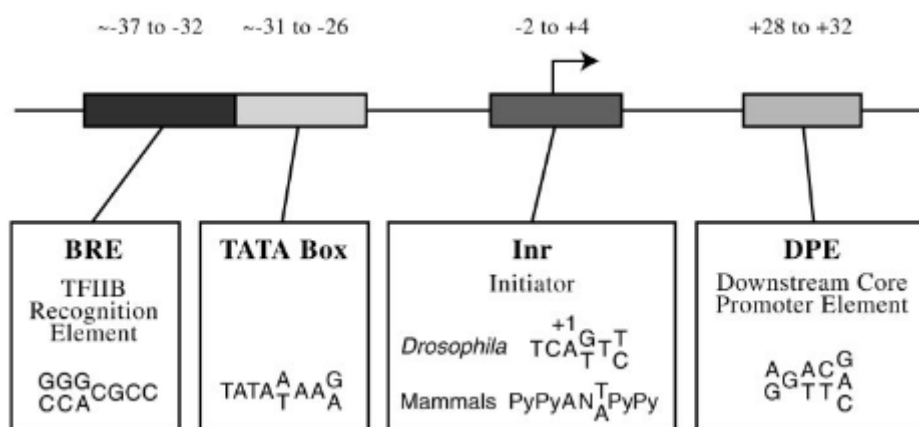
Promotores eucariotos apresentam diferentes regiões com sequências e funções específicas, denominadas elementos *cis*-regulatórios aos quais se ligam as proteínas *trans*-regulatórias. Essas regiões são divididas em três classes de acordo com a sua localização (Maston et al., 2006).

- 1) *Core Promoter*: região de ligação da RNA Polimerase II e de fatores de transcrição próximos ao sítio de iniciação da transcrição (TSS);
- 2) Elementos *cis* proximais: correspondem a sequências próximas ao promotor que auxiliam a ligação da RNA polimerase II e modulam sua atividade e;
- 3) Elementos independentes de distância: os denominados estimuladores, silenciadores e isoladores atuam a distâncias grandes do TSS e influenciam a taxa de transcrição.

O *core promoter* corresponde à região de DNA mínima suficiente para direcionar precisamente o início da transcrição através da maquinaria da RNA polimerase II. Esta região estende-se por aproximadamente 35 pb a montante e/ou a jusante do TSS (Schramm e Hernandez, 2002; Gross e Oelgeschläger, 2006, Juven-Gershon e Kadonaga, 2010).

Elementos gênicos básicos presentes no *core-promoter* são muito variáveis em sua estrutura, e cada sequência é encontrada apenas em um subconjunto de genes, sendo várias combinações possíveis para formar diversos tipos de *core promoter* (Juven-Gershon e Kadonaga, 2010). Apesar da sua complexidade, existem alguns elementos básicos conservados como o TATA Box, o elemento BRE (TFIIB *Recognition Element*) de reconhecimento ao TFIIB, o iniciador (Inr) e o elemento do promotor a jusante DPE (*Downstream Promoter Element*) (Smale e Kadonaga, 2003; Gross e Oelgeschläger, 2006; Baumann et al., 2010). A localização e as sequências consenso dos principais elementos do *core promoter* estão ilustradas na Figura 2.

Além da função de direcionar o acoplamento correto da maquinaria de transcrição da RNA polimerase II, o *core promoter* também atua como uma região de elementos regulatórios *cis* atuantes que determina a especificidade para um complexo transcricional distinto e para outros elementos regulatórios mais distais, como os acentuadores (*enhancers*). Um alto grau de controle transcricional é determinado em função da especificidade do *core promoter* a fatores de transcrição e acentuadores distintos (Butler e Kadonaga, 2002).



**Figura 2.** Motivos conservados do *core promoter*. Diagrama com sequência dos elementos que podem contribuir para a transcrição basal a partir de um *core promoter*, o qual pode conter todos, alguns ou nenhum desses elementos. TATA Box pode funcionar na ausência de elementos BRE, Inr, DPE. Em contraste, o motivo DPE requer a presença de um Inr. O BRE está localizado imediatamente a montante de TATA Box. (Fonte: Smale e Kadonaga, 2003).

O primeiro elemento identificado em eucariontes foi o TATA Box (Goldberg, 1979; Breathnach e Chambon, 1981), localizado tipicamente entre 25-30 nucleotídeos a montante do ponto de início de transcrição. É o sítio de interação da TBP (*TATA Binding Protein*) e determina o local de início da transcrição. Apresenta a sequência consenso TATAAA, podendo ocorrer variações de um ou dois nucleotídeos, e na posição relativa ao sítio de início de transcrição (Butler e Kadonaga, 2002).

Promotores sem TATA Box foram inicialmente relacionados a genes constitutivos em plantas, entretanto, estudos constataram que essa classe de promotores está também associada a genes de fotossíntese (Nakamura et al., 2002). Em mamíferos, de 10 a 15% dos promotores possuem TATA Box (Juven-Gershon e Kadonaga, 2010) e em plantas a prevalência é de 20 a 30% (Yamamoto et al., 2011).

O TATA Box e suas variantes são elementos reguladores envolvidos na formação do complexo de iniciação da transcrição, entretanto, a baixa frequência com que esse elemento é encontrado na maioria dos organismos modelos estudado sugere que diferentes elementos do *core promoter* também estejam vinculados no recrutamento de TBP. Foram identificados em promotores sem TATA Box, elementos TC localizados 30 bases a montante do TSS, mesma localização de TATA Box. As características funcionais e estruturais dos genes que contém elementos TC são distintas dos que contém TATA Box. A análise do transcriptoma de *Arabidopsis* revelou que genes contendo TATA Box geralmente apresentam níveis relativamente elevados de expressão, enquanto os que contêm o elemento TC comumente são expressos em condições específicas, sugerindo que estes elementos podem integrar uma nova classe de elementos reguladores da modulação da expressão gênica em plantas (Bernard et al., 2010).

O elemento Inr, região importante para o processo transcricional, geralmente engloba o sítio de início de transcrição, e é encontrado tanto em promotores que contém TATA Box como nos que não o contém. Esse elemento é caracterizado pela presença de uma adenosina no sítio de início da transcrição (posição +1), uma citosina na posição -1 e algumas pirimidinas ao redor desses nucleotídeos (Smale e Kadonaga, 2003). Funcionalmente, o Inr é semelhante ao TATA Box, pois também é reconhecido pelo TFIID podendo

compensar a ausência desse elemento em promotores sem TATA Box. Entretanto, eles funcionam sinergisticamente em promotores que possuem esses dois elementos separados por 25 pb (Smale e Kadonaga, 2003).

O elemento DPE é um sítio de ligação para o fator de transcrição basal TFIID, que pode se ligar cooperativamente aos elementos Inr e DPE. O DPE se encontra na posição +28 a +32 em relação à posição A+1 do elemento Inr e está comumente presente em promotores sem TATA Box. A ocorrência de mutações em Inr ou DPE causa a perda de afinidade pelo TFIID, e a alteração de um único nucleotídeo de espaçamento entre Inr e DPE pode resultar na diminuição significativa da transcrição (Burke e Kadonaga, 1997; Kutach e Kadonaga, 2000; Juven-Gershon e Kadanoga, 2010).

Estima-se que a sequência consenso para DPE seja A/G+28-G-A/T-C/T-G/A/C, sendo que há uma menor preferência para G em +24 (Kutach e Kadonaga, 2000). Considerando o ponto de vista funcional, o DPE é sempre dependente de Inr, e sua sequência consenso é sempre relativa à posição A+1 de Inr.

O elemento BRE é um sítio de ligação ao TFIIB localizado imediatamente a montante (BREu) ou a jusante (BREd) do TATA Box (Buchel et al., 1999) e pode aumentar ou diminuir as taxas basais de transcrição (Juven-Gershon e Kadonaga, 2010). A sequência consenso de BRE é G/C-G/CG/A-C-G-C-C (onde o C em posição 3' do BRE é seguido por T na posição 5' do TATA Box).

O número de elementos conhecidos do *core promoter* e também de proteínas específicas que atuam através desses elementos está aumentando. Recentes observações sugerem que os motivos TBP, Mot1p e NC2 exerçam efeitos negativos e positivos sobre a montagem do complexo de transcrição dependendo da presença ou ausência de elementos específicos do *core promoter* (Zhou et al., 2009).

Outras sequências de DNA podem ainda contribuir para a atividade do *core promoter*. Entretanto, na análise de um novo motivo, é importante considerar o seu efeito sobre o processo de transcrição basal, tais como a ligação de TFIID ou TFIIB ao *core promoter*, além da frequência de ocorrência do motivo em vários *core promoter*, e de sua localização específica. Por exemplo, o BRE, o TATA, o Inr e o DPE estão localizados em posições

distintas em relação ao local de início, consistente com seus papéis na montagem da iniciação da transcrição (Smale e Kadonaga 2003; Zhou et al., 2009).

As interações que ocorrem no processo transcricional são complexas, não lineares, e frequentemente contexto-dependentes. Assim, para a análise de um promotor não se deve considerar somente a presença ou ausência de sítios de ligação, mas analisar a posição relativa, a orientação e as sequências dos sítios, bem como os perfis de expressão dos fatores de transcrição e co-fatores cognatos, pois todos contribuem para determinar o perfil de transcrição de um gene.

### 2.6.2 Elementos Regulatórios: *trans* e *cis* elementos

O *core promoter* sozinho não possui atividade transcricional. São os elementos regulatórios individualmente ou em combinação que determinam a expressão espaço temporal e o nível transcricional. Dessa forma, na regulação da expressão gênica ocorre a participação de dois tipos de elementos regulatórios: os elementos *trans* e os elementos *cis* (Alberts et al., 2002) (Figura 3).

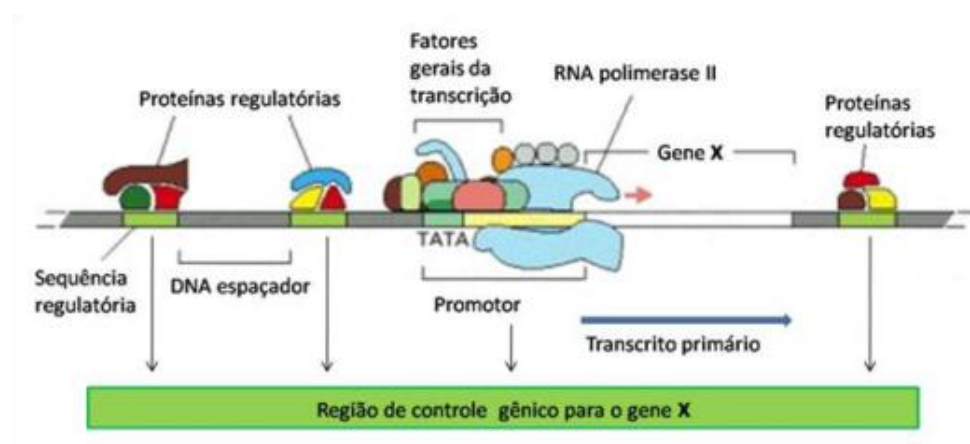
Os fatores *trans* são proteínas ativadoras e fatores gerais de transcrição, que possuem domínios capazes de reconhecer e se ligar a sequências específicas de DNA, os elementos *cis*, e ativar ou reprimir a transcrição. Os FTs atuam em conjunto com outros reguladores transcricionais, incluindo a remodelação da cromatina. Aproximadamente 7% do genoma de plantas são identificados como fatores de transcrição (Udvardi et al., 2007).

Os TFs reconhecem sequências de DNA relativamente curtas, tipicamente 6 a 12 pb, que se localizam próximas ao sítio de início de transcrição, e iniciam (ativadores) ou reprimem (repressores) a transcrição gênica. Cada fator de transcrição tem um conjunto preferido de sítios de ligação do fator de transcrição (*TFBS* - *Transcription Factor Binding Site*). Funcionalmente, para se atingir taxas de transcrição significativas são necessárias associações específicas de FTs às sequências regulatórias *cis* presentes fora do *core promoter* (Carey e Smale, 2000; Lemon e Tjian, 2000; Reineke et al., 2011).

Os elementos *cis* compreendem as sequências reconhecidas pelos fatores *trans*, e são classificados de acordo com sua localização em três grupos: *core promoter*, elementos *cis* proximais e elementos independentes de distância (Maston et al., 2006). Os elementos *cis* proximais compreendem as sequências próximas ao promotor que auxiliam a ligação da RNA polimerase II e modulam sua atividade. Em geral, encontram-se a uma distância de 100 a 200 pb do TSS (Smale e Kadonaga, 2003), e muitos são célula-específicos, realizando um controle da localização e da frequência em que a transcrição deve ocorrer (Lodish et al., 2000).

Os elementos independentes de distância conhecidos como acentuadores (*enhancers*), silenciadores e isoladores podem atuar a centenas ou milhares de pares de base tanto a montante quanto a jusante do TSS, aumentando ou diminuindo a taxa de transcrição (Lodish et al., 2000). Os *cis*-elementos experimentalmente identificados em plantas raramente são maiores que 30 pb apresentando em média 8 pb (Higo et al., 1999; Davuluri et al., 2003; Haberer et al., 2006). Evidências mostram que alguns acentuadores interagem com o *core promoter* dependente de TATA-Inr e outros com o promotor dependente dos elementos Inr-DPE (Juven-Gershon e Kadonaga, 2010).

De modo geral, em plantas, alguns mecanismos evolutivos como a duplicação do genoma geraram diversificação dos motivos de regulação gênica (Reineke et al., 2011). A disposição desses elementos na região de regulação de um gene eucariótico está ilustrado na Figura 3.



**Figura 3.** Elementos regulatórios de um gene eucariótico. Os elementos *cis* são sítios de ligação presentes no DNA reconhecidos pelas proteínas regulatórias (elementos *trans*). (Fonte: Alberts et al., 2002).

## 2.7 Elementos Regulatórios envolvidos em estresses abióticos

Diferentes grupos de *cis*-elementos regulatórios presentes em regiões promotoras e fatores *trans* estão envolvidos na transcrição de genes responsivos a estresse abióticos, sendo alguns deles dependentes de ácido abscísico (ABA-dependentes) e, outros, independentes (ABA-independentes) (Shinozaki e Yamaguchi-Shinozaki, 2000, 2007; Zhu, 2002; Nakashima e Yamaguchi-Shinozaki, 2005, Umezawa et al., 2006; Yamamoto et al., 2011). De modo geral, os FTs interagem com estes *cis*-elementos presentes nas regiões promotoras dos vários genes-alvos relacionados aos estresses abióticos e, assim, regulam sua expressão promovendo a tolerância ao estresse ao qual a planta foi submetida (Agarwal e Jha, 2010). Um grupo de genes controlado por um determinado tipo de fatores de transcrição é conhecido como *regulon* (Saibo et al., 2009).

Embora essa classificação em rotas distintas sugira independência entre as mesmas, em plantas, a regulação transcricional é mediada por um grande número de fatores de transcrição que controlam a expressão de dezenas ou centenas de genes em várias, e às vezes entrelaçadas, cascatas de sinalização (Wellmer e Riechmann, 2005), como demonstrado em ensaios de expressão transiente, onde as proteínas CBF (*C-Repeat Binding Factor* – fator de ligação ao C-repetido) e DREB2 (*Dehydration Responsive Element Binding* – Elemento de ligação responsivo à desidratação) interagiram com as ABFs (*ABA Binding Factors* – fatores de ligação ao ABA) para maior ativação do promotor *rd29A*, ilustrando a intercomunicação entre as vias ABA-dependente e independente (Narusaka et al., 2003).

Na via ABA-dependente, duas principais famílias de FTs estão envolvidas na sinalização e ativação gênica, os bZip (*Basic-domains Leucine Zipper* – domínios básicos de zíper de leucina) e os MYB (*Myelo Blastosis Oncogene*) (Fujita et al., 2011). Os genes bZip contêm, em suas regiões promotoras, o elemento *cis*-atuante ABA responsivo do tipo ABRE (*ABA-Responsive Elements* – elemento responsivo ao ABA), caracterizados pela sequência (C/T ACGTGGC), e reconhecidos pelas proteínas de ligação a ABRE, chamadas de AREB/ABFs (*ABRE Binding Factors* – fatores de ligação ao ABRE) (Guiltinan et al., 1990; Shinozaki e Yamaguchi-Shinozaki, 2000,

Finkelstein et al., 2002; Fujita et al., 2005; Furihata et al., 2006; Yoshida et al., 2010) que são de fato, os potenciais conectores de uma rede de sinalização onde múltiplas vias podem convergir (Halford e Hey, 2009). Análises experimentais *in silico* sugerem que são permitidas variações dentro do núcleo ACGT de modo que a substituição por GCGT ou AAGT resulta na ligação por diferentes AREBs/ABFs (Choi et al., 2000; Hobo et al., 1999, Sarkar e Lahiri, 2013).

A expressão de genes responsivos ao ABA necessita de múltiplos ABREs ou a combinação de um ABRE com um elemento de acoplamento (*Coupling Element*) (Yoshida et al., 2010). Os elementos transcricionais AREBs/ABFs são regulados positivamente pelas quinases SnRK2s e algumas outras proteínas dependentes de cálcio (SnRK3s/CIPKs e CDPKs/CPKs) (Halford & Hey, 2009). Foram identificados em *Arabidopsis* nove homólogos da família de genes AREBs/ABFs, e alguns genes dessa família são relatados em diferentes vias de sinalização em resposta aos estresses abióticos como *ABF1* no estresse ao frio; *AREB1/ABF2* em estresse salino, seca, calor e glicose; *ABF3* no estresse salino e seca e; *AREB2/ABF4* em estresses de frio, salinidade e seca (Kim et al., 2004; Fujita et al., 2005; Yoshida et al., 2010; Fujita et al., 2011).

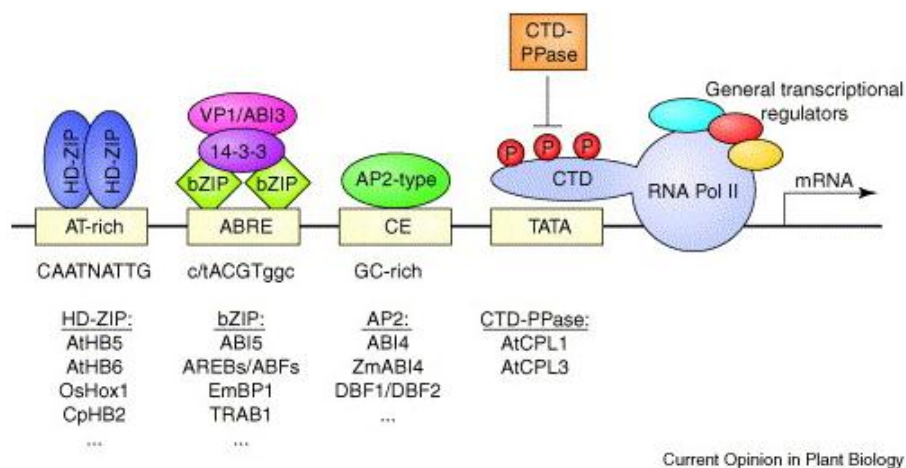
A importância da família AREB/ABF na tolerância celular a estresses abióticos tem sido reforçada por resultados obtidos com PGMs. A super expressão de *OsZIP23*, um membro da subfamília AREB/ABF, em linhas de arroz GM aumentou significativamente a tolerância à seca e alta salinidade no estágio reprodutivo (Xiang et al., 2008). Também foi observado aumento na tolerância à seca e ao calor em plantas transgênicas de *Arabidopsis* super expressando *OsAREB1* (Jin et al., 2010). A superexpressão da forma constitutivamente ativa de *AtAREB1* em plantas de soja transgênicas proporcionou melhor crescimento e desempenho fisiológico das mesmas quando sob déficit hídrico (Barbosa et al., 2012).

Outros fatores de transcrição estão ainda relacionados à genes de resposta lenta ABA-induzidas. Esses genes possuem, em suas sequências promotoras, os elementos MYBR e MYCR, reconhecidos pelas sequências A/TAACCA e C/TAACG/TG; e CANNTG, respectivamente. Estes elementos *cis* são reconhecidos pelos fatores de transcrição MYB e MYC (*Myelo Cytomatosis*

*Oncogene*) (Shinozaki e Yamaguchi-Shinozaki, 2000; Guiltinan et al., 1990; Finkelstein et al., 2002; Fujita et al., 2005; Furihata et al., 2006; Yoshida et al., 2010). Estes FTs possuem um domínio de ligação ao DNA funcional conservado, que consiste em três repetições (R), cada uma contendo de 50-53 aminoácidos que formam uma estrutura em alfa-hélices (hélice-alfa-hélice), que se liga ao DNA (Du et al., 2009). As proteínas MYB em plantas são categorizadas em três grandes grupos de acordo com o número de repetições presentes, R1R2-MYB com duas repetições adjacentes, R1R2R3-MYB com três repetições adjacentes e MYB contendo apenas uma repetição.

Vários genes da família MYB têm sido relatados como responsivos a estresses abióticos. Liao e colaboradores (2008) identificaram 156 genes *GmMYB*, sendo que a expressão de 43 genes foi alterada com a aplicação exógena de ABA, estresses salino, seca e frio. Também em PGMs, a superexpressão de *MYB15* resultou no aumento da tolerância à seca e salinidade em *Arabidopsis* (Ding et al., 2009). Em batata, a superexpressão de *StMYB1R-1* promoveu aumento da tolerância à seca (Shin et al., 2011).

Na Figura 4 são apresentados os principais elementos *cis*-regulatórios e fatores de transcrição envolvidos na expressão gênica ABA-regulada.



**Figura 4.** O complexo transcripcional consiste de RNA polimerase II, em associação com os reguladores gerais de transcrição, incluindo a proteína de ligação TATA Box. A maquinaria de transcrição é controlada por meio da ligação do FT aos *cis* elementos reguladores específicos no promotor. O elemento de regulação de ABA (ABRE) é contatado por FTs bZIP diméricos. O elemento de acoplamento (CE) é alvo de FTs do tipo AP2, que aumentam a transcrição ABA-regulada. As proteínas HD-ZIP se ligam a sequências pseudopalindrômicas ricas em AT. RNA Pol II é regulada pela fosforilação de seu domínio carboxil terminal (CTD). A desfosforilação de uma fosfatase de proteína específica-CTD (CTD ADPG), diminui a expressão dos genes induzíveis por ABA. (Fonte: Himmelbach et al., 2003).

Existem também vias de resposta independentes de ABA, onde vários genes de resposta são diretamente induzidos por seca, temperaturas extremas e salinidade, não necessitando de ABA para serem ativados, embora esses genes possam também responder à adição de ABA (Shinozaki e Yamaguchi-Shinozaki, 1996; Bray, 1997; Xiong et al., 2002). Nesta via também são descritas duas famílias de TFs envolvidas na sinalização e ativação gênica, os CBF/DREB (*C-Repeat-Binding Factor*)/(*Dehydration Responsive Element Binding Protein*) e os NAC (*NAN, ATAF and CUC*) e ZF-HD (*Zinc-Finger Homeodomain*) (Nakashima et al., 2009; Saibo et al., 2009).

Os promotores dos genes da família CBF/DREB contêm o elemento regulatório *cis* DRE (*Dehydration-Responsive Element*); de sequência consenso CCGAC, na qual se ligam os fatores de transcrição conhecidos como DREB (*Dehydration Responsive Element Binding Protein*) ou CBF (*C-Repeat-Binding Factor*). Os FTs DREB estão no topo da cascata de eventos moleculares que finalizam com respostas de defesa contra a desidratação celular. Os membros da família DREB1/CBF, como DREB1B/CBF1, DREB1C/CBF2 e DREB1A/CBF3, pertencem à família multigênica AP2/EREBP2 e são, eles próprios, estresse induzidos. Os fatores de transcrição DREB1 apresentam um maior envolvimento na resposta ao frio, enquanto que o fator de transcrição DREB2 é induzido em resposta à seca, salinidade e ao calor (Liu et al., 1998; Matsukura et al., 2010). Uma segunda rota independente de ABA constituída pela proteína CBF4 também é relacionada em resposta à seca (Haake et al., 2002).

A superexpressão de *DREB1B* em batatas resultou em um aumento significativo da tolerância ao congelamento e seca, com aumento no teor de prolina e teor relativo de água superior sob condições de congelamento e déficit hídrico (Movahedi et al., 2012). A expressão constitutiva de genes *DREB1A*, *DREB1B* e *DREB1C* em plantas transgênicas de *Arabidopsis* ativou a expressão de genes responsivos a estresses abióticos e aumentou a tolerância à seca, salinidade e ao congelamento (Gilmour et al., 2004).

Experimentos realizados em casa de vegetação com plantas de soja GM com o gene *AtDREB1A* controlado pelo promotor estresse induzido *rd29A* demonstraram que as plantas apresentaram maior condutância estomática, taxa fotossintética e transpiratória quando comparadas com plantas não GM,

sugerindo a ativação de mecanismos que podem levar a maior tolerância ao déficit hídrico (Polizel et al., 2011). Linhagens de soja contendo essa mesma construção foram avaliadas tanto em casa de vegetação quanto em condições de campo e alguns componentes de produção foram aumentados quando a seca foi imposta durante a fase vegetativa de desenvolvimento (Rolla et al., 2013).

Análises de microarranjo em *Arabidopsis* GM sugerem que a superexpressão de *DREB2A-CA* induz a expressão de genes responsivos à seca, salinidade e calor. Essas plantas transgênicas também apresentaram aumento da tolerância ao calor, enquanto plantas com *DREB2A* silenciado apresentaram um decréscimo significativo na termotolerância (Sakuma et al., 2006). Plantas de *Arabidopsis* transgênicas superexpressando o gene *OsDREB2B* mostraram aumento da expressão de genes *DREB2A* e melhor tolerância ao estresse de seca e ao choque térmico (Matsukura et al., 2010; Mizoi et al., 2013).

Uma das maiores famílias de fatores de transcrição específicos de plantas é a NAC, com 117 membros identificados em *Arabidopsis* e 151 em arroz (Nuruzzaman et al., 2010). Estes FTs contêm um domínio N-terminal de ligação ao DNA altamente conservado e um domínio C-terminal diversificado (Hu et al., 2008). O elemento regulatório *cis* NACR reconhecido pelos fatores de transcrição NAC apresenta o motivo CATGTG (Tran et al., 2004). Alguns genes NAC já foram descritos como envolvidos na resposta a vários estresses ambientais. A análise de microarranjo em plantas transgênicas de *Arabidopsis* possibilitou a identificação de vários genes alvo dos ativadores de transcrição, tais como, *ANAC019*, *ANAC055* e *ANAC072*. As plantas transgênicas também apresentaram uma melhor tolerância à seca (Tran et al., 2004). Hu e colaboradores (2008), identificaram ainda que *SNAC1* é ativado principalmente em células guarda sob desidratação.

Em soja, 101 proteínas contendo domínios NAC foram identificadas como envolvidas em resposta a estresses abióticos e em eventos de morte celular programada. Também foi avaliada a expressão de seis genes NAC em resposta a uma variedade de condições de estresse. O gene *GmNAC2* foi fortemente induzido por estresse osmótico, já *GmNAC3* e *GmNAC4* além do estresse osmótico também foram induzidos por ABA, ácido jasmônico e

salinidade (Pinheiro et al., 2009). A análise do promotor do gene *ERD1* (*Early Responsive to Dehydration stress 1*), relacionado com a resposta rápida durante estresse hídrico, mostrou que FTs pertencentes à família NAC e ZF-HD são importantes para a ativação deste gene (Tran et al., 2007), demonstrando novamente a interrelação entre diferentes vias de resposta.

Outra família de FTs regulada por fatores ambientais, como seca, altas temperaturas e estresse osmótico, apresentando especificidade em diferentes tecidos e órgãos da planta (Agalou et al., 2008; Elhiti e Stasolla, 2009) é a HD $\beta$  Zip, caracterizada por apresentar dois domínios estruturais: o homeodomínio (HD) e o zíper de leucina (Zip) (Elhiti e Stasolla, 2009). Estes FTs são encontrados exclusivamente no reino vegetal e apresentam além da função estresse-responsivo, papéis específicos no desenvolvimento das plantas (Son et al., 2010; Ré et al., 2011).

Proteínas *Zinc-Finger* (ZFPs)/Cys2/His2 (C2H2) também desempenham um papel fundamental na regulação das respostas de defesa das plantas a condições de estresse biótico e abiótico (Singh et al., 2010). Estes FTs são encontrados em abundância em vegetais e animais, e apresentam motivos de sequência em que cisteínas e/ou histidinas coordenadas com átomo de zinco formam estruturas peptídicas necessárias para as suas funções específicas. A super expressão de um tipo TFIIIA C2H2 ZF, ZFP252, aumentou a tolerância de plantas de arroz à seca, e essa resposta foi correlacionada com a indução de *OsDREB1A* e com um maior acúmulo de prolina e açúcares solúveis livres (Xu et al., 2009).

Os WRKYs representam outra importante família de FTs envolvidos na regulação de múltiplos processos em plantas, incluindo resposta a estresses abióticos (Rushton et al., 2010). Os membros desta família apresentam como principal característica um domínio de ligação ao DNA, denominado “domínio WRKY”, tipicamente composto por 60 aminoácidos, dentre estes a sequência conservada WRKYGQK (triptofano – arginina – lisina – tirosina – glicina – glutamina – lisina) na sua porção N-terminal, além de um motivo “*zinc-finger-like*” na sua porção C-terminal (Eulgem et al., 2000). O domínio WRKY possui como alvo principal de ligação o *cis* elemento do tipo “W-boxes” (C/TTGACT/C), embora tenham sido identificados locais de ligação alternativos, sugerindo que as preferências do sítio de ligação também são, em

parte, determinado pelas sequências de DNA adicionais adjacentes fora do motivo TTGACY (Pandey et al., 2009).

Recentemente, os TFs WRKY foram inseridos na via de sinalização ABA dependente. Alguns deles são reguladores positivos do fechamento dos estômatos mediado por ABA e, portanto, das respostas à seca. Estudos de ligação ao promotor *in vitro* mostram que os genes-alvo para TFs WRKY que estão envolvidos na sinalização ABA incluem genes bem conhecidos como ABA-responsivos, tais como, *ABF2*, *ABF4*, *ABI4*, *ABI5*, *MYB2*, *DREB1A*, *DREB2A* e *RAB18*. Outros genes estresse-induzidos já bem caracterizados tais como *RD29A* e *COR47* também foram encontrados em vias de sinalização a jusante de TFs WRKY (Rushton et al., 2012).

Os primeiros indícios do envolvimento desses FTs com resposta ao estresse abiótico foram obtidos através de perfis de transcrição, no entanto, análises funcionais forneceram evidências mais diretas. O promotor de choque térmico HSP101 orientando a superexpressão de *OsWRKY11* conduziu a uma melhor tolerância ao calor e à seca em arroz (Wu et al., 2009). Da mesma forma, a superexpressão de *OsWRKY45* em *Arabidopsis* resultou em maior tolerância à salinidade e à seca, além de maior resistência a doenças (Qiu et al., 2009). Plantas de *Arabidopsis* superexpressando os genes *GmWRKY21* foram mais tolerantes ao estresse pelo frio, enquanto plantas de *Arabidopsis* superexpressando *GmWRKY54* foram mais tolerantes ao estresse salino e a seca, e a superexpressão de *GmWRKY13* resultou no aumento da sensibilidade ao estresse salino e manitol (Li et al., 2009).

A família de proteínas WRKY inicialmente foi categorizada com base no número de domínios WRKY e nas características do motivo *Zinc-finger*, em três grupos distintos. O grupo (I) tem dois domínios WRKY, o grupo (II), tem um domínio WRKY contendo o mesmo motivo *zinc-finger* Cys2-His2, e o grupo (III) possui um domínio WRKY contendo diferentes motivos *zinc-finger* Cys2-His/Cys Cys2-His2. As proteínas WRKY do grupo II são ainda divididas em subgrupos de A a E com base em motivos estruturais conservados fora do domínio WRKY (Eulgem et al., 2000). Posteriormente, com base na análise filogenética mais precisa, Zhang e Wang (2005) classificaram os fatores WRKY em Grupos I, IIa + IIb, IIc, IId + IIe e III. As proteínas WRKY, tanto para o domínio WRKY altamente conservado quanto para os motivos *Zinc-Finger*

Cys2-His2 ou Cys2-His/Cys, ainda contém as estruturas de zíper de leucina, regiões ricas em serina-treonina, região rica em glutamina, região rica em prolina, e os domínios quinase e TIR-NBS-LRRs (Zhang e Wang, 2005).

Recentemente novos FTs foram relatados como envolvidos na regulação da expressão gênica em resposta a estresses abióticos. Os NF-Y (*Nuclear Factor Y*) são fatores de transcrição, também conhecidos como proteínas heme-associadas HAPs (*Heme Associated Proteins*) e fatores de ligação CCAAT Box CBFs (*CCAAT Binding Factor*), estão emergindo rapidamente como reguladores importantes de numerosas respostas a estresses ambientais. NF-Y é composto por três subunidades, NF-YA, NF-YB, e NF-YC que se ligam como complexos heterotriméricos a CCAAT. Em plantas, tanto em monocotiledôneas como em dicotiledôneas, cada tipo de subunidade é codificada por uma família de 10 genes. Considerando que o CCAAT Box é um dos *cis*-elementos mais comuns presentes em promotores eucarióticos, e também a natureza heterotrimérica do NF-Y, é possível que ocorra a formação de muitos complexos NF-Y, levando a um sistema flexível de combinações de fatores de transcrição que podem promover pequenos ajustes a várias condições ambientais (Petroni et al., 2012).

A interação entre fatores de transcrição é importante na regulação do grupo de genes que está associado com o mesmo padrão de expressão. A análise comparativa de transcriptomas de folhas e raízes de arroz sob condições de frio, seca e estresse salino identificou elementos *cis*-regulatórios altamente conservados em promotores de genes *up*-regulados sob essas condições. Os motivos conservados identificados como envolvidos nas vias de sinalização de estresse abiótico foram pertencentes a famílias de FTs como DREB, ERF, WRKY, MYB, bZIP, helix-loop-helix e NAC (Venu et al., 2013).

Os FTs atuam ainda no controle da expressão em conjunto com outros reguladores transcricionais, incluindo a remodelação da cromatina. As modificações na cromatina e geração de pequenos RNAs (sRNA - *small RNA*) têm sido apresentados como envolvidos no controle transcricional e pós transcricional da expressão de genes responsivos a estresses (Angers et al., 2010). Kucheski e colaboradores (2011) investigaram o possível envolvimento de miRNAs no processo de respostas ao déficit hídrico e à ferrugem asiática em soja, e detectaram 256 miRNA caracterizados em 24 novas famílias e 22

famílias já relatadas em soja, a validação de alguns miRNA revelou diferentes perfis de expressão, onde a maioria dos miRNAs analisados foram *up*-regulados durante o déficit hídrico nas plantas sensíveis e *down*-regulados para o genótipo tolerante.

Outro fator, a metilação de DNA, controlada por fluxos hormonais, é também influenciada por vários fatores bióticos e abióticos (Zhang et al., 2012). Estudos recentes indicam que a dinâmica molecular como homodimerização e heterodimerização específicas, bem como a flexibilidade modular e modificações pós-transducionais determinam a especificidade funcional dos FTs na adaptação ambiental. Processos epigenéticos como a metilação do DNA e modificações pós-traducionais de histonas têm grande influência na eficiência de expressão de genes estresse induzidos. Reguladores centrais como NAC, WRKY e proteínas *Zinc Finger* podem ser modulados por mecanismos como miRNA (micro RNAs) (Goldack et al., 2011).

Os níveis de expressão e metilação do DNA foram avaliados em 49 fatores de transcrição diferencialmente expressos durante a imposição de estresse salino na cultura da soja. Foi observada uma correlação significativa entre os perfis de metilação e expressão de MYB, b-ZIP e famílias de genes de fatores de transcrição AP2/DREB (Song et al., 2012). Em arroz, também foi relatado uma progressão dos mecanismos epigenéticos em conjunto com a ação de fatores de transcrição na regulação das respostas a estresses abióticos (Santos et al., 2011).

Shaik e Ramakrishna (2012) em estudo para identificar prováveis conjuntos de genes controlados epigeneticamente em resposta ao estresse hídrico, e mapear as regiões de metilação presentes em regiões codantes e promotoras nestes genes, analisaram mais de 5.000 genes responsivos a seca em arroz. A análise de agrupamento com relação aos atributos epigenéticos e microRNAs mostrou o enriquecimento de conjuntos *up*-regulados por mecanismos de tolerância à seca, e conjuntos *down*-regulados foram associados a processos de tolerância à seca. Estes pesquisadores apresentaram uma nova abordagem para a compreensão da regulação da resposta ao estresse através da combinação de vários recursos de bioinformática e classificação de subconjuntos de genes envolvidos nas principais vias de resposta.

## 2.8 Promotores constitutivos e estresse induzidos

A identificação de genes de resposta ao déficit hídrico não é garantia para sua utilização com êxito na obtenção de plantas geneticamente modificadas com objetivo de tolerância à um estresse específico. Para se obter sucesso com esta estratégia, o padrão de regulação da expressão gênica feita principalmente pelo promotor é fundamental para iniciar ou interromper a expressão do gene, uma vez que é através deste mecanismo que as células modulam as respostas aos sinais externos ambientais (Wolf et al., 2010; Mastroeni et al., 2011; Zhang et al., 2011).

Os promotores constitutivos dirigem a expressão contínua em praticamente todos os tecidos e são, em grande parte, se não totalmente, independente de fatores ambientais e fase de desenvolvimento. Como sua expressão normalmente não é condicionada por fatores endógenos, promotores constitutivos são normalmente ativos entre espécies e mesmo entre reinos diferentes. Enquanto que nos promotores induzidos o desempenho não está condicionado a fatores endógenos, mas às condições ambientais e estímulos externos. Dentro deste grupo, encontram-se os promotores modulados por fatores abióticos, como a luz, a seca, a salinidade, os níveis de oxigênio, o calor e o frio.

Entre os promotores constitutivos amplamente empregados na produção de plantas geneticamente modificadas destaca-se o CaMV 35S (Odell et al., 1985), do vírus causador do Mosaico da Couve Flor. Esse promotor é ativo na maioria dos tecidos e fases de desenvolvimento vegetal, entretanto, possui limitações com relação ao silenciamento transcricional do gene sob sua regulação, pois as células vegetais podem reconhecer essa sequência como estranha e promover sua metilação, excisão ou remodelamento da cromatina na região onde houve a inserção (Potenza et al., 2004).

As limitações apresentadas pelo promotor 35S podem ser contornadas com a utilização de outros promotores como os que codificam a ubiquitina de milho (ZmUbi-1) (Christensen et al., 1992), a actina de arroz (OsAct1) (McElroy et al., 1990) e o citocromo c (Cominelli e Tonelli, 2010). Ainda assim podem ocorrer mecanismos de silenciamento, ou diferenças nos níveis de ativação quando utilizados em espécies filogeneticamente distantes da qual o promotor

foi isolado, pois os sistemas regulatórios são únicos, podendo não agir da maneira prevista em espécies heterólogas (Potenza et al., 2004, Butaye et al., 2005).

Genes sob o controle de promotores constitutivos são ativos na maioria das células durante todo o ciclo celular em vários níveis (Twyman et al., 2003), e essa característica pode se tornar uma limitação dependendo da estratégia de engenharia genética almejada. Por exemplo, aumentar a expressão de genes de defesa de forma constitutiva normalmente direciona o metabolismo da planta para a síntese destes genes, e este gasto energético pode comprometer o crescimento ou causar problemas no desenvolvimento vegetal, prejudicando características finais relacionadas a produtividade agrônômica da cultura (Chen et al., 2009; Cominelli e Tonelli, 2010; Chen et al., 2013).

Pelo seu papel insubstituível na modulação da expressão espaço-temporal de genes que interagem com TFs, promotores de plantas têm atraído cada vez mais atenção (Liu et al., 2013b). Alguns promotores específicos já foram isolados e até mesmo sinteticamente construídos visando sua aplicação na produção de plantas geneticamente modificadas. Um aspecto limitante, entretanto, refere-se às questões de propriedade intelectual que restringem a utilização de alguns promotores de forma mais abrangente (Rushton et al., 2002).

Nos últimos anos muitos trabalhos têm sido desenvolvidos no sentido de se isolar e avaliar promotores induzidos específicos (Xu et al., 2010; Tittarelli et al., 2009; Kato et al., 2010; Yi et al., 2010; Hernandez-Garcia et al., 2010; Noh et al., 2012; Wang et al., 2013). O emprego destes promotores específicos com sequências regulatórias induzíveis ou órgão específico capazes de realizar o controle preciso da expressão do gene de interesse pode eliminar ou reduzir efeitos negativos da expressão constitutiva (Cominelli et al., 2012). Neste caso, a expressão do transgene seria limitada ao período de duração do estímulo e/ou direcionada para o local, ou momento em que o estímulo foi produzido. Por exemplo, a expressão constitutiva do gene *DREB1/CBF3* conferiu tolerância ao estresse de déficit hídrico em plantas GMs, mas provocou severo atraso no crescimento e floração tardia. No entanto, quando esse mesmo gene foi expresso sob o controle do promotor estresse induzido *rd29A*, estas

características indesejáveis não foram mais observadas (Kasuga et al., 1999; Kasuga et al., 2004; Morran et al., 2011; Saint Pierre et al., 2012).

Kovalchuck e colaboradores (2013) também demonstraram que o uso de promotores induzidos por frio reduziu os efeitos pleiotrópicos negativos no fenótipo de plantas GM com o gene *TaDREB3* em comparação com os resultados observados quando esse mesmo gene foi constitutivamente expresso. Neste contexto, a identificação de promotores que regulam a expressão diferencial de genes no momento, intensidade e tecido de interesse representa um dos maiores objetivos da engenharia genética de plantas na agricultura moderna.

## **2.9 Uso de ferramentas computacionais para identificação de genes de interesse biotecnológico**

Nos últimos anos, uma abundância de dados vêm sendo gerada com um número cada vez maior de sequências completas de genomas de vários organismos e perfis mRNA obtidos através de Microarranjo de DNA e NGS. A integração destes dados permitiram a identificação de novas interações e elementos de regulação da expressão gênica espacial e temporal em resposta a estresses abióticos (Cramer et al., 2011; Hirai et al., 2004; Obata e Fernie 2012; Weckwerth, 2011).

Apesar da quantidade de ferramentas e bancos de dados disponíveis, atualmente, a identificação de sítios de ligação específicos dos FTs em promotores ainda continua sendo um dos desafios em bioinformática, principalmente devido ao pequeno tamanho dos sítios, do número de ocorrência, da posição em relação ao local de início de transcrição (TSS) e também do grau de conservação dos FTs (Nain et al., 2012). Nos últimos anos, algumas novas abordagens que integram a genômica comparativa, estrutural e funcional com algoritmos computacionais têm sido desenvolvidas.

Com base no tipo de informação sobre a sequência de DNA para encontrar os motivos, os algoritmos são classificados em três classes principais: (1) os métodos que utilizam sequências de promotores de genes co-regulados a partir de um único genoma, (2) os métodos que usam o promotor de sequências ortólogas de um único gene, a partir de várias espécies também conhecidos como *footprinting* filogenético e (3) os métodos que utilizam

sequências promotoras de genes co-regulados, bem como a filogenética de *footprinting* (Das e Dai, 2007).

Também se dispõe hoje, de uma série de bases de dados específicos de soja como o Phytozome, o Soybase (Grant et al., 2010), banco de dados de genoma da soja (Shultz et al., 2006) e banco de dados de Microarranjos de DNA (Alkharouf et al., 2004). Além de bancos de dados específicos para FTs de soja como o SoyDB (Wang et al., 2010), e banco de dados que integram FTs de diferentes culturas como o TreeTFDB (Mochida et al., 2013), e ainda programas *on-line* como MEME (*M*ultiple *E*M for *M*otif *E*licitation) (Bailey e Elkan, 1995), Place (Higo et al., 1999), PlantCare (Rombauts et al., 1999) que podem ser utilizados para análise de *motifs* e *cis* elementos regulatórios.

Para lidar com a gestão e integração de dados da genômica, transcriptômica, proteômica e metabolômica de soja juntamente com a anotação da função dos genes e via biológica, bancos de dados como o SoyKB (*Soybean Knowledge Base*) foram projetados (Joshi et al., 2012). Este banco contém informações sobre genes, microRNAs, metabólitos e polimorfismos de nucleotídeo único (SNPs – *Single Nucleotide Polymorphism*).

O banco de dados agriGO (Du et al., 2010) especialmente desenvolvido para espécies agrícolas realiza a análise ontológica e agrupa os resultados em três categorias: processo biológico, função molecular e componente celular. Este banco permite ainda a análise de enriquecimento simples e paramétrica em conjunto de dados. Há ainda, outras ferramentas que agrupam os transcritos em diversas categorias funcionais de acordo com o termo GO (*Gene Ontology*), obtido a partir do banco de dados Gene Ontology (Ashburner et al., 2000) como o MapMan (Usadel et al., 2005).

Diversos servidores de bioinformática ou bancos de dados de promotores de plantas também foram desenvolvidos, embora a maioria com foco na anotação de sítios de ligação de fatores de transcrição em um único gene negligenciando alguns elementos regulatórios importantes como repetições em *tandem* e ilhas CpG/CpNpG em regiões promotoras. Algumas ferramentas como o PlantPAN (*Plant Analysis Promotor Navigator*) foram ainda desenvolvidas para detectar a co-regulação dos fatores de transcrição em um grupo de promotores de genes, além de permitir a identificação de outras funções reguladoras de um promotor de plantas, como as ilhas CpG/CpNpG e

as repetições em série. Utilizando estes recursos, elementos reguladores das regiões conservadas dos promotores de genes homólogos podem ser detectados e apresentados (Chang et al., 2008).

Dentro de todo o contexto apresentado, a identificação de *cis*-elementos presentes nos genes responsivos a estresses abióticos, em especial a deficiência hídrica, pode ajudar na compreensão de padrões espacial e temporal da expressão gênica responsiva ao estresse em plantas, e facilitar a clonagem de promotores estresses induzidos, permitindo sua utilização na engenharia genética para obtenção de plantas GMs mais tolerantes a períodos de seca, sem comprometimento de crescimento/desenvolvimento vegetal e mantendo importantes componentes agrônômicos para produtividade final da cultura.

### 3. Objetivos

#### 3.1 Objetivo Geral

Caracterizar as regiões promotoras de genes diferencialmente expressos em plantas de soja sob condições de déficit hídrico progressivo e de curto período.

#### 3.2 Objetivos Específicos

- Identificar transcritos diferencialmente expressos (TDEs) em condição de déficit hídrico progressivo utilizando a técnica de microarranjos de DNA;
- Identificar transcritos diferencialmente expressos em condição de déficit hídrico de curto período utilizando a técnica de RNA-Seq;
- Classificar a função biológica dos transcritos e caracterizar o perfil de expressão observado em resposta às duas condições de déficit hídrico estudadas;
- Identificar os transcritos induzidos/reprimidos comuns entre os déficits hídricos de curto período e o progressivo e;
- Caracterizar *in silico* os *cis* elementos presentes na sequência promotora dos transcritos diferencialmente expressos identificados em ambas às condições de déficit hídrico.

#### 4. Referências

AGALOU, A.; PURWANTOMO, S.; OVERNAS, E.; JOHANNESSON, H.; ZHU, X.; ESTIATI, A.; DE KAM, R.J.; ENGSTROM, P.; SLAMET-LOEDIN, I.H.; ZHU, Z.; WANG, M.; XIONG, L.; MEIJER, A.H.; OUWERKERK, P.B. A genome-wide survey of HD-Zip genes in rice and analysis of drought-responsive family members. *Plant Molecular Biology*, 66:87-103, 2008.

AGARWAL, P. K.; JHA, B. Transcription factors in plants and ABA dependent and independent abiotic stress signaling. *Biologia Plantarum*, 54:201-212; 2010.

AINSWORTH, E. A.; YENDREK, C. R.; SITCH, S.; COLLINS, W. J.; EMBERSON, L.D. The effects of tropospheric ozone on net primary productivity and implications for climate change. *Annual Review of Plant Biology*. doi:10.1146/annurev-arplant-042110-103829; 2012.

ALBERTS, B.; JOHNSON, A.; JULIAN LEWIS, J. RAFF, M.; ROBERTS, K.; WALTER P. Control of gene Expression. In: (Ed.) *Molecular Biology of the Cell*. 4, New York: Garland Science, 375-466, 2002.

ALKHAROUF, N. W.; MATTHEWS, B. F. SGMD: the Soybean Genomics and Microarray Database. *Nucleic Acids Res*, 32(1):D398-D400, 2004.

AMARA, I.; CAPELLADES, M.; LUDEVID, M. D.; PAGÈS, M.; GODAY, A. Enhanced water stress tolerance of transgenic maize plants over-expressing LEA Rab28 gene. *Journal of Plant Physiology*, 170(9):864–873, 2013.

ANGERS, B.; CASTONGUAY, E.; MASSICOTTE, R. Environmentally induced phenotypes and DNA methylation: how to deal with unpredictable conditions until the next generation and after. *Molecular Ecology* 19:1283–1295, 2010.

ASHBURNER, M.; BALL, C. A.; BLAKE, J. A.; BOTSTEIN, D.; BUTLER, H.; CHERRY, J. M.; SHERLOCK, G. Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25-29, 2000.

ATKINSON, N. J.; URWIN, P. E. The interaction of plant biotic and abiotic stresses: from genes to the field. *Journal of Experimental Botany*, 63:3523-3543, 2012.

AUGUSTINE, R.; ARYA, G. C.; NAMBIAR, D. M.; KUMAR, R.; BISHT, N. C. Translational genomics in Brassica crops: challenges, progress, and future prospects. *Plant Biotechnology Reports*, 1-17, 2013.

BAILEY, T. L.; ELKAN, C. The value of prior knowledge in discovering motifs with MEME++. Technical Report CS95- 413, Department of Computer Science, University of California, San Diego, 1995.

BALDERAS-HERNÁNDEZ, V. E.; ALVARADO-RODRÍGUEZ, M.; FRAIRE-VELÁZQUEZ, S. Conserved versatile master regulators in signaling pathways in response to stress in Plants, *AoB plants*, v. 5, 2013.

BARBOSA, E. G. G.; LEITE, J. P.; MARIN, S. R. R.; MARINHO, J. P.; CARVALHO, J. F. C.; PAGLIARINI, R. F.; FARIAS, J. R. B.; NEUMAIER, N.; MARCELINO-GUIMARÃES, F. C.; OLIVEIRA, M. C. N.; YAMAGUCHI-SHINOZAKI, K.; NAKASHIMA, K.; MARUYAMA, K.; KANAMORI, N.; FUJITA, Y.; YOSHIDA, T.; NEPOMUCENO, A. L. Overexpression of the ABA-Dependent AREB1 Transcription Factor from *Arabidopsis thaliana* Improves Soybean Tolerance to Water Deficit. *Plant Molecular Biology*, Rep doi: 10.1007/s11105-012-0541-4, 2012.

BARTELS, D.; SUNKAR, R. Drought and salt tolerance in plants. *Critical Reviews in Plant Sciences*, Boca Raton, 24(1):23-58, 2005.

BAUMANN, M.; PONTILLER, J.; ERNST, W. Structure and Basal Transcription Complex of RNA Polymerase II Core Promoters in the Mammalian Genome: An Overview. *Molecular Biotechnology*, 45(3):241-247, 2010.

BERNARD, V., BRUNAUD, V., & LECHARNY, A. TC-motifs at the TATA-box expected position in plant genes: a novel class of motifs involved in the transcription regulation. *BMC Genomics*, 11(1):166, 2010.

BERTOLLI, S. C.; RAPCHAN, G. L.; SOUZA, G. M. Photosynthetic limitations caused by different rates of water-deficit induction in Glycine max and Vigna unguiculata. *Photosynthetica*. DOI: 10. 1007/s11099-012-0036-4, 2012.

BHATNAGAR-MATHUR, P.; VADEZ, V.; SHARMA, K. K. Transgenic approaches for abiotic stress tolerance in plants: retrospect and prospects. *Plant Cell Reports*, 27:411–424, 2008.

BRAY, E. A. Plant responses to water deficit. *Trends Plant Science*, 2:48-54, 1997.

BREATHNACH, R.; CHAMBON, P. Organization and expression of eukaryotic split genes coding for proteins. *Annual Review of Biochemistry.*, 50:349-383, 1981.

BURKE, E. J.; BROWN, S. J.; CHRISTIDIS, N. Modeling the recent evolution of global drought and projections for the twenty-first century with the Hadley centre climate model. *Journal Hydrometeorology*, 7:1113-1125, 2006.

BURKE, T. W.; KADONAGA, J. T. The downstream core promoter element, DPE, is conserved from *Drosophila* to humans and is recognized by TAFII60 of *Drosophila*. *Genes & Development*, 11:3020-3031, 1997.

BUTAYE, K. M.; CAMMUE, B. P.; DELAURÉ, S. L.; DE BOLLE, M. F. Approaches to minimize variation of transgene expression in plants. *Molecular Breeding*, 16(1):79-91, 2005.

BUTLER, J.E.F.; KADONAGA, J.T. The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes & Development*, 16:2583-2592, 2002.

CAREY, M.; SMALE, S.T. Transcription regulation in eukaryotes: concepts, strategies, and techniques. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NewYork, 2000.

CHAE, L.; SUDAT, S.; DUDOIT, S.; ZHU, T.; LUAN, S. Diverse transcriptional programs associated with environmental stress and hormones in the

Arabidopsis receptor-like kinase gene family. *Molecular Plant*, 2(1):84-107, 2009.

CHANG, W. C.; LEE, T. Y.; HUANG, H. D.; HUANG, H. Y.; PAN, R. L.. PlantPAN: Plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups. *BMC Genomics*, 9(1): 561, 2008.

CHEN, M.; XU, Z.; XIA, L.; LI, L.; CHENG, X.; DONG, J.; WANG, Q.; MA, Y. Cold induced modulation and functional analyses of the DRE-binding transcription factor gene, GmDREB3, in soybean (*Glycine max L.*). *Journal of Experimental Botany*, 60(1):121-35, 2009.

CHEN, Z.; WANG, J.; YE, M. X.; LI, H.; JI, L. X.; LI, Y.; CUI, D. Q.; LIU, J.M.; AN, X. M. A Novel Moderate Constitutive Promoter Derived from Poplar (*Populus tomentosa* Carrière). *International Journal of Molecular Sciences*, 14(3):6187-6204, 2013.

CHEONG, Y. H.; CHANG, H. S.; GUPTA, R.; WANG, X.; ZHU, T.; LUAN, S. Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in *Arabidopsis*. *Plant Physiology*, 129:661-677, 2002.

CHOI, I. H.; HONG, H. J.; HA, O. J.; KANG, Y. J.; KIM, Y. S. ABFs, a family of ABA responsive element binding factors. *Journal of Biological Chemistry*, 275:1723-1730, 2000.

COLLINS, N C.; TARDIEU, F.; TUBEROSA, R. Quantitative trait loci and crop performance under abiotic stress: Where do we stand? *Plant Physiology*, 147:469-486, 2008.

COMINELLI, E.; CONTI, L.; TONELLI, C.; GALBIATI, M. Challenges and perspectives to improve crop drought and salinity tolerance. *New Biotechnology*, 30(4), 2013.

COMINELLI, E.; TONELLI, C. Transgenic crops coping with water scarcity. *New Biotechnology*, 27(5):473-477, 2010.

CRAMER, G. R.; URANO, K.; DELROT, S.; PEZZOTTI, M.; SHINOZAKI, K. Effects of abiotic stress on plants: a systems biology perspective. *BMC Plant Biology*, 11:1471-2229, 2011.

DAS M. K.; DAI H. K. A survey of DNA motif finding algorithms. *BMC Bioinformatics*, 8:1471-2105, 2007.

DAS, A.; PARIDA, S. K. Advances in biotechnological applications in three important food legumes; *Plant Biotechnology Reports*, ISSN: 1863-5466; 2013.

DAVULURI, R. V., SUN, H., PALANISWAMY, S. K., MATTHEWS, N., MOLINA, C., KURTZ, M.; GROTEWOLD, E. AGRIS: Arabidopsis gene regulatory information server, an information resource of Arabidopsis cis-regulatory elements and transcription factors. *BMC Bioinformatics*, 4:25, 2003.

DING, Z.; LI, S.; AN, X.; LIU, X.; QIN, H.; WANG, D. Transgenic expression of MYB15 confers enhanced sensitivity to abscisic acid and improved drought tolerance in *Arabidopsis thaliana*. *Journal of Genetics & Genomics*, 36:17-29, 2009.

DIVI, U. K.; RAHMAN, T.; KRISHNA, P. Brassinosteroid-mediated stress tolerance in *Arabidopsis* shows interactions with abscisic acid, ethylene and salicylic acid pathways. *BMC Plant Biology*, 10:151, 2010.

DU, H.; ZHANG, L.; LIU, L.; TANG, X. F.; YANG, W. J.; WU, Y. M.; TANG, Y. X. Biochemical and molecular characterization of plant MYB transcription factor family. *Biochemistry (Moscow)*, 74(1):1-11, 2009.

DU, Z.; ZHOU, X.; LING, Y.; ZHANG, Z.; SU, Z. agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Research*, 38(2):W64-W70, 2010.

DWEVEDI, A.; KAYASTHA A. M. Soybean: a Multifaceted Legume with Enormous Economic Capabilities, *Soybean- Biochemistry, Chemistry and Physiology*, Prof. Tzi-Bun Ng (Ed.), 978-9-53307-219-7 InTech, 165-188, 2011.

ELHITI, M.; STASOLLA, C. Structure and function of homodomain-leucine zipper (HD-Zip) proteins. *Plant Signaling and Behavior*, 4:86-88, 2009.

EMBRAPA, 2010, Histórico da soja no Brasil. Disponível em: <[http://www.cnpso.embrapa.br/index.php?op\\_page=113&cod\\_pai=35](http://www.cnpso.embrapa.br/index.php?op_page=113&cod_pai=35)> Acesso em: 31/03/2013,

EULGEM, T.; RUSHTON, P. J.; ROBATZEK, S.; SOMSSICH, I. E. The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* 5:199–206, 2000.

FIELD, B.; FISTON-LAVIER, A. S.; KEMEN, A.; GEISLER, K.; QUESNEVILLE, H.; OSBOURN, A. E. Formation of plant metabolic gene clusters within dynamic chromosomal region. *Proceedings of the National Academy of Sciences*, 108(38):16116-16121, 2011.

FINKELSTEIN, R. R.; GAMPALA, S. S.; ROCK, C. D. Abscisic acid signaling in seeds and seedlings. *Plant Cell*, 14(Suppl):S15–S45, 2002.

FUJITA, Y.; FUJITA, M.; SATOH, R.; MARUYAMA, K.; PARVEZ, M. M.; SEKI, M.; HIRATSU, K.; OHME-TAKAGI, M.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. AREB1 Is a Transcription Activator of Novel ABRE-Dependent ABA Signaling That Enhances Drought Stress Tolerance in *Arabidopsis*. *The Plant Cell*, 17:3470-3488, 2005.

FUJITA, Y.; FUJITA, M.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. ABA-mediated transcriptional regulation in response to osmotic stress in plants. *Journal of Plant Research*, 124:509-525, 2011.

FURIHATA, T.; MARUYAMA, Y.; FUJITA, Y.; UMEZAWA, T.; YOSHIDA, R.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Abscisic acid-dependent multisite phosphorylation regulates the activity of a transcription activator AREB1. *Proceedings of the National Academy of Sciences of the United States of America-PNAS*, 103(6):1988-1993, 2006.

GILMOUR, S. J.; FOWLER, S. G.; THOMASHOW, M. F.; *Arabidopsis* transcription activators CBF1, CBF2 and CBF3 have matching functional activities. *Plant Molecular Biology*, 54:767781, 2004.

GODFRAY, C. J.; BEDDINGTON, J. R.; CRUTE, I. R.; HADDAD, L.; LAWRENCE, D.; MUIR, J. F.; PRETTY, J.; ROBINSON, S.; THOMAS, S. M.;

TOULMIN, C. Food security: the challenge of feeding 9 billion people. *Science*, 327 (5967):812–818, (2010).

GOLDBERG, M. L.; Sequence analysis of *Drosophila* histone genes. Ph.D. dissertation, Stanford University, California, 1979.

GOLLDACK, D.; LU<sup>2</sup>KING, I.; YANG, O. Plant tolerance to drought and salinity: stress regulating transcription factors and their functional significance in the cellular transcriptional network. *Plant Cell Reports*, 30:1383-1391, 2011.

GRANT, D.; NELSON, R. T.; CANNON, S. B.; SHOEMAKER, R. C. SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Research*, 38:D843-D846, 2010.

GRIFFITHS, A.J.F.; *et al.* Regulation of gene transcription In: *Modern Genetic Analysis*. 3<sup>rd</sup> edition, W.H. Freeman and Company, New York, NY, 2000.

GROSS, P.; OELGESCHLÄGER, T. Core promoter-selective RNA polymerase II transcription. *Biochemistry Society Symposium.*, 2006; (73):225-236, 2006.

GUBIS, J.; VANKOVÁ, R.; CERVENÁ, V.; DRAGÚŇOVÁ, M.; HUDCOVICOVÁ, M.; LICHTNEROVÁ, H. Transformed tobacco plants with increased tolerance to drought. *South African Journal of Botany*, 73:505-11, 2007.

GULTINAN, M. J.; MARCOTTE, W. R.; QUATRANO, R. S.A plant leucine zipper protein that recognizes an abscisic acid response element. *Science*, 250:267–271, 1990.

HAAG, J. R.; PIKAARD, C. S. Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. *Molecular Cell Biology*, 12:483-492, 2011.

HAAKE, V.; COOK, D.; RIECHMANN, J. L.; PINEDA, O.; THOMASHOW, M. F.; ZHANG, J. Z. Transcription factor CBF4 is a regulator of drought adaptation in *Arabidopsis*. *Plant Physiology*, 130:639–648, 2002.

HABERER, G., MADER, M. T.; KOSAREV, P.; SPANNAGL, M.; YANG, L.; MAYER, K. F. X. Large-scale cis-element detection by analysis of correlated

expression and sequence conservation between *Arabidopsis* and *Brassica oleracea*. *Plant Physiology*, 142:1589–1602, 2006.

HALFORD, N. G.; HEY, S. J. SNF1-related protein kinases (SnRKs) act within an intricate network that links metabolic and stress signalling in plants. *Biochemical Journal*, 419:247-259, 2009.

HARTMAN, G. L., WEST, E. D., HERMAN, T. K. Crops that feed the world 2. Soybean-worldwide production, use, and constraints caused by pathogens and pests. *Food Security*, 3(1):5-17, 2011.

HERNANDEZ-GARCIA, C. M.; BOUCHARD, R. A.; RUSHTON, P. J.; JONES, M. L.; CHEN, X.; TIMKO, M. P.; FINER, J. J. High level transgenic expression of soybean (*Glycine max*) GmERF and Gmubi gene promoters isolated by a novel promoter analysis pipeline. *BMC Plant Biology*, 10(1):p237, 2010.

HERNANDEZ-GARCIA, C. M.; FINER, J. J. Identification and validation of promoters and cis-acting regulatory elements. *Plant Science*, 217:109-119, 2014.

HIGO, K.; UGAWA, Y.; IWAMOTO, M.; KORENAGA, T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research* 27:297-300, 1999.

HIMMELBACH, A.; YANG, Y.; GRILL, E. Relay and control of abscisic acid signaling. *Current Opinion In Plant Biology*, 6(5):470-479, 2003.

HIRAI, M. Y.; YANO, M.; GOODENOWE, D. B.; KANAYA, S.; KIMURA, T.; AWAZUHARA, M.; ARITA, M.; FUJIWARA, T.; SAITO, K. Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in *Arabidopsis thaliana*. *Proceedings of the National academy of Sciences of the United States of America*, 101:10205-10210, 2004.

HOBO, T.; ASADA, M.; KOWYAMA, Y.; HATTORI, T. ACGT-containing abscisic acid response element (ABRE) and coupling element 3 (CE3) are functionally equivalent. *The Plant Journal*, 19:679-689, 1999.

HU, H.; YOU, J.; FANG, Y.; ZHU, X.; QI, Z.; XIONG, L. Characterization of transcription factor gene SNAC2 conferring cold and salt tolerance in rice. *Plant Molecular Biology*, 67:169-181, 2008.

HUANG, G-T.; MA, S-L.; BAI, L-P.; ZHANG, L.; MA, H.; JIA, P.; LIU, J.; ZHONG, M.; GUO; Z-F. Signal transduction during cold, salt, and drought stresses in plants; *Molecular Biology Reports*, 39:969-987, 2012.

HUBBARD, K. E.; NISHIMURA, N.; HITOMI, K.; GETZOFF, E. D.; SCHROEDER, J. I. Early abscisic acid signal transduction mechanisms: newly discovered components and newly emerging questions. *Genes & Development*, 24:1695-1708, 2010.

JAMES, C. Global Status of Commercialized Biotech/GM Crops: ISAAA Brief. n. 43. Ithaca, NY: 2012.

JIN, X. F.; JIONG, A. S.; PENG, R. H.; LIU, J. G.; GAO, F.; CHEN, J. M.; YAO, Q. H. OsAREB1, an ABRE binding protein responding to ABA and glucose, has multiple functions in *Arabidopsis*. *BMB Reports* 43:34-39, 2010.

JOSHI, T.; PATIL, K.; FITZPATRICK, M. R.; FRANKLIN, L. D.; YAO, Q.; COOK, J. R.; XU, D. Soybean Knowledge Base (SoyKB): a web resource for soybean translational genomics. *BMC Genomics*, 13(1):15, 2012.

JUVEN-GERSHON, T.; KADONAGA, J. T. Regulation of Gene Expression via the Core Promoter and the Basal Transcriptional Machinery. *Developmental Biology*, 339(2):225–229, 2010.

KASUGA, M.; MIURA, S.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. A Combination of the *Arabidopsis* DREB1A Gene and Stress-Inducible *rd29A* Promoter Improved Drought- and Low-Temperature Stress Tolerance in Tobacco by Gene Transfer. *Plant Cell Physiology*, 45(3):346–350, 2004.

KATO, H.; XIE, G.; SATO, Y.; IMAI, R. Isolation of anther-specific gene promoters suitable for transgene expression in rice. *Plant Molecular Biology Reporter*, 28:381–387, 2010.

KEY, S.; MA, J. K. C.; DRAKE, P. M. W. Genetically modified plants and human health. *Journal of the Royal Society of Medicine*, 101:290–298, 2008.

KIM, S. Y.; KANG, J. Y.; CHO, D. I.; PARK, J. H.; KIM, S. Y. ABF2, an ABRE-binding bZIP factor, is an essential component of glucose signaling and its overexpression affects multiple stress tolerance. *The Plant Journal*, 40:75-87, 2004.

KOVALCHUK, N.; JIA, W.; EINI, O.; MORRAN, S.; PYVOVARENKO, T.; FLETCHER, S.; BRAZANOVA, N.; HARRIS, J.; BECK-OLDACH, K.; SHAVRUKOV, Y.; LANGRIDGE, P.; LOPATO, S. Optimization of TaDREB3 gene expression in transgenic barley using cold-inducible promoters. *Plant Biotechnology Journal*, 2013.

KULIK, A.; ANIELSKA-MAZUR, A.; BUCHOLC, M.; KOEN, E.; SZYMAŃSKA, K.; ŻMIENKO, A.; DOBROWOLSKA, G. SNF1-related protein kinases type 2 are involved in plant responses to cadmium stress. *Plant Physiology*, 160(2):868-883, 2012.

KULIK, A.; WAWER, I.; KRZYWINSKA, E.; BUCHOLC, M.; DOBROWOLSKA, G. SnRK2 protein kinases key regulators of plant response to abiotic stresses. *OMICS*, 15:859-872, 2011.

KULCHESKI, F. R.; DE OLIVEIRA, L. F.; MOLINA, L. G.; ALMERÃO, M. P.; RODRIGUES, F. A.; MARCOLINO, J.; BARBOSA, J. F.; STOLF-MOREIRA, R.; NEPOMUCENO, A.L.; MARCELINO-GUIMARÃES, F. C.; ABDELNOOR, R.V.; NASCIMENTO, L. C.; CARAZZOLLE, M. F.; PEREIRA, G. A. G.; MARGIS, R. Identification of novel soybean microRNAs involved in abiotic and biotic stresses. *BMC genomics*, 12(1):307-323, 2011.

KUTACH, A.K.; KADONAGA, J.T. The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters. *Molecular and Cellular Biology*, 20:4754-4764, 2000

LEMON, B.; TJIAN, R. Orchestrated response: a symphony of transcription factors for genecontrol. *Genes & Development*, 14:2551-2569, 2000.

LI, S.; FU, Q.; HUANG, W.; YU, D. Functional analysis of an Arabidopsis transcription factor WRKY25 in heat stress. *Plant Cell Reports*, 28(4):683-693, 2009.

LIAO, Y.; ZHANG, J. S.; CHEN, S. Y.; ZHANG, W. K.; Role of soybean GmbZip132 under abscisic acid and salt stresses. *Journal of Integrative Plant Biology*, 50:221-230, 2008.

LIU, Q.; KASUGA, M.; SAKUMA, Y.; ABE, H.; MIURA, S.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Two transcription factors, DREB1 and DREB2, with EREBP/AP2 DNA binding domain separate two cellular signal pathways in drought and low-temperature responsive gene expression, respectively, in Arabidopsis. *Plant Cell*. 10:1391-1406, 1998.

LIU, Y.; JI, X.; ZHENG, L.; NIE, X.; WANG, Y. Microarray Analysis of Transcriptional Responses to Abscisic Acid and Salt Stress in Arabidopsis thaliana. *International Journal of Molecular Sciences*, 14(5):9979-9998, 2013.

LIU, Y.; YIN, J.; XIAO, M.; MASON, A. S.; GAO, C., LIU, H.; FU, D. Characterization of Structure, Divergence and Regulation Patterns of Plant Promoters. *Journal of Molecular Biology Research*, 3(1):p23, 2013b.

LODISH, H.; BERK, A.; ZIPURSKY, S. L.; MATSUDAIRA, P.; BALTIMORE, D.; DARNELL, J. Regulation of transcription Initiation. In (Ed.) *Molecular Cell Biology*. 4 New York: W. H. Freeman and Company, 2000.

LORENZ, W. W.; ALBA, R.; YU, Y. S.; BORDEAUX, J.; SIMÕES, M.; DEAN, J. Microarray analysis and scale-free gene networks identify candidate regulators in drought-stressed roots of loblolly pine (*P. taeda* L.). *BMC Genomics*, 12(1):264, 2011.

MAPA – Ministério da Agricultura, Pecuária e Abastecimento. Projeções para o agronegócio Brasil 2011/2012 a 2021/2022. Disponível no site <<http://www.agricultura.gov.br>>

MARUYAMA, K.; TODAKA, D.; MIZOI, J.; YOSHIDA, T.; KIDOKORO, S.; MATSUKURA, S.; TAKASAKI, H.; SAKURAI, T.; YAMAMOTO, Y. Y.;

YOSHIWARA, K.; KOJIMA, M.; SAKAKIBARA, H.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Identification of cis-Acting Promoter Elements in Cold- and Dehydration- Induced Transcriptional Pathways in Arabidopsis, Rice, and Soybean. *DNA Research*, 19:37–49, 2012.

MASTON, G. A.; EVANS, S. K., GREEN, M. R. Transcriptional regulatory elements in the human genome. *Annual Review of Genomics and Human Genetics*, 7:29-59, 2006.

MASTROENI, P.; JANIS, C.; GRANT, A. J.; MCKINLEY, T. J.; MORGAN, F. J. E.; JOHN, V. F.; DOUGAN, G. In vivo regulation of the VI antigen in salmonella and induction of immune responses with an in vivo-inducible promoter. *Infection and Immunity*, 79(6):2481-2488, 2011.

MATSUKURA, S.; MIZOI, J.; YOSHIDA, T.; TODAKA, D.; ITO, Y.; MARUYAMA, K.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Comprehensive analysis of rice DREB2-type genes that encode transcription factors involved in the expression of abiotic stress-responsive genes. *Molecular Genetics and Genomics*, 283:185–196, 2010.

MATSUMURA, H.; YOSHIDA, K.; LUO, S.; KIMURA, E.; FUJIBE, T. High-Throughput SuperSAGE for Digital Gene Expression Analysis of Multiple Samples Using Next Generation Sequencing. *PLoS ONE* 5(8):12010, 2010. doi:10.1371/journal.pone.001, 2010.

MAZZUCOTELLI, E.; MASTRANGELO, A. M.; CROSATTI, C.; GUERRA, D.; STANCA, A. M.; CATTIVELLI, C. Abiotic stress response in plants: When post-transcriptional and posttranslational regulations control transcription. *Plant Science*, 174:420-431, 2008.

MCELROY, D.; ZHANG, W.; CAO, J.; WU, R. Isolation of an efficient actin promoter for use in rice transformation. *Plant Cell*, 2:163–171, 1990.

MIYASAKA, S.; MEDINA, J.C. *A soja no Brasil*. Ed. Miyasaka e Medina, Brasil.1981.

MOCHIDA, K.; YOSHIDA, T.; SAKURAI, T.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K.; TRAN, L. In silico analysis of transcription factor repertoire and prediction of stress responsive transcription factors in soybean. *DNA Research*, 16(41):353-369, 2009.

MOCHIDA, K.; YOSHIDA, T.; SAKURAI, T.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K.; TRAN, L. S. P. TreeTFDB: An Integrative Database of the Transcription Factors from Six Economically Important Tree Crops for Functional Predictions and Comparative and Functional Genomics. *DNA Research*, 20(2):151-162, 2013.

MORRAN, S.; EINI, O.; PYVOVARENKO, T.; PAREN, T. B.; SINGH, R.; ISMAGUL, A.; ELIBY, S.; SHIRLEY, N.; LANGRIDGE, P.; LOPATO, S. Improvement of stress tolerance of wheat and barley by modulation of expression of DREB/CBF factors. *Plant Biotechnology Journal*, 9:230–249, 2011.

MOVAHEDI, S.; SAYED TABATABAEI, B. E.; ALIZADE, H.; GHOBADI, C.; YAMCHI, A.; KHAKSAR, G. Constitutive expression of Arabidopsis DREB1B in transgenic potato enhances drought and freezing tolerance; *Biologia Plantarum*, 56(1):37-42, 2012.

NAIN, V.; SAHI, S.; KUMAR, P. A. In silico identification of regulatory elements in promoters. *Computational Biology and Applied Bioinformatics*, 2012.

NAKAMURA, M.; TSUNODA, T. OBOKATA, J. Photosynthesis nuclear genes generally lack TATA-boxes: a tobacco photosystem I gene responds to light through na initiator. *Plant Journal*, 29(1):1-10, 2002.

NAKASHIMA, K.; ITO, Y.; YAMAGUCHI-SHINOZAKI, K. Transcriptional Regulatory Networks in responsive to abiotic stresses in *Arabidopsis thaliana* and grasses. *Plant Physiology*, 149:88-95, 2009.

NAKASHIMA, K.; YAMAGUCHI-SHINOZAKI, K. Molecular Studies on Stress-Responsive Gene Expression in *Arabidopsis* and Improvement of Stress Tolerance in Crop Plants by Regulon Biotechnology. *Jarq*, 39:221-229, 2005.

NARUSAKA, Y., NAKASHIMA, K., SHINWARI, Z. K., SAKUMA, Y., FURIHATA, T., ABE H., NARUSAKA, M., SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Interaction between two cis-acting elements, ABRE and DRE, in ABA-dependent expression of Arabidopsis rd29A gene in response to dehydration and high-salinity stresses. *The Plant Journal*, 34:137-148, 2003.

NOH, S. A.; LEE, H. S.; HUH, G. H.; OH, M. J.; PAEK, K. H.; SHIN, J. S.; BAE, J. M. A sweet potato SRD1 promoter confers strong root, taproot and tuber-specific expression in Arabidopsis, carrot and potato. *Transgenic Research*, 21:265–278, 2012.

NURUZZAMAN, M.; MANIMEKALAI, R.; SHARONI, A.M.; SATOH, K.; KONDOH, H.; OOKA, H.; KIKUCHI, S. Genome-wide analysis of NAC transcription factor family in rice. *Gene*, 465:30–44, 2010.

OBATA, T.; FERNIE, A. R. The use of metabolomics to dissect plant responses to abiotic stresses. *Cellular and Molecular Life Sciences*, 69:3225-3243, 2012.

ODELL, J. T.; NAGY, F.; CHUA, N. H. Identification of DNA sequences for activity of the cauliflower mosaic virus 35S promoter. *Nature*, 313(6005):810-812, 1985.

OSOSKI, A. L.; KENNELLY, E. J. Phytoestrogens: a review of the present state of research. *Phytotherapy Research*, 17:84-86, 2003.

PANDEY, S. P.; SOMSSICH, I. E. The role of WRKY transcription factors in plant immunity. *Plant Physiology*, 150, 1648–1655, 2009.

PAYTON, P.; KOTTAPALLI, K. R.; ROWLAND, D.; FAIRCLOTH, W.; GUO, B.; BUROW, M.; PUPPALA, N.; GALLO, M. Gene expression profiling in peanut using high density oligonucleotide microarrays. *BMC Genomics*, 10:265, 2009.

PETRONI, K.; KUMIMOTO, R. W.; GNESUTTA, N.; CALVENZANI, V.; FORNARI, M.; TONELLI, C., MANTOVANI, R. The promiscuous life of plant nuclear factor  $\gamma$  transcription factors. *The Plant Cell Online*, 24(12):4777-4792, 2012.

PINHEIRO, C.; CHAVES, M. M. Photosynthesis and drought: can we make metabolic connections from available data? *Journal of Experimental Botany*, 62(3):869-882, 2011.

PINHEIRO, G. L.; MARQUES, C. S.; COSTA, M. D. B. L.; REIS, P. A. B.; ALVES, M.S.; CARVALHO, C. M.; FIETTO, L. G.; FONTES, E. P. B. Complete inventory of soybean NAC transcription factors: Sequence conservation and expression analysis uncover their distinct roles in stress response. *Gene*, 444:10-23, 2009.

POLIZEL, A.; MEDRI, M. E.; NAKASHIMA, K.; YAMANAKA, N.; FARIAS, J. R.; OLIVEIRA, M. C. N.; MARIN, S. R. R.; ABDELNOOR, R. V.; MARCELINO GUIMARAES, F. C.; FUGANTI, R.; RODRIGUES, F.A.; STOLF, R.; BENEVENTI, M. A.; ROLLA, A. A. P.; NEUMAIER, N.; YAMAGUCHI-SHINOZAKI, K.; CARVALHO, J. F. C.; NEPOMUCENO, A. L. Molecular, anatomical and physiological properties of a genetically modified soybean line transformed with rd29A:AtDREB1A for the improvement of drought tolerance. *Genetics and Molecular Research*, 10, 2011.

POPP, M. P.; DILLON, C. R.; KEISLING, T. C. Economic and weather influences on soybean planting strategies on heavy soils. *Agricultural Systems*, Essex, 76:969-984, 2003.

POTENZA, C.; ALEMAN, L.; SENGUPTA-GOPALAN, C. Targeting transgene expression. In research, agricultural, and environmental applications: Promoters used in plant transformation. *In Vitro Cellular & Developmental Biology – Plant*, 40(1):1-22, 2004.

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

QUILIS, J.; PENAS, G.; MESSEGUER, J.; BRUGIDOU, C.; SEGUNDO, B. C. The *Arabidopsis* AtNPR1 inversely modulates defense responses against fungal, bacterial, or viral pathogens while conferring hypersensitivity to abiotic stresses in transgenic rice. *Mol Plant Microbe Interact*, 21:1215-1231, 2008.

RÉ, D. A.; DEZAR, C. A.; CHAN, R. L.; BALDWIN, I. T.; BONAVENTURE, G. *Nicotiana attenuata* NaHD20 plays a role in leaf ABA accumulation during water stress, benzylacetone emission from flowers, and the timing of bolting and flower transitions. *Journal of Experimental Botany*, 62:155-166, 2011.

REINEKE, A. R.; BORNBERG-BAUER, E.; GU, J. Evolutionary divergence and limits of conserved non-coding sequence detection in plant genomes. *Nucleic Acids Research*, 1–15, 2011.

REYNOLDS, M. P.; ORTIZ, R. Adapting crops to climate change: a summary. In: Reynolds MP (eds) *Climate change and crop production*, CAB international, 1–8, 2010.

RODRIGUES, F. A.; MARCOLINO-GOMES, J.; CARVALHO, J. F. C.; NASCIMENTO, L. C.; NEUMAIER, N.; FARIAS, J. R. B.; CARAZZOLLE, M. F.; MARCELINO, F. C.; NEPOMUCENO, A. L. Subtractive libraries for prospecting differentially expressed genes in the soybean under water deficit. *Genetics and Molecular Biology*, 35(10):304-314, 2012.

RODRÍGUEZ, M.; CANALES, E.; BORRÁS-HIDALGO, O. Molecular aspects of abiotic stress in plants. *Biotechnologia Aplicada*, La Habana, 22:1-10, 2005.

ROLLA, A. A. P.; CARVALHO, J. F. C.; FUGANTI-PAGLIARINI, R.; ENGELS, C.; RIO, A.; MARIN, S. R. R.; OLIVEIRA, M. C. N.; BENEVENTI, M. A.; MARCELINO-GUIMARÃES, F. C.; FARIAS, J. R. B.; NEUMAIER, N.; NAKASHIMA, K.; YAMAGUCHI-SHINOZAKI, K.; NEPOMUCENO, A. L. Phenotyping soybean plants transformed with rd29A: AtDREB1A for drought tolerance in the greenhouse and field. *Transgenic research*, 1-13, 2013.

ROMBAUTS, S.; DÉHAIS, P.; VAN MONTAGU, M.; ROUZÉ, P. PlantCARE, a plant cis-acting regulatory element database. *Nucleic Acids Research*, 27(1):295-296, 1999.

RONDE, J. A. D.; CRESS, W. A.; KRUGERD, G. H. J.; STRASSERD, R. J.; VAN STADEN, J. Photosynthetic response of transgenic soybean plants, containing an Arabidopsis P5CR gene, during heat and drought stress. *J Plant Physiol*, 161:1211–1224, 2004.

RUSHTON, D.L.; TRIPATHI, P.; RABARA, R.C.; LIN, J.; RINGLER, P.; BOKEN, A.K.; LANGUM, T.J.; SMIDT, L.; BOOMSMA, D.D.; EMME, N.J.; et al. WRKY transcription factors: Key components in abscisic acid signalling. *Plant Biotechnol. Journal*, 10:2–11, 2012.

RUSHTON, P. J.; REINSTÄDLER, A.; LIPKA, V.; LIPPOK, B.; SOMSSICH, I. E. Synthetic plant promoters containing defined regulatory elements provide novel insights into pathogen-and wound-induced signaling. *The Plant Cell Online*, 14(4):749-762, 2002.

RUSHTON, P.J.; SOMSSICH, I.E.; RINGLER, P.; SHEN, Q.J. WRKY transcription factors. *Trends Plant Sci*, 15:247–258, 2010.

SAIBO, N. J. M.; LOURENÇO, T.; OLIVEIRA, M. M. Transcription factors and regulation of photosynthetic and related metabolism under environmental stresses. *Annals of Botany*, 103:609-623, 2009.

SAINT PIERRE, C.; CROSSA, J. L.; BONNETT, D.; YAMAGUCHI-SHINOZAKI, K.; REYNOLDS, M. P. Phenotyping transgenic wheat for drought resistance. *Journal of experimental botany*, 63(5):1799-1808, 2012.

SAKAI, T.; KOGISO, M. Soy isoflavones and immunity. *Journal of Medical Investigation*, 55:167-173, 2008.

SAKUMA, Y.; MARUYAMA, K.; QIN, F.; OSAKABE, Y.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Dual function of an Arabidopsis transcription factor DREB2A in water-stress responsive and heat-stress-responsive gene expression. *Proceedings of National Academy of Sciences USA*, 103:18822–18827, 2006.

SALEKDEH, G. H.; REYNOLDS, M.; BENNETT, J.; & BOYER, J. Conceptual framework for drought phenotyping during molecular breeding. *Trends in plant science*, 14(9):488-496, 2009.

SANTOS, A. P.; SERRA, T.; FIGUEIREDO, D. D.; BARROS, P.; LOURENÇO, T.; CHANDER, S.; OLIVEIRA, M. M.; SAIBO, N. J. Transcription regulation of

abiotic stress responses in rice: a combined action of transcription factors and epigenetic mechanisms. *OMICS*, 15:839–857, 2011.

SARKAR, A. K.; LAHIRI, A. Specificity determinants for the abscisic acid response element, *FEBS Open Bio*, 3: 101-105, 2013.

SCHMUTZ, J.; CANNON, S. B.; SCHLUETER, J.; MA, J.; MITROS, T.; NELSON, W.; HYTEN, D. L.; SONG, Q.; THELEN, J. J.; CHENG, J.; XU, D.; HELLSTEN, U.; MAY, G. D.; YU, Y.; SAKURAI, T.; UMEZAWA, T.; BHATTACHARYYA, M. K.; SANDHU, D.; VALLIYODAN, B.; LINDQUIST, E.; PETO, M.; GRANT, D.; SHU, S.; GOODSTEIN, D.; BARRY, K.; FUTRELL-GRIGGS, M.; ABERNATHY, B.; DU J.; TIAN, Z.; ZHU, L.; GILL, N.; JOSHI, T.; LIBAULT, M.; SETHURAMAN, A.; ZHANG, X. C.; SHINOZAKI, K.; NGUYEN, H. T.; WING, R. A.; CREGAN, P.; SPECHT, J.; GRIMWOOD, J.; ROKHSAR, D.; STACEY, G.; SHOEMAKER, R. C.; JACKSON, S. A. Genome sequence of the palaeopolyploid soybean. *Nature*, 463:178-83, 2010.

SCHRAMM, L.; HERNANDEZ, N. Recruitment of RNA polymerase III to its target promoters. *Genes & Development*, 16:2593-2620, 2002.

SHAIK, R.; RAMAKRISHNAM W. Bioinformatic analysis of epigenetic and micro-RNA mediated regulation of drought responsive genes in rice. *PLoS ONE* 7(11):e49331, 2012.

SHEN, Y. G.; DU, B. X.; ZHANG, W. K.; ZHANG, J. S.; CHEN, S. Y. AhCMO, regulated by stresses in *Atriplex hortensis*, can improve drought tolerance in transgenic tobacco. *Theor Appl Genet*, 105:815–21, 2002.

SHIN, D.; MOON, S-J.; HAN, S.; KIM, B-G.; PARK, S. R.; LEE, S-K.; YOON, H-J.; LEE, H-E.; KWON, H-B.; BAEK, D.; YI, B. Y.; BYUN M-O. Expression of StMYB1R-1, a novel potato single MYB-like domain transcription factor, increases drought tolerance. *Plant Physiology*, 155:421-432, 2011.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Gene networks involved in drought stress response and tolerance. *Journal of Experimental Botany*, 58(2): 221-227, 2007.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Curr Opin Plant Biol*, 3:217-223, 2000.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Molecular responses to drought and cold stress. *Curr Opin Biotechnol*, 7:161-167, 1996.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K.; SEKI, M. Regulatory network of gene expression in the drought and cold stress responses. *Curr Opin. Plant Biol*, 6: 410-17, 2003.

SHULAEV, V. et al. Metabolomics for plant stress response. *Physiol. Plant*, 132:199–208, 2008.

SHULTZ, J. L.; KURUNAM, D.; SHOPINSKI, K.; IQBAL, M. J.; KAZI, S.; ZOBRIK, K.; BASHIR, R.; YAEGASHI, S.; LAVU, N.; AFZAL, A. J.; YESUDAS, C. R.; KASSEM, M. A.; WU, C.; ZHANG, H. B.; TOWN, C. D.; MEKSEM, K.; LIGHTFOOT, D. A. The Soybean Genome Database (SoyGD): a browser for display of duplicated, polyploid, regions and sequence tagged sites on the integrated physical and genetic maps of *Glycine max*. *Nucleic Acids Res*, 34(Database issue):D758-D765, 2006.

SHURTLEFF, W.; AOYAGI, A. History Of Edamame, Green Vegetable Soybeans And Vegetable-Type Soybeans (1275 - 2009): Extensively annotated bibliography and sourcebook. Soyinfo Center, USA, 2009.

SINGH, A. K.; SOPORY, S. K.; WU, R.; SINGLA-PAREEK, S. L. Transgenic Approaches. In: *Abiotic Stress Adaptation in Plants: Physiological, Molecular and Genomic Foundation*, A Pareek, SK Sopory, HJ Bohnert and Govindjee (eds.), 417-450, 2010.

SINGH, O. V.; GHAI, S.; PAUL, D.; JAIN, R. K. Genetically modified crops: Success safety assessment and public concern. *Applied Microbiology and Biotechnology*, 71(5):598-607, 2006.

SMALE, S. T.; KADONAGA, J. T. THE RNA POLYMERASE II CORE PROMOTER *Annual Review of Biochemistry*, 72: 449-479, 2003.

SON, O.; HUR, Y. S.; KIM, Y. K.; LEE, H. J.; KIM, S.; KIM, M. R.; NAM, K. H.; LEE, M. S.; KIM, B.-Y.; PARK, J.; LEE, S. C.; HANADA, A.; YAMAGUCHI, S.; LEE, I.-J.; KIM, S. K.; YUN, D. J.; SÖDERMAN, E.; CHEON, C.I. ATHB12, an ABA-inducible homeodomain-leucine zipper (HD-Zip) protein of Arabidopsis, Negatively regulates the growth of the inflorescence stem by decreasing the expression of a Gibberellin 20-oxidase gene. *Plant and Cell Physiology*, 51:1537-1547, 2010.

SONG, Y.; JI, D.; LI, S.; WANG, P.; LI, Q.; XIANG, F. The dynamic changes of DNA methylation and histone modifications of salt responsive transcription factor genes in soybean. *PLoS ONE* 7(7):e41274, 2012.

TANAKA, T.; KOYANAGI, K. O.; ITOH, T. Highly diversified molecular evolution of downstream transcription start sites in rice and Arabidopsis. *Plant Physiology*, 149(3):1316-1324, 2009.

TITTARELLI, A.; SANTIAGO, M.; MORALES, A.; MEISEL, L.A.; SILVA, H. Isolation and functional characterization of cold-regulated promoters, by digitally identifying peach fruit cold-induced genes from a large EST dataset. *BMC Plant Biololy*, 9:121, 2009.

TRAN, L. S.; NAKASHIMA, K.; SAKUMA, Y.; OSAKABE, Y.; QIN, F.; SIMPSON, S. D.; MARUYAMA, K.; FUJITA, Y.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Co-expression of the stress inducible zinc finger homeodomain ZFHD1 and NAC transcription factors enhances expression of the ERD1 gene in *Arabidopsis*. *The Plant Journal*, 49:46-63, 2007.

TRAN, L. S. P.; NAKASHIMA, K.; SAKUMA, Y.; SIMPSON, S. D.; FUJITA, Y.; MARUYAMA, K.; FUJITA, M.; SEKI, M.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Isolation and functional analysis of Arabidopsis stress-inducible NAC transcription factors that bind to a droughtresponsive cis-element in the early responsive to dehydration stress 1 promoter. *The Plant Cell*, 16:2481-2498, 2004.

TWYMAN, R. M. Growth and development: Control of gene expression, regulation of transcription. Encyclopedia of Applied Plant Sciences. London UK: Elsevier Science, 558-567, 2003.

UDVARDI, M. K.; KAKAR, K.; WANDREY, M.; MONTANRI, O.; MURRAY, J.; ANDRAIANKAJA, A.; ZHANG, J-Y.; BENEDITO, V.; HOFER, J. M. I.; CHENG, F.; TOWN, C. D. Legume transcription factors: global regulators of plant development and response to the environment. *Plant Physiology*, 144: 538-549. 2007.

UMEZAWA, T.; FUJITA, M.; FUJITA, Y.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future. *Current Opinion in Biotechnology* 17:113-122, 2006.

USADEL, B.; NAGEL, A.; THIMM, O.; REDESTIG, H.; BLAESING, O. E.; PALACIOS-ROJAS, N.; SELBIG, J.; HANNEMANN, J.; PIQUES, M. C.; STEINHAUSER, D.; SCHEIBLE, W. R.; GIBON, Y.; MORCUENDE, R.; WEICHT, D.; MEYER, S.; STITT, M. Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. *Plant Physiology*, 138:1195-204, 2005.

U.S. Department of Agriculture, National Agricultural Statistics Service, 2012  
<http://www.nass.usda.gov/Statistics>

VENU, R. C.; SREEREKHA, M. V.; MADHAV, M. S.; NOBUTA, K.; MOHAN, K. M.; CHEN, S.;...WANG, G. L. Deep transcriptome sequencing reveals the expression of key functional and regulatory genes involved in the abiotic stress signaling pathways in rice. *Journal of Plant Biology*, 56(4):216-231, 2013.

VISARADA, K. B. R. S.; MEENA, K.; ARUNA, C.; SRUJANA, S.; SAIKISHORE, N.; SEETHARAMA, N. Transgenic breeding: Perspectives and prospects. *Crop Science*, 49(5):1555-1563, 2009.

WANG, L.; LI, L.; XU, L.; ZHOU, J.; ZHUANG, H.; GONG, X.; WANG, M.; SUN, S.S.M.; ZHUGE, Q. Isolation and functional analysis of the Poplar RbcS gene promoter. *Plant Molecular Biology Reporter*, 31:120–127, 2013.

WANG, Q.; WAN, L.; LI, D.; ZHU, L.; QIAN, M.; DENG, M. Searching for bidirectional promoters in *Arabidopsis thaliana*. BMC Bioinformatics, 10(1):S29, 2009.

WANG, W. X.; VINOCUR, B.; SHOSEYOV, O.; ALTMAN, A. Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. Planta, 218:1-14, 2003.

WANG, X.; HABERER, G.; MAYER, K. Discovery of cis-elements between sorghum and rice using co-expression and evolutionary conservation. BMC Genomics, 10:1471-2164, 2009.

WANG, Z.; LIBAULT, M.; JOSHI, T.; VALLIYODAN, B.; NGUYEN, H., XU, D. Stacey G, Cheng J. SoyDB: a knowledge database of soybean transcription factors. BMC Plant Biology, 10:14, 2010.

WECKWERTH W. Green systems biology - From single genomes, proteomes and metabolomes to ecosystems research and biotechnology. Journal Proteomics, 75:284-305, 2011.

WEI, W.; PELECHANO, V.; JARVELIN, A. I.; STEINMETZ, L. M. Functional consequences of bidirectional promoters. Trends in Genetics, 27(7):267-276, 2011.

WELLMER, F.; RIECHMANN J. L.; Gene network analysis in plant development by genomic technologies. International Journal of Developmental Biology, 49:745-759, 2005.

WOLF, E.; AIGNER, B.; KLYMIUK, N. Transgenic pigs for xenotransplantation: selection of promoter sequences for reliable transgene expression. Current Opinion in Organ Transplantation, 15(2):201-206, 2010.

WU, X.; SHIROTO, Y.; KISHITANI, S.; ITO, Y.; TORIYAMA, K. Enhanced heat and drought tolerance in transgenic rice seedlings overexpressing OsWRKY11 under the control of HSP101 promoter. Plant Cell Reports, 28(1):21-30, 2009.

XIANG, Y.; TANG, N.; DU, H.; YE, H.; XIONG, L. Characterization of OsbZIP23 as a key player of the basic leucine zipper transcription factor family for conferring abscisic acid sensitivity and salinity and drought tolerance in rice. *Plant Physiology*, 148:1938–1952, 2008.

XIONG, L.; LEE, H.; ISHITANI, M.; ZHU, J. K. Regulation of osmotic stress responsive gene expression by LOS6/ABA1 locus in *Arabidopsis*. *Journal of Biological Chemistry*, 277:8588-8596, 2002.

XU, D.Q.; HUANG, J.; GUO, S.Q.; YANG, X.; BAO, Y.M.; TANG, H.J.; ZHANG, H.S. Overexpression of a TFIIIA-type zinc finger protein gene ZFP252 enhances drought and salt tolerance in rice (*Oryza sativa* L.). *FEBS Letters*, 582:1037–1043, 2009.

XU, L.; YE, R.; ZHENG, Y.; WANG, Z.; ZHOU, P.; LIN, Y.; LI, D. Isolation of the endosperm-specific LPAAT gene promoter from coconut (*Cocosnucifera* L.) and its functional analysis in transgenic rice plants. *Plant Cell Reports*, 29:1061–1068, 2010.

XU, Z. S.; CHEN, M.; LI, L. C.; MA, Y. Z. Functions of the ERF transcription factor family in plants. *Botany*, 865:969–977, 2008.

YAMADA, M.; MORISHITA, H.; URANO, K.; SHINOZAKI, N.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Effects of free proline accumulation in petunias under drought stress. *Journal of Experimental Botany*, 56:1975–1981, 2005.

YAMAMOTO, Y. Y.; YOSHIOKA, Y.; HYAKUMACHI, M., MARUYAMA, K.; YAMAGUCHI-SHINOZAKI, K.; TOKIZAWA, M.; KOYAMA, H. Prediction of transcriptional regulatory elements for plant hormone responses based on microarray data. *BMC Plant Biology*, 11:39, 2011.

YI, N.; KIM, Y. S.; JEONG, M. H.; OH, S.J.; JEONG, J. S.; PARK, S. H.; JUNG, H.; CHOI, Y.D.; KIM, J. K. Functional analysis of six drought-inducible promoters in transgenic rice plants throughout all stages of plant growth. *Planta*, 232:743–754, 2010.

YOSHIDA, T.; FUJITA, Y.; SAYAMA, H.; KIDOKORO, S.; MARUYAMA, K.; MIZOI, J.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. AREB1, AREB2, and ABF3 are master transcription factors that cooperatively regulate ABRE-dependent ABA signaling involved in drought stress tolerance and require ABA for full activation. *The Plant Journal*, 61:672-685, 2010.

ZHANG, C.; WANG, J.; HUA, X.; FANG, J.; ZHU, H.; GAO, X. A mutation degree model for the identification of transcriptional regulatory elements. *BMC Bioinformatics*, 12:262, 2011.

ZHANG, G. H.; SU, Q.; AN, L.J.; WU, S. Characterization and expression of a vacuolar Na<sup>+</sup>/H<sup>+</sup> antiporter gene from the monocot halophyte *Aeluropus litoralis*. *Plant Physiology and Biochemistry*, 46:117–26, 2008.

ZHANG, L.; WANG, Y.; ZHANG, X.; ZHANG, M.; HAN, D.; QIU, C.; HAN, Z. Dynamics of phytohormone and DNA methylation patterns changes during dormancy induction in strawberry (*Fragaria 9 ananassa* Duch.). *Plant Cell Reports* 31:155–165, 2012.

ZHANG, Y.; WANG, L. The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evolutionary. Biology*, 5(1):1, 2005.

ZHOU, Z.; LIN, I. J.; DARST, R. P.; & BUNGERT, J. Maneuver at the transcription start site: Mot1p and NC2 navigate TFIID/TBP to specific core promoter elements. *Epigenetics*, 4(1):1-4, 2009.

ZHU, J. K. Salt and drought stress signal transduction in plants. *Annual Review of Plant Biology*, 53:247-273, 2002.

## CAPÍTULO 2

### 5 Artigo

#### **Identificação de regiões promotoras de genes expressos em soja sob déficit hídrico**

Artigo a ser submetido à revista BMC Plant Biology

## 5.1. Resumo

Estudos de promotores induzidos responsáveis pela modulação gênica em condições específicas podem melhorar a compreensão sobre a regulação de genes e expandir as opções de promotores para uso em engenharia genética vegetal. Os elementos presentes no *core promoter*, elementos proximais e distais estão envolvidos no controle transcricional dos muitos genes relacionados com a complexa rede de resposta das plantas ao estresse. Os avanços nas tecnologias de transcriptomas têm permitido análises em larga escala de transcritos, e estudo dos componentes de suas regiões promotoras para confirmação de sua funcionalidade como sequências reguladoras. Neste estudo foram analisadas regiões promotoras de genes diferencialmente expressos em plantas de soja sob condições de déficit hídrico. A cultivar brasileira de soja BR 16 foi avaliada sob duas condições de déficit hídrico: progressivo (oito dias) e de curto período (25 a 150 min). Os transcritos identificados nos dois tratamentos foram comparados e 134 transcritos diferencialmente expressos e com mesma regulação foram identificados em comum. A análise de 1kb *upstream* da região promotora desses transcritos revelou a presença de 85 *cis*-elementos comumente encontrados em regiões promotoras de plantas, sendo 35 deles presentes em pelo menos 50% dos TDEs, incluindo TATA Box, CCAAT Box e GC Box e INR necessários para o acoplamento da maquinaria de transcrição e *cis* elementos especificamente induzidos por fatores abióticos tais como ABRE, DREB, HEAT e SALT. Todos os modelos de promotores apresentaram combinação binária dos *cis* elementos GTBX e AHBP podendo haver uma relação entre esse modelo e a regulação transcricional em condição de déficit hídrico.

## 5.2. Introdução

A cultura da soja ocupa um lugar de destaque na economia brasileira e mundial, com perspectivas de expansão, devido a maior demanda alimentar em razão do crescimento da população global e as novas possibilidades de utilização deste grão. Entretanto, diversos fatores bióticos e abióticos podem interferir diretamente na produtividade, provocando prejuízos financeiros significativos.

A seca é um dos mais importantes fatores que afetam a regulação do crescimento e desenvolvimento das plantas e uma das principais causas para a redução do rendimento das culturas. Fatores como a duração, a severidade e a fase de desenvolvimento da planta em que o estresse é imposto, influenciam no padrão de resposta e na dimensão das perdas de produtividade que ocorrem (Pinheiro e Chaves, 2011, Bertolli et al., 2012; Saint Pierre et al., 2012).

A complexa rede de respostas das plantas ao estresse envolve o controle transcricional de muitos genes (Venu et al., 2013). A regulação da expressão gênica é o mecanismo utilizado pelas células para modular a resposta aos sinais externos, sendo essa regulação feita principalmente pela região promotora do gene, fundamental para iniciar ou interromper sua expressão (Wolf et al., 2010; Mastroeni et al., 2011; Zhang et al., 2011).

As vias de transdução de sinal podem regular a expressão de genes responsivos à seca através de diversos fatores de transcrição (FTs) que se ligam a *cis*-elementos específicos na região promotora dos genes alvo (Balderas-Hernandez et al., 2013). Os *cis*-elementos regulatórios (CERs) de forma independente ou em conjunto promovem a ativação ou repressão desses genes específicos e desencadeiam mecanismos fisiológicos, bioquímicos e moleculares de resposta permitindo que a planta sobreviva ou se adapte (Zou et al., 2011).

Diferentes grupos de CERs presentes em regiões promotoras e fatores *trans* são conhecidos por estarem envolvidos na transcrição de genes responsivos a estresse abióticos, sendo alguns deles dependentes de ácido abscísico (ABA-dependentes) e, outros, independentes (ABA-independentes)

(Shinozaki e Yamaguchi-Shinozaki, 2000, 2007; Zhu, 2002; Nakashima e Yamaguchi-Shinozaki, 2005, Umezawa et al., 2006; Yamamoto et al., 2011 ).

Pelo seu papel insubstituível na modulação da expressão espaço-temporal de genes, promotores de plantas têm atraído cada vez mais atenção (Liu et al., 2013), e representam ferramentas potenciais para utilização na geração de plantas geneticamente modificadas (PGMs) com características específicas e melhor adaptadas às mais diversas condições ambientais.

Nos últimos anos muitos trabalhos têm sido desenvolvidos no sentido de se isolar e avaliar promotores induzidos específicos (Xu et al., 2010; Tittarelli et al., 2009; Kato et al., 2010; Yi et al., 2010; Hernandez-Garcia et al., 2010; Noh et al., 2012; Wang et al., 2013). A identificação de promotores que regulam a expressão diferencial de genes no momento, intensidade e tecido de interesse representa um dos grandes objetivos da engenharia genética de plantas na agricultura moderna, uma vez que genes sob o controle de promotores constitutivos são ativos na maioria das células durante todo o ciclo celular em vários níveis (Twyman et al., 2003), e essa característica pode se tornar uma limitação dependendo da estratégia de engenharia genética almejada.

### 5.3. Material e Métodos

#### 5.3.1. Experimento de déficit hídrico progressivo

##### 5.3.1.1 Obtenção do material biológico

A cultivar de soja convencional BR 16 sensível à seca (Oya et al., 2004) foi cultivada em potes de plástico contendo 220 g de substrato composto por uma mistura de terra:vermiculita (1:1) esterilizada por autoclavagem. Os potes foram mantidos em caixa plástica contendo nove litros de água, de maneira que a absorção da água ocorreu por capilaridade (Figura 5).

As plantas permaneceram sob condições de 12h de luz (28°C) sob  $50 \pm 10 \mu\text{mol photons/m}^2/\text{s}$  e 12h de escuro (25°C) até atingirem o estágio de desenvolvimento  $V_1$ - $V_2$  (aproximadamente 14 dias). Após esse período a umidade do solo foi medida e oito plantas foram então transferidas para uma caixa plástica sem água, para aplicação do déficit hídrico. Outras oito plantas foram mantidas na caixa com água (condição controle). Transcorridos oito dias (tempo de exposição ao déficit hídrico = T8) foram realizadas novas medições da umidade do solo e a coleta da parte aérea de quatro plantas submetidas ao déficit hídrico e de quatro plantas da condição controle. Todo o material coletado foi imediatamente congelado em nitrogênio líquido e armazenado em freezer -80 °C até o momento da realização das análises moleculares.



**Figura 5** Sistema utilizado para o cultivo das plantas de soja cultivar BR 16 e aplicação do déficit hídrico progressivo.

### 5.3.1.2. Extração e Purificação de RNA Total

A extração do RNA total foi realizada utilizando-se o reagente Trizol<sup>®</sup> (Invitrogen), com algumas modificações nas condições sugeridas pelo fabricante. Aproximadamente 200 mg de tecido foliar foram pulverizados em nitrogênio líquido e adicionados 1mL de Trizol<sup>®</sup> aquecido, a 55°C. A solução foi homogeneizada por 5 min em agitador e posteriormente centrifugada por 10 min, a 15000 x g, a 4°C. O sobrenadante foi transferido para novo tubo e 400 µL de clorofórmio:isoamilico (24:1) foram adicionados, seguido de homogeneização e centrifugação nas mesmas condições anteriores. Essa etapa foi repetida duas vezes.

O RNA foi precipitado com 250 µL da solução *Salt High-Solution* (0,8M citrato de sódio e 1,2M cloreto de sódio) e 250 µL de 2-Propanol. Após homogeneização por 10 min em agitador e centrifugação por 10 min, a 15000 x g, a 4°C, o sobrenadante foi descartado e o *pellet* lavado com 1mL de etanol 75%, seguido de nova centrifugação a 15000 x g por 15 min, a 4°C. Novamente o sobrenadante foi descartado e o *pellet* ressuspendido em 100 µL de água livre de RNase. Uma nova precipitação do RNA foi feita com a adição de 10 µL de Acetato de sódio 3M e 250 µL de Etanol 100%. Após homogeneização por 5 min em agitador e centrifugação por 15 min, a 15000 x g, a 4°C, o sobrenadante foi descartado e o *pellet* lavado com 1 mL de etanol 75%, seguido de mais uma centrifugação a 15000 x g por 15 min, a 4°C. Finalmente o sobrenadante foi descartado e o *pellet* ressuspendido em 30 µL de água livre de RNase. O RNA obtido foi quantificado por espectrofotometria em Nanodrop P100.

As amostras de RNA foram tratadas com DNaseI (Invitrogen) para eliminação de DNA, empregando metade dos volumes e tempos sugeridos pelo fabricante. Após o tratamento, o RNA foi purificado. O volume de cada amostra foi ajustado para 100 µL com água livre de RNase, e adicionados 100 µL de fenol:clorofórmio:isoamil (25:24:1). A solução foi homogeneizada e centrifugada a 12000 x g, por 5 min, a temperatura ambiente. O sobrenadante foi transferido para novo tubo e novamente clorofórmio foi adicionado, seguindo-se a homogeneização e centrifugação. O sobrenadante foi mais uma vez transferido para novo tubo e adicionados 10 µL de Acetato de Sódio 3M e 250 µL de

etanol 100% gelado. A solução foi mantida a  $-80^{\circ}\text{C}$ , por 20 min, seguindo-se uma centrifugação a  $12000 \times g$ , por 10 min, a  $4^{\circ}\text{C}$ . O sobrenadante foi então removido, o *pellet* lavado com etanol 75% gelado e centrifugado a  $12000 \times g$ , por 5 min, a  $4^{\circ}\text{C}$ . Novamente, o sobrenadante foi descartado e o *pellet* ressuspendido em 20  $\mu\text{L}$  de água livre de RNase.

### 5.3.1.3. Síntese de cDNA

A síntese de cDNA foi realizada utilizando-se o kit SuperScript III (Invitrogen) seguindo as recomendações do fabricante. Após a síntese de cDNA a ausência de contaminação com DNA genômico foi verificada pela amplificação do cDNA via PCR convencional, com *primers* de actina de soja (Forward 5' CCC CTC AAC CCA AAG GTC AAC AG 3' e Reverse 5' GGA ATC TCT CTG CCC CAA TTG TG 3'). Quando na presença de DNA genômico, a reação produz um fragmento de 520 pb, entretanto na presença de cDNA observa-se um fragmento de 440 pb devido à remoção de um íntron de 80 pb.

A reação de PCR foi realizada em volume final de 25  $\mu\text{L}$ , sendo o *mix* composto por 60 ng de cDNA, 5 $\mu\text{M}$  de cada *primer* F e R, 2,5mM dNTP's (Invitrogen), 2mM cloreto de magnésio ( $\text{MgCl}_2$ ) (Sigma), 2,5  $\mu\text{L}$  de tampão de reação 10X, e 1U de *Taq* DNA polimerase (Invitrogen). As amplificações foram realizadas em termociclador Veriti (Life Tech), com programação composta por um período inicial de desnaturação a  $94^{\circ}\text{C}$  por 1 min, seguido por 35 ciclos de  $94^{\circ}\text{C}$  por 30 s,  $55^{\circ}\text{C}$  por 30 s e  $72^{\circ}\text{C}$  por 30 s finalizando com um período de extensão a  $72^{\circ}\text{C}$ , por 5 min. O produto da amplificação foi analisado por eletroforese em gel de agarose 1,3%, corado com brometo de etídeo e visualizado sob luz ultravioleta.

### 5.3.1.4. Análise via PCR em tempo real

Para validação do estresse de déficit hídrico aplicado nas plantas da cultivar de soja BR 16, foram realizadas análises de RTqPCR e *Northern Blot*.

As reações de PCR quantitativo foram conduzidas em equipamento ABI PRISM<sup>®</sup> 7300 Sequence Detection System (Applied Biosystems), utilizando o *kit* Platinum<sup>®</sup> SYBR Green<sup>®</sup> qPCR SuperMix-UDG with ROX (Invitrogen),

conforme as instruções do fabricante. Inicialmente, foi determinada a eficiência de amplificação de três conjuntos de *primers* para o gene alvo *LEA* e dois conjuntos para os genes normalizadores (*GmB-actina* e *18S*) (Tabela 01). O conjunto de *primers* com melhor relação entre as eficiências (*Gm09g31740* e *Gmβ-Actina*) foi utilizado para validação do estresse.

**Tabela 01:** Conjuntos de *primers* testados para amplificação do gene alvo *LEA* e para os genes normalizadores (*GmB-actina* e *18S*). O conjunto de *primers* (alvo e normalizador) com melhor relação entre as eficiências foi utilizado para validação do déficit hídrico progressivo.

<b>Primers Normalizadores</b>	<b>Sequência 5' – 3'</b>	<b>Eficiências (%)</b>	
GmRNAr18S-F	AAACGGCTACCACATCCAAGG	90,8	
GmRNAr18S-R	CCTTCAATGGATCCATCGTTA		
Gmβ-Actina-F	GAGCTATGAATTGCCTGATGG	96,5	
Gmβ-Actina-R	CGTTTCATGAATTCCAGTAGC		

<b>Primer Alvos</b>	<b>Sequência 5' – 3'</b>	<b>Acesso GenBank</b>	<b>Eficiências (%)</b>
Glyma03g30044.1-F	CTGCAAACATTGGTGCTTCT	AK244314	54,0
Glyma03g30044.1-R	GCATAGCTGACGTCTGATGG		
Glyma13g17980.1-F	AGGGGCTGCTGAATATGCTA	U02966	62,7
Glyma13g17980.1-R	GTCACCTGCCTTCTCTGCTG		
Glyma09g31740.1-F	CTGCTGCTGGTGGTTACAGT	AJ704825	99,8
Glyma09g31740.1-R	CCATGTTGCCTACCAGTGTC		

### 5.3.1.5. Análise da expressão gênica via *Northern Blot*

O *Northern Blotting* foi realizado utilizando-se o *kit* DIG DNA Labeling and Detection (Roche), que se baseia na detecção de hibridização por imunoenensaio enzimático de *primer* de DNA marcado com digoxigenin-dUTP.

As amostras contendo 5 µg de RNA, e 23,3 µL de tampão Northern Loading Buffer e brometo de etídeo foram mantidas sob agitação por 1 min, aquecidas a 65°C por 15 min, e transferidas para gelo por 10 min, e então aplicadas em gel de agarose (1,1% em tampão contendo 0,54% de acetato de sódio, 2,0% de MOPS, 5mM de EDTA e 9% de formamida).

A eletroforese foi realizada por duas horas e vinte minutos a 80 V e 400 mA, mantendo-se o tampão de corrida sob agitação. Após a eletroforese o RNA foi transferido por capilaridade para membrana durante 20 h, utilizando-se um aparato de transferência contendo a solução 10X SSC (1,5M de Cloreto de sódio e 0,15M de citrato de sódio).

A sonda foi preparada por PCR com primer Glyma09g31740 (*LEA 8*) utilizando o *kit* DIG DNA Labeling and Detection (Roche), seguindo as orientações do fabricante. O produto de amplificação foi purificado utilizando o *kit* SV Gel and PCR CleanUp Wizard (Promega), conforme recomendação do fabricante. A membrana foi pré-hibridizada em tampão (0,5M fosfato pH 7,0, 7% SDS, 10mM EDTA pH 8,0) a 65°C, por 30 min, em forno de hibridização. Em seguida aproximadamente 25 ng da sonda desnaturada foi adicionada ao tubo de hibridização, mantido a 65°C por aproximadamente 20 h.

A membrana foi submetida a duas lavagens seguidas com tampão 2X (2X SSC e 0,1% SDS), a 65°C, com rotação por 10 min, seguindo-se mais duas lavagens com tampão 0,1X (0,1X SSC e 0,1% SDS), a 65°C, com rotação por 15 min, e uma lavagem com tampão TBST 1X por 5 min, a temperatura ambiente. O bloqueio da membrana e o preparo para detecção da fluorescência foi feito de acordo com a recomendação do fabricante do *kit* DIG DNA Labeling and Detection (Roche). A fluorescência foi detectada em equipamento LAS 3000 Fuji.

#### **5.3.1.6. Microarranjos de DNA**

A qualidade das amostras de RNA foi determinada pela avaliação em Bioanalyzer (Agilent Technologies). O preparo das amostras para análise de microarranjo de DNA foi feito utilizando o *kit* Two-Color Microarray-Based Gene Expression Analysis (Agilent Technologies), seguindo as recomendações do fabricante. Foi utilizado chip G2534 no formato 4x 44K (Agilent Technologies) previamente desenhado por Maruyama e colaboradores (2012) contendo 43.640 genes de soja codificadores de proteínas. Após a hibridização das lâminas, a imagem foi capturada pelo Scanner G2505C (Programa de controle de scan, Versão A.8.5.1; Agilent Technologies Inc) e transformada em valores

numéricos com o programa Feature Extractions (Versão 10.10.1.1; Agilent Technologies).

Os dados foram analisados pelo programa GeneSpring GX (versão 11.5.1; Agilent Technologies), considerando as réplicas técnicas e biológicas, o qual realiza a normalização usando o método Lowess (*Locally Weighted Scatterplot Smoothing*) (Quackenbush, 2002) e calcula as razões de expressão gênica corrigidas pelo método de Benjamini e Hochberg (Benjamini e Hochberg, 1995) com taxa de falsa descoberta FDR (*False Discovery Rate*) a 1% de significância.

### 5.3.2. Experimento de déficit hídrico de curto período

#### 5.3.2.1. Cultivo das plantas e tratamento experimental

As plantas de soja da cultivar BR 16 foram cultivadas em sistema hidropônico com solução nutritiva de Hoagland (Hoagland, 1950) como descrito por Rodrigues *et al.* (2012). O experimento foi conduzido em casa-de-vegetação, disposto em três blocos (contendo 10 plantas em cada) totalmente casualizados, sob temperatura de  $30 \pm 4^\circ\text{C}$  e umidade relativa de  $68 \pm 8\%$ .

As plantas foram mantidas até o estágio vegetativo V<sub>4</sub> (Fehr, 1971), quando foram submetidas ao déficit hídrico pela remoção das mesmas da solução nutritiva e exposição ao ar durante zero minuto (plantas controle), 25, 50, 75, 100, 125 e 150 minutos. O tecido foliar foi imediatamente coletado após cada período de tempo, congelado em nitrogênio líquido e armazenado a  $-80^\circ\text{C}$ .



**Figura 6** **A)** Sistema hidropônico utilizado para o cultivo das plantas de soja cultivar BR 16 e **B)** aplicação do déficit hídrico de curto período.

### 5.3.2.2. Construção das bibliotecas e sequenciamento

O RNA total de tecido foliar foi extraído utilizando-se o reagente Trizol<sup>®</sup> (Invitrogen) seguindo as recomendações do fabricante. Após avaliar a qualidade e integridade do material, quantidades equimolares de RNA total purificado das réplicas biológicas foram misturadas para produzir uma amostra única de plantas provenientes de dois períodos de exposição ao estresse. Assim, foram combinadas as seguintes amostras para a síntese das bibliotecas: 0 min; 25-50 min; 75-100 min; 125-150 min sob condições de déficit hídrico. As amostras de RNA total foram enviadas para sequenciamento na plataforma Illumina pela empresa Fasteris SA (Suíça), realizado a partir do kit True-Seq (Illumina). As bibliotecas foram sequenciadas em dois canais no sequenciador HiSeq 2000 (Illumina) usando 1x100 bp *single reads*.

### 5.3.2.3. Alinhamento dos *reads* contra os transcritos de soja

Os *reads* obtidos a partir de cada biblioteca foram alinhados contra os 73.321 transcritos de soja (*Gmax v1.1*) disponíveis no banco de dados Phytozome (<ftp://ftp.jgi-psf.org/pub/compngen/phytozome/v9.0/Gmax/>). Os transcritos foram previamente filtrados, optando-se pela maior variante de *splicing* (46.367 transcritos). O programa SOAP foi utilizado para o mapeamento com no máximo 2 *mismatches* permitidos.

O cálculo de expressão de cada transcrito foi baseado no número de *reads* que alinhou em cada transcrito, de acordo com o método descrito por Mortazavi *et al.* (2008). Para esse cálculo foram considerados somente os *reads* que apresentaram alinhamento único contra os transcritos de soja, conforme a fórmula abaixo:

$$FPKM = (N * 10^9) / S * T$$

onde,

FPKM = Fragmentos Por Kilobase de exon por Milhão de *reads* mapeados

N = Número de *reads* (fragmentos) mapeados em um transcrito

S = tamanho do transcrito

T = Total de *reads* mapeados da biblioteca

#### **5.3.2.4. Identificação dos transcritos diferencialmente expressos (TDEs)**

A identificação de TDEs em cada biblioteca sequenciada foi baseada no número de *reads* alinhados em cada transcrito de soja. Foram considerados somente os *reads* com alinhamento único contra os transcritos. Dois pacotes estatísticos foram utilizados nesta etapa: edgeR (Robinson *et al.*, 2010) e baySeq (Hardcastle *et al.*, 2010). Somente transcritos que apresentaram a taxa de confiança (ou FDR, False Discovery Rate)  $\leq 0,01$  por ambos os métodos, e, com *fold-change* de FPKM  $\leq -2$  ou  $\geq 2$  foram determinados como diferencialmente expressos na condição avaliada.

#### **5.3.3. Classificação funcional dos transcritos diferencialmente expressos**

Os transcritos identificados como diferencialmente expressos no déficit hídrico progressivo e de curto período foram analisados utilizando o software MapMan (Usadel *et al.* 2005), que organiza os transcritos em categorias funcionais de acordo com o termo GO, obtido a partir do banco de dados Gene Ontology (Ashburner *et al.*, 2000).

#### **5.3.4. Seleção dos transcritos comuns**

Para seleção dos TDEs identificados comuns a ambos os déficits hídricos aplicados, uma comparação foi realizada utilizando-se o software Venny (Oliveiros, 2007). Para isto, foram considerados os transcritos diferencialmente expressos identificados no déficit hídrico progressivo, e os transcritos obtidos em cada um dos três períodos de déficit hídrico de curto período (T25-50, T75-100 e T125-150 min).

Os TDEs que apresentaram mesmo padrão de expressão gênica em resposta aos tratamentos induzidos (déficit hídrico progressivo e de curto período) foram selecionados para investigação dos *cis*-elementos presentes na região promotora.

### 5.3.5. Análise de enriquecimento dos transcritos comuns

Os TDEs comuns aos déficits hídrico de curto período e progressivo foram submetidos à análise de enriquecimento utilizando o banco de dados AgriGO (<http://bioinfo.cau.edu.cn/agriGO/analysis.php>). Os transcritos foram associados aos seus respectivos termos GOs (Plant GO slim) e comparados com os termos observados no genoma da soja utilizando os testes de Fisher e correção de Hochberg (FDR) ( $p = 0,05$ ). A percentagem relativa dos transcritos foi calculada pela razão do número de genes mapeados em cada termo GO e o número de genes observados no genoma referência para cada classe.

### 5.3.6. Identificação dos *cis*-elementos putativos nas regiões promotoras dos transcritos selecionados

A região promotora dos transcritos comuns ao déficit hídrico progressivo e de curto período foi determinada a 1000 pares de bases (pb) a partir do sítio ATG. As sequências foram obtidas a partir do banco de dados Phytozome (<http://www.phytozome.net/>) e submetidas ao software Genomatix (Cartharius et al., 2005). O módulo MatInspector foi utilizado considerando-se o valor de matriz de similaridade igual ou superior a 0,80.

Foi realizada também uma busca por modelos de promotores no mesmo banco de dados utilizando o módulo FrameWorker, com os parâmetros de similaridade = 0,8, variância entre dois elementos = 25 pb, distância entre dois elementos = 5 pb a 250 pb e número mínimo de sequências com framework de 50%. O módulo FrameWorker apresenta modelos complexos que são comuns às sequências analisadas e que atendam aos parâmetros estabelecidos. Os modelos são formados pelos FTs que ocorrem na mesma ordem e dentro de um intervalo de distância nas sequências submetidas.

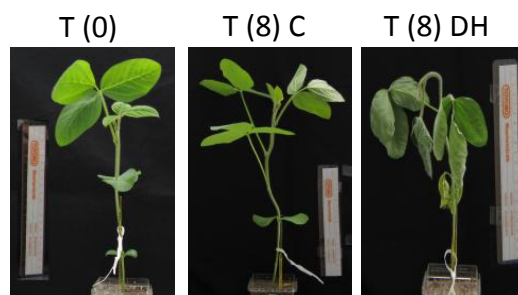
Para as análises realizadas nos diferentes módulos do software Genomatix, os TDEs comuns aos déficits hídricos progressivo e de curto período foram separados em três grupos, considerando a regulação da expressão foram separados em *Down* e *Up* regulados e ainda um grupo formado pelos TDEs sem classificação funcional (*Unknown*). Também foi realizada a análise dos TDEs, identificados nas intersecções entre dois tempos

dos déficits hídricos de curto período e progressivo, relacionados na categorização funcional com estresse de calor e seca/salinidade. Os três TDEs *Unknown* que apresentaram maior nível de expressão nos dois déficits hídricos foram analisados no módulo Gene2Promoter, para busca de regiões promotoras conservadas em genes ortólogos de outras espécies de plantas que pudessem dar alguma informação de função biológica para esses transcritos.

## 5.4. Resultados

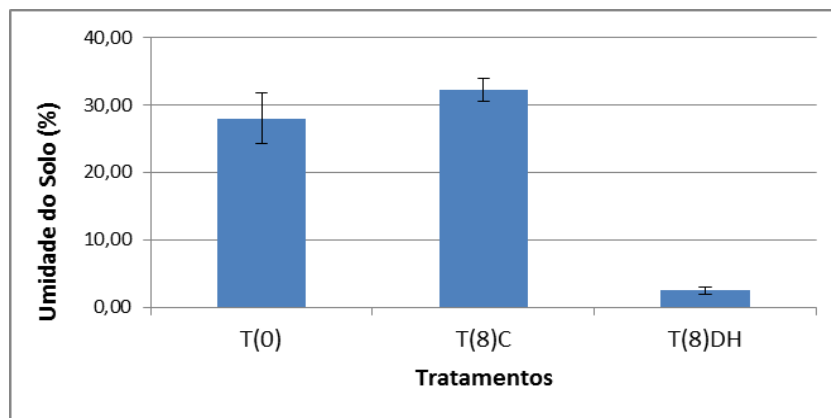
### 5.4.1 Validação do déficit hídrico progressivo aplicado nas plantas de soja

Com a imposição do déficit hídrico progressivo foi possível observar em comparação às plantas controles, que as plantas sob tratamento experimental apresentaram gradual redução da turgidez das folhas (Figura 7), efeito biofísico mais precoce e significativo do déficit hídrico. Após oito dias sob déficit hídrico as plantas apresentaram significativa alteração na arquitetura e murchamento das folhas, indicando que a perda de água por transpiração excedeu o ganho por absorção.



**Figura 7.** Aspecto observado nas plantas de soja cultivar BR 16 submetidas ao déficit hídrico progressivo e seus respectivos controles. Início do experimento (T0); oito dias condição controle T(8)C e oito dias sob déficit hídrico T(8)DH.

A umidade do solo foi medida para avaliar o conteúdo hídrico presente no substrato. Nas plantas controle a umidade do solo foi mantida próxima a 30% da capacidade de campo (CC), enquanto que, nas plantas submetidas ao tratamento experimental a umidade registrada no dia da coleta apresentou valores de 2,5%, caracterizando níveis severos de deficiência hídrica (Figura 8).



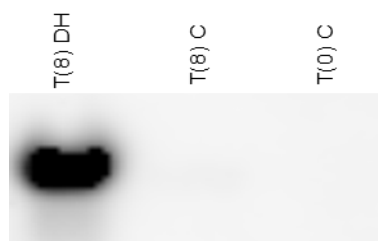
**Figura 8.** Umidade relativa do solo durante aplicação do déficit hídrico progressivo no tratamento e controle. Início do experimento (T0); oito dias condição controle T(8)C e oito dias sob déficit hídrico T(8)DH.

Os genes (Glyma03g30044, Glyma13g17980, Glyma09g31740) codificadores de proteínas LEA com expressão induzida em resposta à seca foram quantificados por RT-qPCR para a validação do tratamento de déficit hídrico imposto.

O conjunto de *primers* que apresentou melhor eficiência de amplificação foi para o gene alvo *LEA-8* (Glyma09g31740) e para o gene normalizador *Gm $\beta$ -Actina*, com valores próximos a 100% e sem ocorrência de dímeros de *primers* ou amplificação de produtos inespecíficos sendo, portanto, utilizados para validar a aplicação do tratamento experimental (Tabela 1).

A expressão relativa do gene *LEA-8* (Glyma09g31740) nas plantas no tempo de oito dias foi aproximadamente 20.000 vezes maior em relação às plantas controle (dados não mostrados). Esse aumento significativo na expressão do gene demonstra que o período de privação de água imposto foi suficiente para induzir o déficit hídrico e supostamente ativar os mecanismos moleculares de resposta à seca nas plantas.

O resultado de indução gênica observado para o gene *LEA-8* no RTqPCR foi confirmado na análise de expressão por *Northern Blot*, onde apenas nas plantas submetidas ao déficit hídrico a expressão do gene *LEA-8* foi detectada (Figura 9).

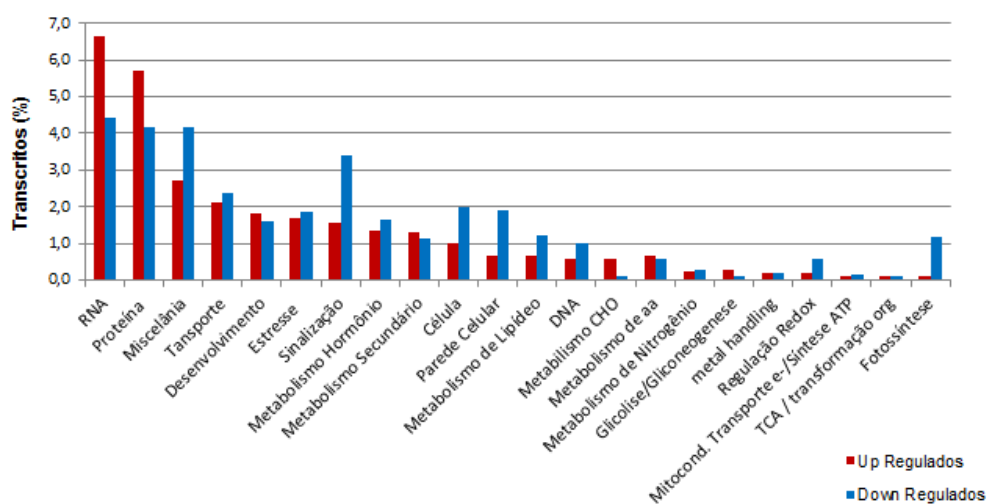


**Figura 9.** Northern Blot do gene *LEA-8* nas amostras das plantas de soja submetidas aos tratamentos de oito dias sob déficit hídrico progressivo T(8)DH e plantas controle T(8)C e no início do experimento T(0)C. A hibridização da sonda ocorreu somente no período de déficit hídrico progressivo.

#### 5.4.2. Identificação de transcritos diferencialmente expressos no déficit hídrico progressivo

Foram identificados 10.004 transcritos com expressão diferencial sob condições de déficit hídrico progressivo ( $FCA \geq +2$  e  $\leq -2$  e  $FDR 0,000085$  a  $0,02$ ); sendo 4.570 induzidos (*up* regulados) e 5.434 reprimidos (*down* regulados).

Uma das categorias mais abundante foi RNA (6,64% *up* regulado e 4,44% *down* regulado), onde foram identificados fatores de transcrição associados à estresses abióticos tais como: AP2/EREB, Homeobox, bHLH,  $C_2H_2$ , WRKY, bZip, MYB, CCAAT Box, (Figura 10).



**Figura 10.** Classificação por categoria funcional dos transcritos diferencialmente expressos identificados durante o déficit hídrico progressivo em plantas de soja cultivar BR 16. A frequência de cada classe foi normalizada pelo total de genes organizados nas diferentes categorias ( $n = 8.945$ ). Aproximadamente 10,6% dos transcritos não possui função atribuída.

A categoria de estresse (1,70% *up* regulado e 1,84% *down* regulado) contempla transcritos relacionados a respostas a estresses bióticos e abióticos. Os transcritos *up* e *down* regulados agrupados nesta categoria e relacionados a estresses abióticos foram identificados principalmente como responsivos a calor, à seca e à salinidade, sendo que a maioria desses transcritos apresentou perfil de *up*-regulação.

Na categoria sinalização (1,54% *up* regulado e 3,41% *down* regulado) que envolve a sinalização de receptores quinases, sinalização de cálcio e açúcares, os transcritos diferencialmente expressos em maior quantidade foram identificados como *down* regulados. O mesmo foi observado para a categoria fotossíntese (0,11% *up* regulado e 1,16% *down* regulado).

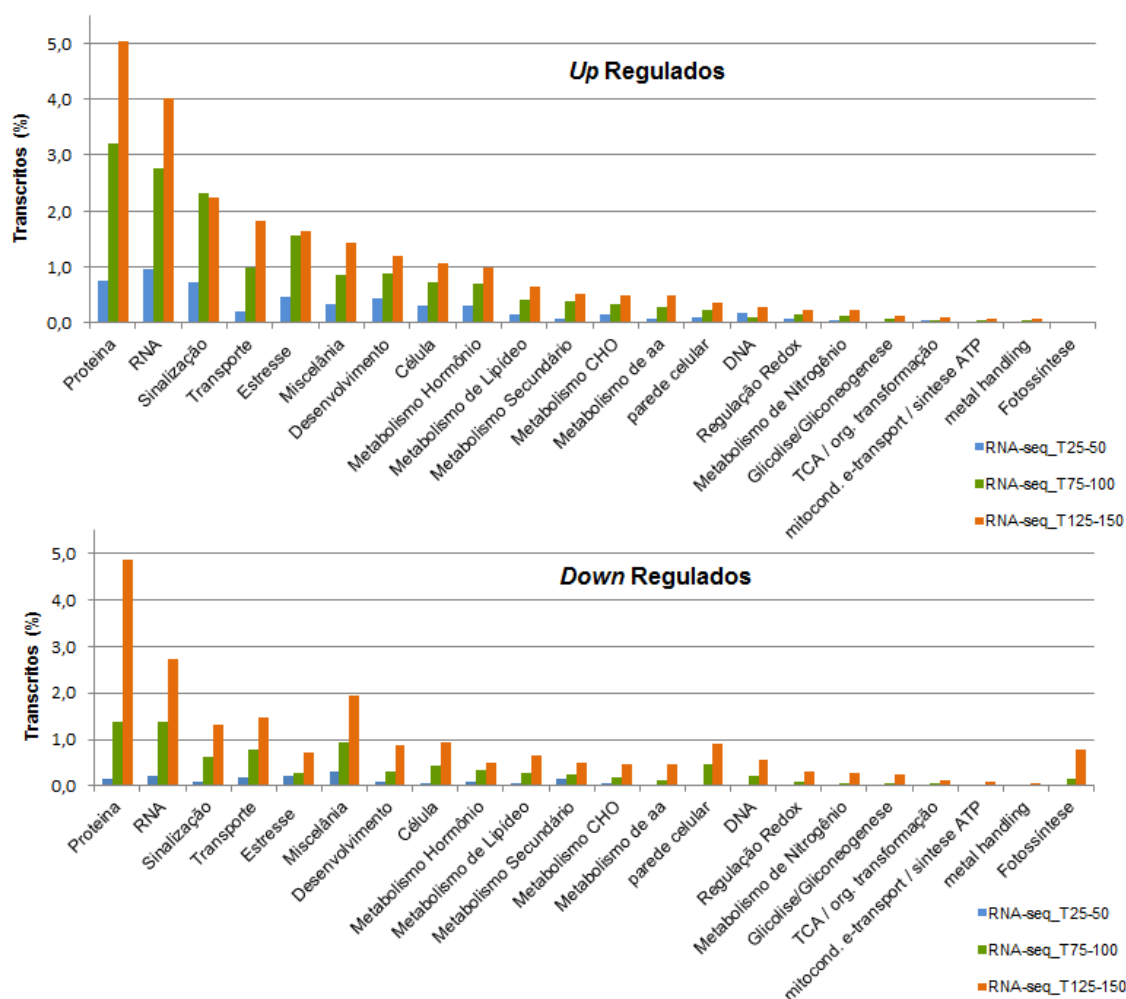
Foi também identificado na categoria de metabolismo de hormônio (1,33% *up* regulado e 1,64% *down* regulado), transcritos relacionados com sinal de transdução, síntese e degradação de ácido abscísico (ABA) e NCED (9-*cis*-epóxicarotenóide dioxigenase), enzima chave na via de regulação da biossíntese deste importante fitohormônio regulador central da tolerância à estresses abióticos em plantas.

#### **5.4.3. Identificação de transcritos diferencialmente expressos no déficit hídrico de curto período**

A quantidade total de transcritos diferencialmente expressos (TDEs) sob déficit hídrico de curto período foi de 10562 (6329 *up* regulados e 4233 *down* regulados). No tempo de T25-50 min de déficit hídrico foram 994 TDEs (734 *up* regulados e 260 *down* regulados), em T75-100 min foram 3444 TDEs (2.290 *up* regulados e 1.154 *down* regulados) e no tempo T125-150 min foram 6.124 TDEs (3.305 *up* regulados e 2.819 *down* regulados).

O agrupamento dos transcritos de acordo com as categorias funcionais revelou uma relação crescente entre o número de TDEs e o aumento no tempo de imposição de déficit hídrico aos quais as plantas foram expostas, em praticamente todas as categorias funcionais. Esta relação, contudo, foi mais acentuada para as categorias de proteína, RNA e transporte. Para as categorias sinalização e estresse, no entanto, houve um aumento em relação

ao período inicial (T25-50 min), mas entre os períodos T75-100 min e T125-150 min não ocorreu diferença (Figura 11).



**Figura 11.** Classificação por categoria funcional dos transcritos diferencialmente expressos identificados nos três períodos analisados do déficit hídrico de curto período em plantas de soja cultivar BR 16. A frequência de genes observados em cada categoria foi normalizada pelo total de genes identificados nos três períodos ( $n = 10.562$ ).

Igualmente ao observado nos resultados do déficit hídrico progressivo, dentro da categoria RNA foram identificados fatores de transcrição associados à estresses abióticos tais como AP2/EREB, Homeobox, bHLH, C<sub>2</sub>H<sub>2</sub>, WRKY, bZip, MYB, CCAAT Box, entretanto em menor número. O mesmo ocorreu com a categoria estresse, onde a maioria dos transcritos relacionados a estresses abióticos foram identificados como responsivos ao calor, à seca e à salinidade apresentando-se em maior quantidade transcritos *up* regulados. Na categoria fotossíntese a maioria dos transcritos diferencialmente expressos foi *down*

regulados principalmente no período T125-150 min. Este mesmo perfil foi observado nos resultados de déficit hídrico progressivo.

Na classe sinalização foi observado que os transcritos *up* regulados apresentam-se em maior número nos dois primeiros períodos de déficit hídrico (25-50 min e 75-100 min), ocorrendo uma tendência de diminuição no último período (Figura 11). Para os transcritos *down* regulados, no entanto, o número de TDEs aumentou progressivamente com a severidade da imposição do déficit hídrico (Figura 11). Na categoria de metabolismo de hormônios, transcritos relacionados com sinal de transdução, síntese e degradação de ácido abscísico (ABA) foram identificados nos três períodos de tratamento, somente nos tempos T75-100 min e T125-150 min, transcritos associados ao gene NCED foram identificados.

Uma síntese de alguns dos principais processos biológicos relacionados ao déficit hídrico encontrados na categorização funcional dos transcritos diferencialmente expressos nos tempos do déficit hídrico de curto período e no déficit hídrico progressivo é apresentada na Tabela 2.

**Tabela 2.** Processos observados na classificação funcional dos transcritos diferencialmente expressos durante os déficits hídricos. Número de transcritos *up* e *down* regulados identificados nos três tempos de déficit hídrico de curto período e no déficit hídrico progressivo.

Categoria Funcional	T25-50min		T75-100min		T125-150min		Progressivo	
	<i>Up</i>	<i>Down</i>	<i>Up</i>	<i>Down</i>	<i>Up</i>	<i>Down</i>	<i>Up</i>	<i>Down</i>
Sinalização de Ca	44	5	85	16	81	30	33	68
Sinalização de Açúcares	6	0	10	2	9	3	5	15
Sinalização Receptores Quinases	23	4	129	33	102	64	46	192
Transporte	23	19	105	84	196	160	189	241
Regulação da Transcrição (RNA)	100	20	279	129	396	229	527	423
Estresse Abiótico - Seca/Salinidade	2	2	9	10	12	26	11	15
Estresse Abiótico – Calor	11	6	51	5	61	16	59	27
Síntese Açúcares	0	0	0	1	0	1	2	5
Degradação Açúcares	1	0	7	3	10	3	4	5
Trealose	2	0	12	2	10	2	6	2
Síntese Amido	0	0	1	3	1	19	4	5
Degradação Amido	5	3	5	3	6	8	5	6
Síntese de Proteína	1	0	14	48	22	259	36	70
Degradação de Proteína	41	11	153	41	278	98	292	173
Metabolismo ABA	6	0	17	1	17	2	20	10

Na Tabela 3 estão relacionados, os principais fatores de transcrição descritos como estresse responsivos que foram identificados nas duas condições experimentais estudadas.

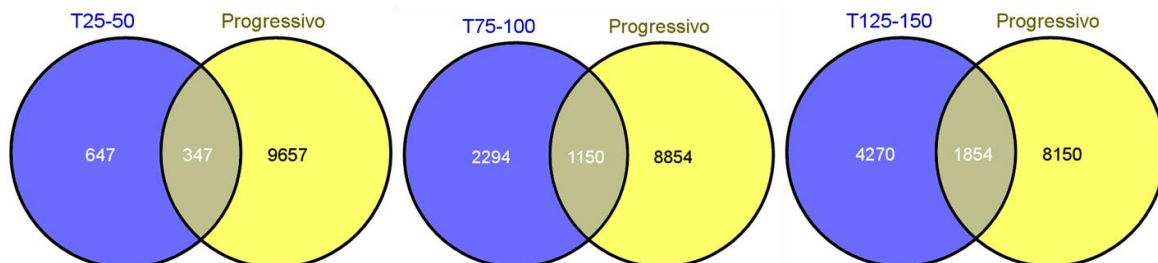
**Tabela 3.** Principais fatores de transcrição relacionados a estresses abióticos. Número de transcritos *up* e *down* regulados identificados nos três tempos de déficit hídrico de curto período e no déficit hídrico progressivo.

Categoria Funcional	T25-50min		T75-100min		T125-150min		Progressivo	
	<i>Up</i>	<i>Down</i>	<i>Up</i>	<i>Down</i>	<i>Up</i>	<i>Down</i>	<i>Up</i>	<i>Down</i>
MYB domain TF family	15	4	27	11	35	26	70	49
AP2/EREBP protein family	5	1	14	1	19	1	37	38
bHLH,Basic Helix-Loop-Helix family	--	1	11	11	19	16	28	62
HB,Homeobox TF family	1	3	7	10	12	16	28	25
C2H2 zinc finger family	8	--	18	5	26	9	24	--
WRKY domain TF family	20	--	44	--	39	--	23	19
bZIP TF family	7	1	27	3	39	5	22	13
C2C2(Zn) DOF zinc finger family	--	1	3	2	1	5	21	3
CCAAT box binding factor family	1	--	5	--	5	--	17	1
C3H zinc finger family	4	--	10	1	10	2	14	3
HSF,Heat-shock TF family	5	--	16	2	17	2	13	6
C2C2(Zn) CO-like, zinc finger family	--	--	2	5	4	7	12	7
G2-like TF family, GARP	1	--	3	3	5	3	12	7
PHOR1	6	--	10	--	8	--	9	3
MADS box TF family	--	--	1	--	2	--	8	5
Trihelix, Triple-Helix TF family	--	--	2	3	6	4	6	3
ARF, Auxin Response Factor family	1	--	3	5	4	10	5	3
AtSR TF family	2	--	5	--	6	1	5	--
Aux/IAA family	2	--	2	9	3	21	5	16
EIN3-like(EIL) TF family	--	--	2	--	4	--	5	--
ARR	--	1	--	2	3	--	3	3
C2C2(Zn) GATA TF family	1	1	1	4	1	5	1	8
JUMONJI family	3	--	2	--	2	2	1	1
NAC domain TF family	1	1	1	1	1	1	1	3
SNF7	--	--	4	--	7	--	1	--

#### 5.4.4. Transcritos identificados no déficit hídrico progressivo e de curto período

A comparação entre os transcritos diferencialmente expressos obtidos no déficit hídrico progressivo e em cada um dos tempos avaliados no déficit

hídrico de curto período possibilitou a seleção de transcritos comuns expressos em ambos os experimentos de déficit hídrico (Figura 12).



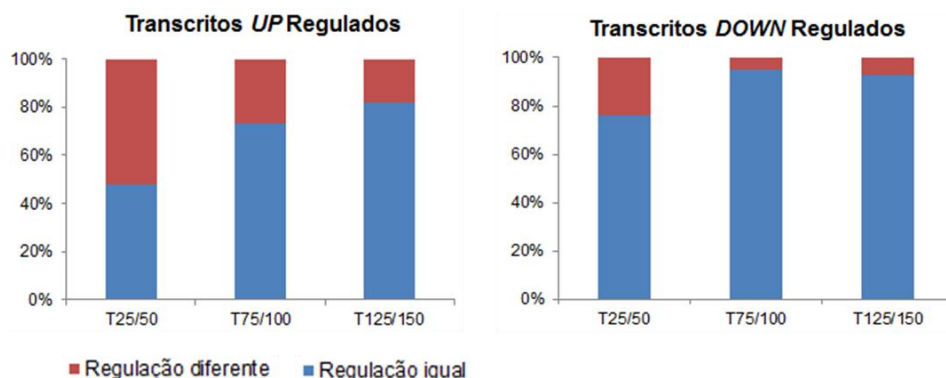
**Figura 12.** Diagrama de Venn. Comparação entre os transcritos diferencialmente expressos identificados no déficit hídrico progressivo e de curto período. Observa-se o número de transcritos exclusivos e comuns entre os déficits hídricos progressivo de curto período. A quantidade de transcritos *up* e *down* regulados entre os transcritos comuns em cada comparação foi T25-50 min / Progressivo 264 *up* e 83 *down* regulados, T75-100 min / Progressivo 764 *up* e 205 *down* regulados e T125-150 / Progressivo 1042 *up* e 812 *down* regulados.

No transcorrer do tratamento de déficit hídrico de curto período, houve um aumento progressivo da quantidade de transcritos identificados como diferencialmente expressos. Entre 30% e 35% dos transcritos identificados no déficit hídrico de curto período apresentaram correspondentes no déficit hídrico progressivo (Figura 12).

Considerando o total de transcritos obtidos no déficit hídrico progressivo, um aumento no número de transcritos comuns a ambos os experimentos foi observado. Aproximadamente 3,5% dos transcritos foram comuns nos períodos de T25-50 min, enquanto nos períodos T50-75 min e T125-150 min o percentual de transcritos em comum aumentou para 11,5% e 18,5% respectivamente (Figura 12). Assim, um maior número de transcritos comuns foi identificado à medida que os períodos de exposição ao déficit hídrico de curto período se tornaram mais prolongados.

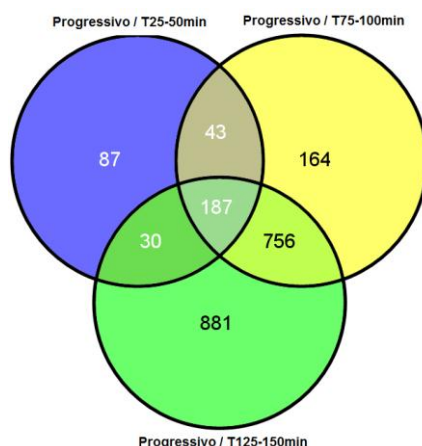
Entre os transcritos induzidos, no tempo inicial de déficit hídrico (T25-50 min) 50% dos transcritos em comum apresentaram o mesmo perfil de regulação, enquanto nos tempos T75-100 min e T125-150 min, essa relação aumentou para 73% e 82%, respectivamente (Figura 13). Com o aumento do tempo de exposição ao déficit hídrico, foi observado um aumento na semelhança do padrão de expressão, com os transcritos apresentando o

mesmo perfil de regulação da expressão gênica entre os processos avaliados (déficit hídrico progressivo e de curto período).



**Figura 13.** Regulação dos transcritos diferencialmente expressos comuns ao déficit hídrico progressivo e nos diferentes tempos de déficit hídrico de curto período.

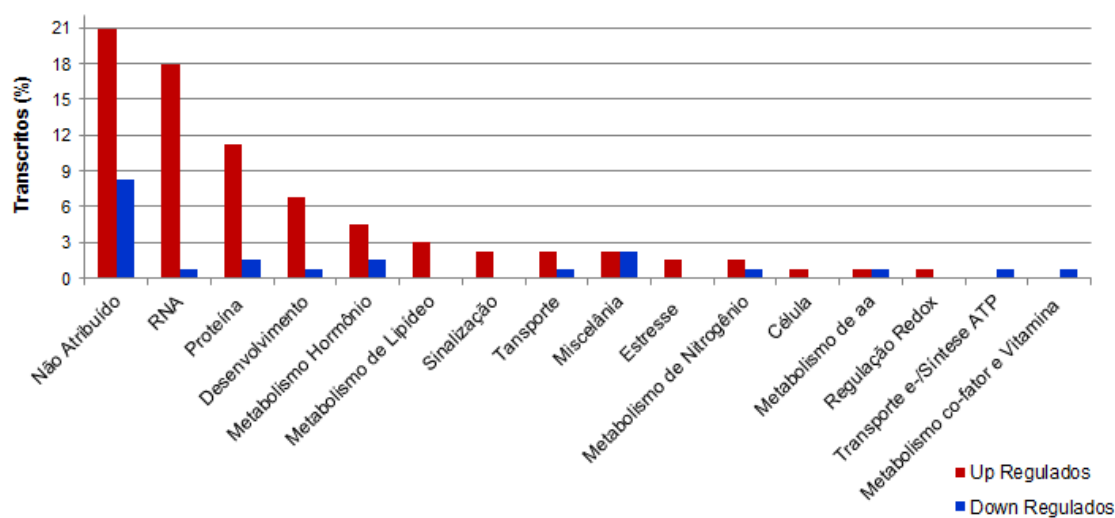
Após comparar os transcritos identificados como diferencialmente expressos nos três períodos avaliados no déficit hídrico de curto período e no déficit hídrico progressivo, 187 foram identificados em comum, considerados ativados e expressos em resposta ao déficit hídrico progressivo, e ao longo dos 150 min do déficit hídrico de curto período (Figura 14). Deste total, 134 transcritos apresentaram o mesmo perfil de regulação em ambos os tratamentos de déficits hídricos, sendo 107 *up* regulados e 27 down regulados, e por esta razão foram selecionados para a análise de busca por *cis* elementos putativos na região promotora dos mesmos.



**Figura 14.** Diagrama de Venn, comparação entre os transcritos diferencialmente expressos comuns em cada um dos tempos de déficit hídrico de curto período e o déficit hídrico progressivo.

### 5.4.5 Categorização Funcional

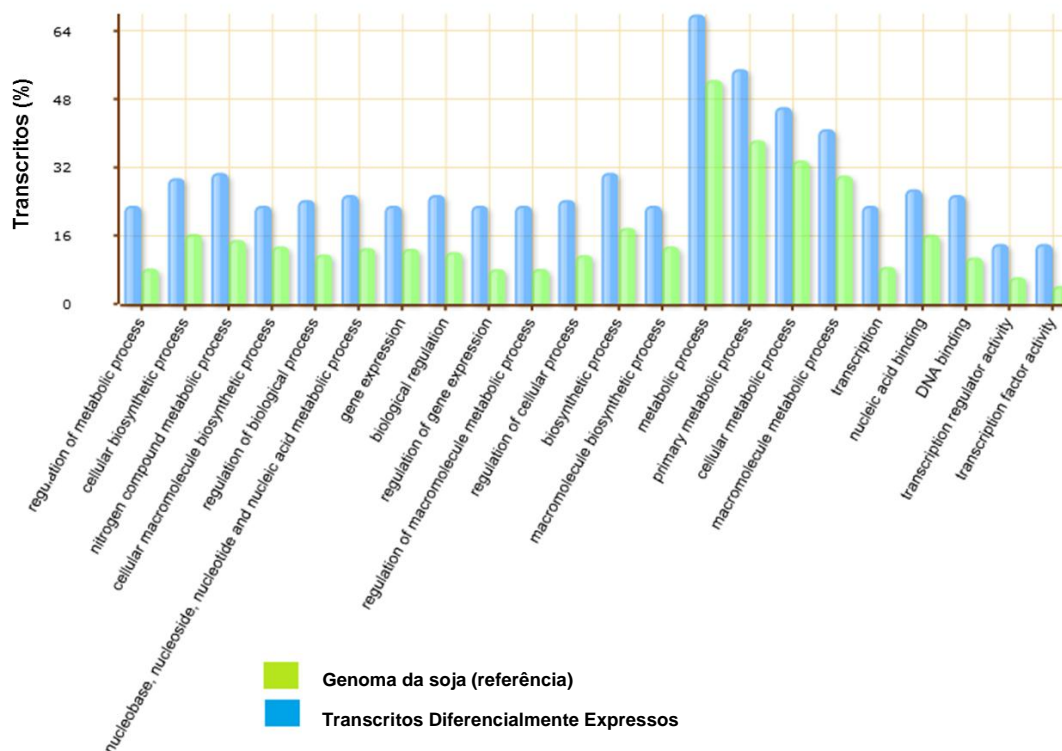
Para uma melhor compreensão das alterações moleculares decorrentes da aplicação do déficit hídrico, e identificação de TDEs potencialmente importantes em ambos os processos avaliados (déficit hídrico progressivo e déficit hídrico de curto período), os 134 transcritos comuns identificados como diferencialmente expressos que apresentaram o mesmo perfil de regulação gênica foram organizados de acordo com a categorização funcional das respectivas proteínas. Os transcritos *up* regulados foram categorizados em 13 categorias funcionais, enquanto que, os transcritos *down* regulados em 10. Para aproximadamente 21% dos transcritos *up* e 8% dos transcritos *down* regulados não foi atribuída pelo programa utilizado nenhuma categoria funcional. As categorias de sinalização (2,24%), estresse (1,49%), e regulação redox (0,75%) foram exclusivamente *up* reguladas enquanto que as categorias transporte de elétrons/síntese de ATP (0,75%) e metabolismo co-fator e vitamina (0,75%) foram exclusivamente *down* reguladas (Figura 16).



**Figura 15.** Classificação por categoria funcional dos transcritos diferencialmente expressos comuns ao déficit hídrico progressivo e de curto período. A frequência de TDEs observados foi normalizada pela somatória dos genes categorizados comuns e com mesma regulação da expressão (n = 134).

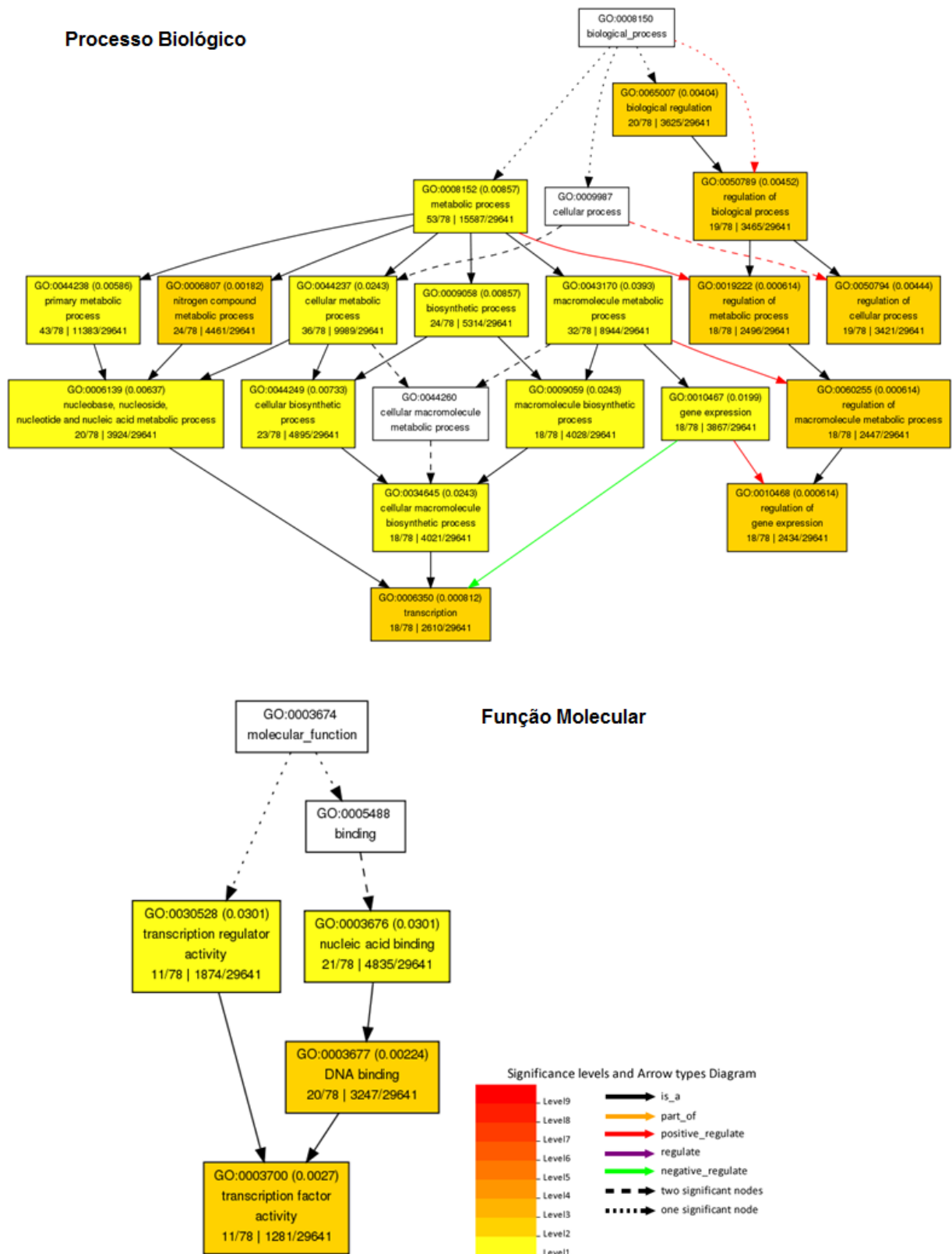
A análise de enriquecimento simples (SEA – *Simple Enrichment Analysis*) realizada no banco de dados agriGO, agrupou 78 dos transcritos

diferencialmente expressos em 18 diferentes níveis de processos biológicos e quatro níveis de funções moleculares todos enriquecidos em relação ao genoma de referência (Figura 17).



**Figura 16.** Análise de enriquecimento dos transcritos diferencialmente expressos comuns ao déficit hídrico progressivo e de curto período e que apresentaram mesma regulação da expressão.

Diagramas hierárquicos com os diferentes termos GOs agrupando os transcritos diferencialmente expressos com seus respectivos níveis de significância para processos biológicos e função molecular são apresentados na Figura 18. Funções moleculares relacionadas com fatores de transcrição e ligantes de DNA apresentaram-se mais enriquecidas.



**Figura 17.** Diagrama hierárquico de categoria funcional dos transcritos diferencialmente expressos comuns ao déficit hídrico progressivo e de curto período e com mesma regulação da expressão. Os quadros coloridos indicam termos com significância ( $p \leq 0,05$ ), e os termos não significativos são mostrados em quadros brancos. O grau de saturação da cor dos quadros é positivamente correlacionado com o nível de enriquecimento do termo.

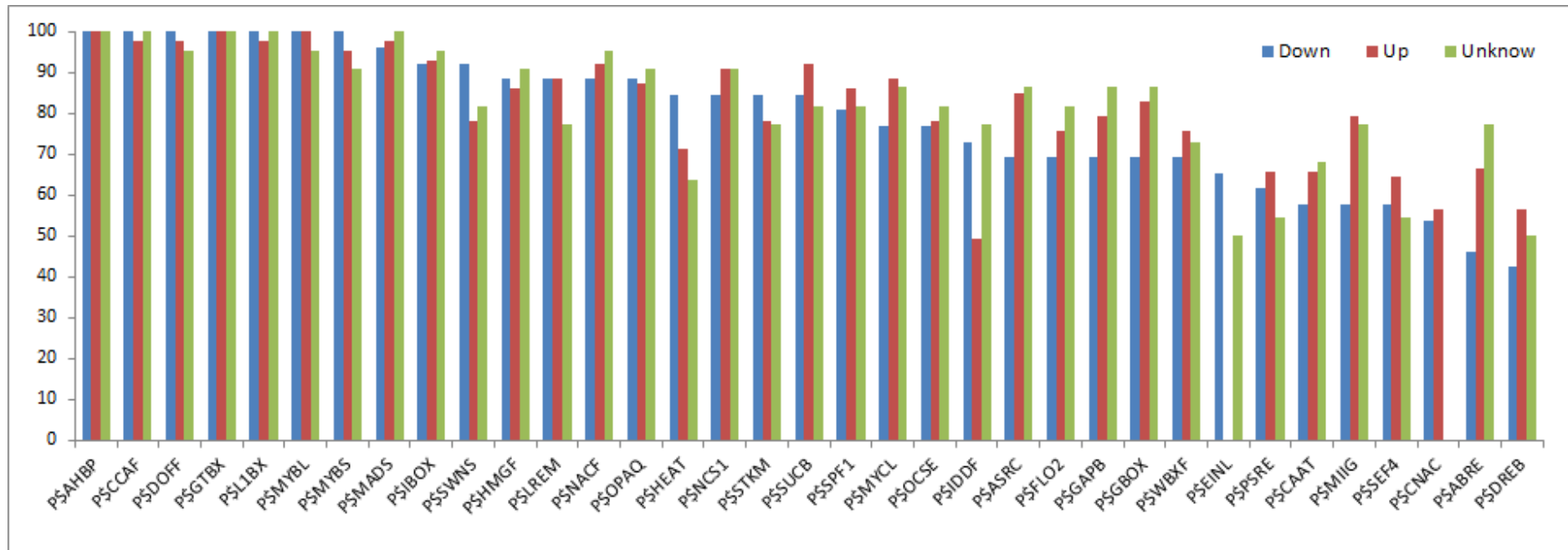
#### 5.4.4. Análise de *cis*-elementos putativos

O ponto de início de transcrição gênica determina o posicionamento relativo dos *cis*-elementos putativos encontrados nas regiões promotoras dos TDEs analisados. Para esta análise foram consideradas somente as fitas de orientação positivas.

A análise dos transcritos utilizando as informações disponíveis no banco de dados Genomatix (Módulo MatInspector) revelou a presença de 85 motivos comumente encontrados em regiões promotoras de plantas com frequência entre 2 e 100%. Além da presença dos bem caracterizados TATA Box, CCAAT Box e GC Box e INR encontrados em promotores de eucariotos e necessários para o acoplamento da maquinaria de transcrição, foram também encontrados diversos *cis*-elementos especificamente presentes em regiões promotoras de genes induzidos por fatores abióticos (Tabela 2 – Material suplementar) tais como ABRE, DREB, HEAT e SALT envolvidos respectivamente com estresse de seca, calor e salinidade. Outros elementos putativos relacionados a processos de respostas a estresses bióticos e abióticos também foram identificados tais como SUCR e WBox.

Sítios putativos de ligação de fatores de transcrição como MYB, MYC e NAC também foram identificados na região promotora dos TDEs comuns ao déficit hídrico de curto período e ao déficit hídrico progressivo. Além de elementos putativos de ligação de hormônios vegetais (AREF, EREF e GCC-Box), do ciclo circadiano (EE e CCA1) e responsivos a luz (GAP-Box e LREM).

Os *cis*-elementos que apresentaram frequência  $\geq$  a 50% entre os 134 TDEs foram selecionados e apresentados na Figura 19. Alguns *cis*-elementos foram encontrados em todos os transcritos tais como proteína HDZip ATHB5, GT1 Box, fator de ciclo circadiano e MYB Cov1. O número de vezes que o mesmo *cis* elemento putativo foi encontrado nos transcritos variou de 1 a 15 vezes.



**Figura 18.** Frequência de *cis*-elementos putativos encontrados nos 134 transcritos diferencialmente expressos identificados em comum a ambos os experimentos de déficit hídrico progressivo e de curto período.

Em cada grupo de TDEs analisado pelo módulo MatInspector (*Down* regulados, *Up* regulados e *Unknown*) foram identificados 35 *cis*-elementos comuns com frequência  $\geq$  a 50%. No módulo Frame Worker para todos os grupos de TDEs analisados, foram encontrados modelos de promotores formados por combinações entre os *cis*-elementos GTBX, AHBP e TBP (Tabela 4).

**Tabela 4.** Elementos encontrados nos modelos de promotores obtidos pelo módulo FramWorker do software Genomatix nos transcritos diferencialmente expressos comuns ao déficit hídrico progressivo e de curto período.

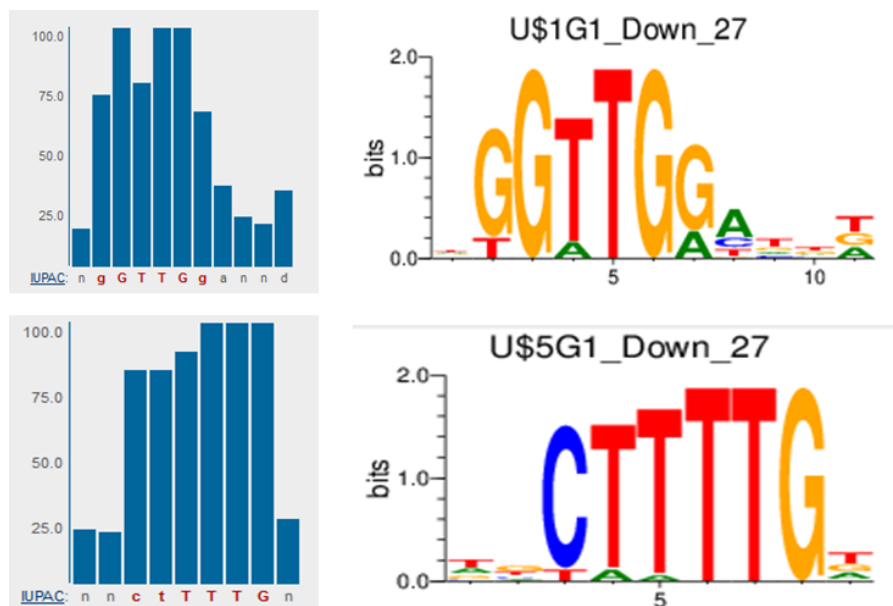
Grupo de TDEs	Quantidade de elementos	Elementos
<i>Unknown</i>	2	GTBX-AHBP AHBP-GTBX
<i>Down</i> Regulados	2	TBP-AHBP-GTBX AHBP-GTBX-TBP
<i>Up</i> Regulados	4	TBP-AHBP-GTBX AHBP-TBP AHBP-GTBX-TBP AHBP
Seca/Calor	6	1LBX-MYB 1LBX-TBP-AHBP TBP-AHBP AHBP-GTBX AHBP TBP-GTBX-AHBP

Nos três grupos de TDEs analisados, muitos CREs do tipo TATA-Box foram preditos em regiões distantes do ponto de início de transcrição (-115 pb a -986 pb), apresentando principalmente os sítios de TATA *Box1* (gctcTTTAAAAAAA) e TATA *Box2*. (ttccTATAtatctct). Nos três grupos foram preditos sítios putativos TATA *Box2* a aproximadamente 30 nucleotídeos anteriores ao ponto de início de transcrição, que aparenta ser o TATA-box funcionalmente mais plausível pelo seu posicionamento.

A comparação dos três TDEs *Unknown* com sequências promotoras de genes ortólogos revelou similaridade entre as regiões promotoras de genes

relacionados com glutamina sintetase GS (GeneID:100266191) e 1,4-dihidroxi-2-naftoato polipreniltransferase (GeneID:842354) em relação aos genomas de uva e *Arabidopsis* respectivamente.

Apesar do grande número de cis-elementos já descritos (para plantas são até o momento 209 matrizes em 84 famílias) ainda há elementos que não foram identificados, o módulo CoreSearch do software Genomatix permite a definição de motivos desconhecidos. Entre os TDEs *down* regulados foram identificados novos motivos que necessitam ser avaliados quanto à relação com o déficit hídrico (Figura 19).



**Figura 19.** Novos putativos cis-elementos encontrados pelo módulo CoreSearch do software Genomatix nos transcritos diferencialmente expressos *down* regulados comuns ao déficit hídrico progressivo e de curto período.

## 5.5. Discussão

A associação entre o fenômeno físico de murcha observado (Figura 7) e as análises moleculares (Figura 9) indicaram que as plantas de soja estavam sob déficit hídrico. A diminuição da turgescência é a primeira e mais pronunciada resposta ao déficit hídrico, ocorrendo à medida que o conteúdo de água da planta descrece, e que as células vegetais se contraem reduzindo a pressão de turgidez contra as paredes celulares (Taiz e Zeiger, 2010).

Além de desencadear respostas fisiológicas ao déficit hídrico, as plantas modulam alterações moleculares visando à aclimatação à seca. Genes que codificam para proteínas LEA são considerados importantes neste processo. Essas proteínas atuam como moléculas de agregação de água, no sequestro de íons e na estabilização das membranas, e estudos indicaram ainda um aumento de sua concentração em tecidos vegetais expostos ao déficit hídrico (Olvera et al., 2010). A alta expressão do gene *LEA-8* observada na análise de RTqPCR corrobora esta função biológica observada em condições de déficit hídrico, servindo como indicador de que o tratamento imposto foi apropriado para ativar genes de resposta à seca.

De modo geral, nas respostas das plantas aos estresses bióticos ou abióticos alguns genes têm seus níveis de expressão aumentados, enquanto outros são reprimidos. Essas alterações na expressão gênica podem estar envolvidas no mecanismo de tolerância ao estresse como também resultar de danos celulares sofridos pela planta (Rodriguez et al., 2005). Segundo Harb e colaboradores (2010) a categorização funcional de transcritos diferencialmente expressos em diversos processos biológicos constitui uma importante ferramenta para compreensão das alterações moleculares decorrentes da aplicação do déficit hídrico. O aumento progressivo de transcritos *up* e *down* regulados nas diferentes categorias funcionais, observado à medida que níveis mais severos de déficit hídrico foram impostos às plantas (Figura 11), durante o experimento de déficit hídrico de curto período, demonstra o envolvimento de muitos genes e a interação com múltiplas vias moleculares que visam o restabelecimento da homeostase celular durante eventos de estresse abiótico (Bansal et al., 2011).

As alterações no nível de expressão gênica de transcritos relacionados à fotossíntese (Figuras 10 e 11) eram esperadas, uma vez que a fotossíntese é um dos primeiros processos biológicos a responder ao déficit hídrico. Metabolicamente a eficiência da reação fotossintética na planta é afetada pela privação de água através de vários mecanismos como alteração no nível dos metabólicos intermediários, a inibição do sistema de transporte de elétrons fotossintéticos, o fechamento dos estômatos reduzindo a entrada de CO<sub>2</sub> e a alteração na taxa de respiração, sendo que a inibição da fotossíntese aumenta à medida que o estresse torna-se mais rigoroso (Taiz e Zeiger, 2010).

Em condições de rápida desidratação, as plantas desenvolvem ainda o estresse oxidativo como um efeito secundário, potencialmente muito prejudicial à maquinaria fotossintética (Chaves et al., 2009; Xiao et al., 2009). Transcritos diferencialmente expressos para a categoria regulação redox (Figuras 10 e 11) foram associados ao ascorbato, a glutathione S-transferase e a catalase, demonstrando a ocorrência de estresse oxidativo no tecido vegetal e a necessidade de produção regulada de espécies reativas de oxigênio por sistemas enzimáticos visando reestabelecer a homeostase redox celular em condições de estresse.

Outra resposta das plantas à seca bastante conhecida é o ajuste osmótico, o qual é atingido através da produção e acúmulo de osmólitos e solutos orgânicos tais como prolina e açúcares. Neste estudo não foram identificados transcritos relacionados à síntese de sacarose nos períodos iniciais de déficit hídrico de curto período (T25-50 min e T75-100 min), contudo esses transcritos foram detectados no tempo T125-150 min e no déficit hídrico progressivo (Tabela 2). Por outro lado, a degradação de sacarose apresentou-se progressivamente *up* regulada nos tempos do déficit hídrico de curto período e no déficit hídrico progressivo. Este processo de degradação da sacarose no citosol ocorre em razão da redução na concentração de hexoses fosfato, consequência da redução das taxas fotossintéticas em condições adversas como a seca (Taiz e Zeiger, 2010). Nestas condições, a assimilação de carbono e o metabolismo primário são também afetados, e desta forma importantes metabólitos como açúcares, álcool açúcares e aminoácidos sofrem alterações de concentração (Krasensky e Jonak, 2012; Valerio et al., 2011).

Transcritos associados a outros açúcares como trealose e rafinose foram também progressivamente *up* regulados em função do déficit hídrico (Tabela 2), provavelmente, associados ao processo de ajuste osmótico (Reguera et al., 2012; Kondrak et al., 2012). Além desta função, os açúcares podem atuar também como moléculas sinalizadoras (Rolland et al., 2006; Muller et al., 2011) e estão envolvidos no balanço redox celular influenciando a produção de ROS em condições adversas (Couée et al., 2006). Cabe ressaltar que, nem todos os açúcares solúveis desempenham papéis similares em eventos associados ao metabolismo de plantas sob estresse. A sacarose e a trealose podem atuar como substratos para a respiração celular, ou como osmólitos para manter a homeostase celular, enquanto a frutose não está relacionada com osmoproteção e parece estar relacionada com a síntese de metabólitos secundários (Rosa et al., 2009).

De forma semelhante ao processo de síntese de sacarose, o metabolismo de amido é ainda mais alterado em condições adversas. Neste trabalho, foi observado que o número de transcritos diferencialmente expressos envolvidos com a síntese e degradação de amido aumentou progressivamente à medida que o déficit hídrico de curto período se tornou mais severo (Tabela 2). Para a etapa de síntese de amido, um maior número de transcritos *down* regulados foi observado e na degradação os transcritos *up* regulados foram identificados em maior quantidade. Esse mesmo perfil gênico foi também identificado para os TDEs identificados no déficit hídrico progressivo (Tabela 2). A degradação do amido ocorreu em razão da baixa concentração de açúcar que aumenta a mobilização de reservas e a exportação, enquanto que, quando açúcares estão em abundância a estocagem de carboidratos é promovida (Taiz & Zeiger, 2010).

Também é de conhecimento que os processos de sinalização são essenciais para as respostas das plantas ao déficit hídrico, pois após a percepção de alterações do potencial hídrico, uma cascata de sinais transducionais envolvendo fosforilação e desfosforilação de proteínas mediada por várias proteínas quinases e fosfatases é ativada. Muitas proteínas quinases regulatórias envolvidas na sinalização de estresse abiótico são dependentes de hidrogênio ativado (MAPK) e de  $\text{Ca}^{2+}$  (CDPK). Esses níveis de  $\text{Ca}^{2+}$  intracelular podem ser modulados por mensageiros secundários, iniciando uma cascata de

fosforilação de proteínas diretamente envolvidas na proteção celular ou de fatores de transcrição culminando na alteração da expressão de diversos genes específicos (Shao, 2007; Seki et al., 2007). Várias MAPKs e CDPKs têm sido identificadas em plantas relacionadas ao déficit hídrico (Umezawa, 2006). Corroborando suas funções biológicas, transcritos associados ao cálcio, destacando-se receptores quinase, fosfatase e MAP (Tabela 2) foram identificados nas categorias transporte e sinalização, com aumento no número de TDEs (*up* e *down* regulados) a medida que o déficit hídrico se tornou mais severo.

Em condições de estresse, o metabolismo de aminoácidos é também amplamente alterado, sendo a síntese de proteínas diminuída e a proteólise aumentada. Esse padrão foi observado para os TDEs identificados associados ao metabolismo de proteína em ambos os déficit hídricos aplicados (Tabela 2). Essa é uma tendência consistente nas respostas ao estresse abiótico em relação à regulação no início do metabolismo energético e na síntese de proteínas. Isso pode indicar uma conservação de energia e refletir em uma mudança de crescimento da planta como mecanismos de proteção (Cramer, 2011).

Para o metabolismo de nitrogênio foram encontrados mais transcritos *down* regulados à medida que o déficit hídrico de curto período foi mais severo bem como no progressivo (Tabela 2). Como parte do metabolismo global das plantas, o metabolismo do nitrogênio também é afetado pelo déficit hídrico. A conversão do nitrato a amônio é influenciado negativamente por fatores abióticos. Os aminoácidos glutamina e glutamato produzidos nesta reação servem para translocar nitrogênio orgânico de fontes para drenos, bem como para a incorporação a outros aminoácidos por meio de reações de transaminações (Taiz e Zeiger, 2010).

Neste trabalho, na categoria de estresse, transcritos diferencialmente expressos responsivos ao calor, à seca e a salinidade foram identificados, sendo que ocorreu um aumento dos transcritos associados à resposta ao calor com o aumento do tempo de exposição ao déficit hídrico (Tabela 2). Esses dados demonstram que múltiplos estresses podem desencadear um padrão similar de respostas, além de evidenciar a relação existente entre as respostas das plantas ao déficit hídrico e ao calor (Seki et al., 2002; Venu et al., 2013). Os

transcritos relacionados à resposta ao calor podem estar envolvidos com o processo de transpiração, inibido pelo déficit hídrico, através do qual as plantas dissipam calor. Estes resultados mostram que transcritos envolvidos nas respostas ao estresse por calor podem também estar associados a respostas ao déficit hídrico (Rodríguez et al., 2005; Venu et al., 2013).

Além de desempenhar um papel crucial como moléculas sinalizadoras na regulação do desenvolvimento da planta, os hormônios vegetais tais como auxinas, ABA, citocininas, giberelinas, etileno, brassinosteróides, jasmonatos e ácido salicílico, também participam de respostas aos estresses ambientais. Transcritos diferencialmente expressos associados a todos esses hormônios foram identificados neste estudo, especificamente, relacionado ao hormônio ABA, foram observados transcritos *up* regulados em maior número nos níveis mais severos do déficit hídrico de curto período e no déficit hídrico progressivo (Tabela 2).

O hormônio ABA desempenha um papel crítico nas respostas aclimatativas a estresses como a seca e a alta salinidade. Nestas condições, ocorre o acúmulo de ABA nas células vegetais, promovendo o fechamento dos estômatos em células-guarda e regulando a expressão de muitos genes cujos produtos podem proteger os tecidos vegetativos da desidratação celular ou da alta pressão osmótica (Umezawa et al., 2011). Transcritos associados ao gene *NCED* (9-cis-epóxicarotenóide dioxigenase), enzima chave na síntese de ABA, foram identificados nos tempos de 75-100 min e 125-150 min do déficit hídrico de curto período. Consistente com a importância desta enzima na regulação da biossíntese de ABA; no genoma de *Arabidopsis*, existem cinco genes *NCED*, dos quais o *AtNCED3* está especificamente envolvido nas respostas ao déficit hídrico e à alta salinidade (Iuchi et al., 2000; Behnam et al., 2013). Estudos realizados em plantas de feijão (Qin e Zeevaart, 1999), tomate (Burbidge et al., 1999), feijão de corda (Iuchi et al., 2000), abacate (Chernys e Zeevaart, 2000), amendoim (Wan & Li., 2005) e laranja (Rodrigo et al., 2006) constataram a indução de genes *NCED* e o acúmulo de ABA associados ao déficit hídrico e alta salinidade, corroborando a importância deste gene para as respostas à seca dependentes do metabolismo de ABA.

O papel desempenhado por ABA em respostas aos estresses abióticos é tão significativo que muitos dos mecanismos de defesa vegetal são ativados

por fatores de transcrição (FTs) em duas vias de resposta, uma ABA dependente e outra independente de ABA. Estes fatores de transcrição específicos regulam a expressão de genes responsivos a estresse (Shinozaki e Yamaguchi-Shinozaki, 2007) pela ligação/interação com *cis*-elementos presentes na região promotora dos genes alvos, que quando ativados resultam na tolerância ao estresse abiótico (Agarwal e Jha, 2010).

Neste trabalho, transcritos diferencialmente expressos foram associados a fatores de transcrição previamente relacionados a respostas vegetais ao déficit hídrico (Tabela 3). Dentre estes, estão os genes das famílias DREB e AREB, que atuam, respectivamente, nas vias independente e dependente de ABA, ligando-se às regiões *cis* específicas DRE e ABRE localizadas na região promotora de genes alvos. Os FTs DREB e AREB ativam a transcrição de genes *downstream* responsáveis pelas respostas defensivas de proteção celular ao déficit hídrico (Yamaguchi-Shinozaki e Shinozaki, 2006; Shinozaki e Yamaguchi-Shinozaki, 2007; Xu et al., 2008; Nakashima et al., 2009; Wang et al., 2011; Mizoi et al., 2012).

Além dos FTs já amplamente descritos como responsivos ao déficit hídrico também foram identificados outros FTs que recentemente foram relacionados às respostas das plantas a estresses abióticos tais como (CBFs) CAAT Box que podem promover pequenos ajustes a várias condições ambientais (Petroni et al., 2012). Estes dados evidenciam a complexidade das respostas ao déficit hídrico e sugerem o envolvimento de outros fatores reguladores da transcrição que necessitam ser ainda melhor estudados.

A análise das regiões promotoras dos 134 TDEs que apresentaram a mesma regulação da expressão e foram identificados em comum nos déficits hídricos de curto período e progressivo, identificou *cis*-elementos relacionados a vias metabólicas de respostas a estresses abióticos, tais como SUCB, ABRE, GTBX, NAC, CCAF, EE e HEAT (Tabela 4). SUCR é uma sequência putativa de um motivo presente na região promotora de genes responsivos à açúcares. É sabido que a sacarose atua como um soluto. Em plantas de feijão comum (*Phaseolus vulgaris*), respostas ao déficit hídrico foram associadas ao metabolismo de sacarose (Cuellar-Ortiz et al., 2008). ABRE é um *cis*-elemento relacionado ao ácido abscísico, um hormônio diretamente relacionado à resposta ao déficit hídrico, que como já mencionado, atua no fechamento

estomático e induz a expressão de genes responsivos à seca (Yamaguchi-Shinozaki e Shinozaki 2006; Fujita et al., 2011; Umezawa, 2011).

Além desta função, estudos recentes apontaram uma correlação entre o ABA, ciclo circadiano (CAAF e EE motivos) e as respostas das plantas à seca, sugerindo uma conexão entre estas vias metabólicas (Legnaioli et al., 2009; Wilkins et al., 2010, Sanchez et al., 2011). Legnaioli e colaboradores demonstraram que o gene *TOC1A* (*Timing Of Cab expression*) se liga a genes ABA-responsivos (*ABAR/CHLH/GUN5*) dirigindo sua expressão circadiana. Em outro estudo, Covington e colaboradores (2007) apontaram que não apenas genes ABA-induzidos são expressos durante o dia, mas que enzimas chaves na biossíntese do hormônio, como NCED, são também controladas pelo relógio circadiano. Juntamente com os motivos CCAF e EE, o *cis*-elemento GTBX é também responsivo à luz (Gilmartin et al., 1990, Simpson et al., 2003).

A presença de *cis* elementos putativos circadianos e estresse-responsivos reforçam as vias cruzadas que percebem as oscilações do ambiente externo (dia/noite/estação do ano) e a regulação do relógio biológico vegetal. Estas conexões coordenam o metabolismo da planta sob condições de estresse para uma sincronização fina do transcriptoma global e dos processos fisiológicos, visando reduzir o dano celular devido à privação de água e ativar a expressão gênica em condições ambientais específicas.

Outro *cis* elemento putativo relacionado ao déficit hídrico identificado nos transcritos em comum foi o NAC (*NAM, ATAF1/2, CUC2*) (Zheng et al., 2009; Nakashima et al., 2012; Lu et al., 2012). A super expressão de genes NAC aumentou a tolerância à desidratação em milho, (*ZmSNAC1*) (Lu et al., 2012), arroz (*ONAC045*) (Zheng et al., 2009) e tabaco (Ramegowda et al., 2012). Ainda, a presença do *cis*-elemento putativo HEAT aponta a importância das respostas cruzadas à seca e ao calor, uma vez que, transcritos associados à fatores de transcrição responsivos ao calor também foram identificados. *Cis* elementos putativos associados ao desenvolvimento vegetal também foram identificados tal como MADS, um elemento associado ao florescimento (Schwarz-Sommer et al., 1990; Svensson, 2000).

Ainda, elementos relacionados ao sítio de início da transcrição foram identificados tais como PTBP e CAAT e outros relacionados à atividade transcricional tais como DOFF, uma família de fatores de transcrição,

identificada em associação com diversos promotores de genes específicos de plantas, envolvidos no controle positivo/negativo da transcrição (Schneidereit et al., 2008) e HMGF apontado como importante nos processos de replicação, transcrição e montagem do nucleossomo, mas também apresentando mudanças na expressão gênica em situações de estresse em *Arabidopsis* como reportado por Kwak e colaboradores (2006), onde a expressão dos genes *HMGB2*, *HMGB3* e *HMGB4* foi *up* regulada em resposta ao frio, enquanto que, a expressão de *HMGB2* e *HMGB3* foi marcadamente *down* regulada por seca e salinidade.

Apesar de terem sido identificados muitos *cis* elementos comuns em todos os transcritos, os modelos encontrados apresentaram apenas de três a seis elementos compostos em sua maioria pela combinação entre os *cis*-elementos TBP, AHBP e GTBX. O *cis* elemento GTBX (tgtagaGTTAttgtttt) é um GT *Box* com domínio de ligação bZip relacionado com respostas defesa e via de sinalização de ácido salicílico, respostas à luz, além do controle positivo/negativo da transcrição. O *cis* elemento AHBP (tgaATAAttgt) é identificado por FTs HDZip, e está relacionado com via de sinalização mediada por ABA. Enquanto TBP (tcaaTATAtatattt) é um fator de ligação TATA *Box*.

A dificuldade no desenho de modelos de promotores com mais *cis* elementos deve-se principalmente à posição em que os elementos se encontravam na sequência em relação ao TSS, que foi muito variável, e também provavelmente à grande variedade de famílias gênicas que foram diferencialmente expressas em respostas as condições de déficit hídrico impostas, que possivelmente ativaram a expressão de genes de respostas iniciais (T25-50 min) e genes de respostas mais tardias (T75-100 min, T125-150 min e progressivo).

O número/quantidade e a localização dos *cis* elementos na sequência promotora são importantes parâmetros para a determinação da sua função. Estudos demonstraram que a presença de um elemento GT no promotor induziu a expressão transiente de PMEC em tabaco enquanto que a presença de dois elementos levou a repressão do mesmo gene (Mehrotra et al., 2005; Mehrotra e Panwar, 2009). O espaçamento entre os *cis*-elementos é relevante pois determina as interações entre os TFs que se ligam a esses elementos, e podem em alguns casos controlar a natureza da indução do promotor.

Experimentos conduzidos com promotores sintéticos construídos com múltiplos elementos ACGT em diferentes espaçamentos revelaram que esses motivos quando separados por 5 pb apresentam indução específica pelo ácido salicílico e parcial pelo ABA, enquanto que, quando separados por 25 pb, o promotor foi unicamente induzido por ABA (Mehrotra e Mehrotra, 2010). É sabido também que a combinação de diferentes *cis* elementos resulta no controle transcricional dos genes, como mostrado por Tran e colaboradores (2007) que descreveram a atividade coordenada do GT Box e da sequência de reconhecimento MYC, ambas necessárias para a ativação de resposta por desidratação do promotor *erd1*.

## 5.6. Considerações finais

De modo geral, no déficit hídrico, fatores como a duração e a severidade da seca, além do estágio fenológico vegetal, influenciam o padrão das respostas moleculares, bioquímicas e fisiológicas permitindo que a planta evite ou tolere o estresse. Os mecanismos de tolerância envolvem basicamente o ajustamento osmótico e a remoção de espécies reativas de oxigênio formadas como consequência do metabolismo perturbado. A análise da categorização funcional dos transcritos diferencialmente expressos em ambos os estresses hídricos avaliados neste estudo, evidenciam a ativação de mecanismos de tolerância à seca.

Considerando que as plantas deste estudo encontravam-se no mesmo estágio fenológico quando o déficit hídrico foi imposto, esse fator possivelmente não influenciou significativamente a expressão gênica dos transcritos, entretanto a duração (déficit hídrico prolongado ou de curto período) e a severidade do tratamento, associado com os estresses de calor e oxidativo podem ter ativado diferentes vias de sinalização resultando nas respostas/mecanismos de defesa das plantas.

Cabe ressaltar que as metodologias utilizadas para indução do déficit hídrico foram distintas, bem como a técnica de detecção de transcritos expressos. Na técnica hidropônica o tratamento foi de curta duração e considerado rápido e intenso, enquanto que as plantas cultivadas no substrato sofreram um déficit progressivo, com evaporação lenta e gradual da água do solo. Entretanto, transcritos diferencialmente expressos comuns aos dois tratamentos (déficit de período curto e prolongado) foram identificados independentemente da metodologia de imposição utilizada ou da técnica de sequenciamento (Microarranjos de DNA ou RNASeq), reforçando o envolvimento destes nas respostas moleculares de ativação de mecanismos de aclimação e defesa da planta ao déficit hídrico.

A quantidade e variedade de *cis*-elementos relacionados a estresses abióticos e especificamente ao déficit hídrico identificados nos TDEs, sugere que um grande número de fatores de transcrição estão envolvidos na regulação da expressão, e aponta para um sistema complexo e cooperativo de controle da expressão gênica. O controle combinatorial refere-se à forma como

grupos de proteínas trabalham juntos para determinar a expressão de um único gene. Infere-se que nos TDEs analisados ocorre uma regulação combinatória, essencial para que a expressão de um determinado gene ocorra na célula correta, em resposta a uma dada condição, no tempo certo e no nível requerido.

Modelos de promotores contendo os *cis*-elementos AHBP e GTBX foram observados em todos os TDEs estudados, sugerindo que a presença combinada desses elementos pode estar relacionada com o controle transcricional de genes envolvidos com a resposta ao déficit hídrico. O desafio é entender como as diferentes permutações dos mesmos fatores de ligação de DNA podem alterar os padrões de expressão gênica.

## 5.7. Referências

AGARWAL, P. K.; JHA, B. Transcription factors in plants and ABA dependent and independent abiotic stress signaling. *Biologia Plantarum*, 54:201-212; 2010.

ASHBURNER, M.; BALL, C. A.; BLAKE, J. A.; BOTSTEIN, D.; BUTLER, H.; CHERRY, J. M.; SHERLOCK, G. Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25-29, 2000.

BALDERAS-HERNÁNDEZ, V. E.; ALVARADO-RODRÍGUEZ, M.; FRAIRE-VELÁZQUEZ, S. Conserved versatile master regulators in signaling pathways in response to stress in Plants, *AoB plants*, v. 5, 2013.

BANSAL, K. C.; LENKA, S. K.; TUTEJA, N. Abscisic Acid in Abiotic Stress Tolerance: An 'Omics' Approach; *Omics and Plant Abiotic Stress Tolerance*; 143-150, 2011.

BENJAMINI, Y.; HOCHBERG, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*, 289-300, 1995.

BERTOLLI, S. C.; RAPCHAN, G. L.; SOUZA, G. M. Photosynthetic limitations caused by different rates of water-deficit induction in *Glycine max* and *Vigna unguiculata*. *Photosynthetica*. DOI: 10. 1007/s11099-012-0036-4, 2012.

BURBIDGE, A.; GRIEVE, T. M.; JACKSON, A.; THOMPSON, A.; MCCARTY, D. R.; TAYLOR, I. B. Characterization of the ABA-deficient tomato mutant *notabilis* and its relationship with maize Vp14. *The Plant Journal* 17(4):427–431, 1999.

CARTHARIUS, K.; FRECH, K.; GROTE, K.; KLOCKE, B.; HALTMEIER, M.; KLINGENHOFF, A.; FRISCH, M.; BAYERLEIN, M.; WERNER, T. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics*, 21:2933-42, 2005.

CHAVES, M. M.; FLEXAS, J.; PINHEIRO, C. Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. *Annals of Botany*, 103(4):551-560, 2009.

CHERNYS, J. T.; ZEEVAART, J. A. D. Characterization of the 9-Cis-Epoxycarotenoid Dioxygenase Gene Family and the Regulation of Abscisic Acid Biosynthesis in Avocado. *Plant Physiology*, 124:343–353, 2000.

COUÉE, I.; SULMON, C.; GOUESBET, G.; EL AMRANI, A. Involvement of soluble sugars in reactive oxygen species balance and responses to oxidative stress in plants. *Journal of Experimental Botany*, 57:449–459, 2006.

COVINGTON, M. F.; HARMER, S. L. The circadian clock regulates auxin signaling and responses in *Arabidopsis*. *PLoS Biology*, 5(8):222, 2007.

CRAMER, G. R.; URANO, K.; DELROT, S.; PEZZOTTI, M.; SHINOZAKI, K. Effects of abiotic stress on plants: a systems biology perspective. *BMC Plant Biology*, 11:1471-2229, 2011.

CUELLAR-ORTIZ, S. M.; ARRIETA-MONTIE, M. DE L. P.; ACOSTA-GALLEGOS, J.; COVARRUBIA, A. A. Relationship between carbohydrate partitioning and drought resistance in common beans. *Plant Cell & Environment* 31(10):1399-1409, 2008.

FEHR, W. R.; CAVINESS, C. E.; BURMOOD, D. T.; PERNNIGTON, J. S. Stage of development description for soybeans [*Glycine Max (L.) Merrill*]. *Crop Science*, 11:929-931, 1971.

FUJITA, Y.; FUJITA, M.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. ABA-mediated transcriptional regulation in response to osmotic stress in plants. *Journal of Plant Research*, 124:509-525, 2011.

GILMARTIN, P. M.; CHUA, N. H. Spacing between GT-1 binding sites within a light-responsive element is critical for transcriptional activity. *Plant Cell* 2:447-455, 1990.

HARB, A.; KRISHNAN, A.; AMBAVARAM, M. M.; PEREIRA, A. Molecular and physiological analysis of drought stress in *Arabidopsis* reveals early responses leading to acclimation in plant growth. *Plant Physiology*, 154(3):1254-1271, 2010.

HARDCASTLE, T. J.; KELLY, K. A. BaySeq: empirical Bayesian methods for identifying differential expression in sequence count data. *BMC Bioinformatics*, 11(1):422, 2010.

HERNANDEZ-GARCIA, C. M.; BOUCHARD, R. A.; RUSHTON, P. J.; JONES, M. L.; CHEN, X.; TIMKO, M. P.; FINER, J. J. High level transgenic expression of soybean (*Glycine max*) GmERF and Gmubi gene promoters isolated by a novel promoter analysis pipeline. *BMC Plant Biology*, 10(1):p237, 2010.

IUCHI, S.; KOBAYASHI, M.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. A stress-inducible gene for 9-cis-epoxycarotenoid dioxygenase involved in abscisic acid biosynthesis under water stress in drought tolerant cowpea. *Plant Physiology*, 123:553-562, 2000.

KATO, H.; XIE, G.; SATO, Y.; IMAI, R. Isolation of anther-specific gene promoters suitable for transgene expression in rice. *Plant Molecular Biology Reporter*, 28:381–387, 2010.

KONDRÁK, M.; MARINCS, F.; ANTAL, F.; JUHÁSZ, Z.; BÁNFALVI, Z. Effects of yeast trehalose-6-phosphate synthase 1 on gene expression and carbohydrate contents of potato leaves under drought stress conditions. *BMC Plant Biology*, 12(1):74, 2012.

KRASENSKY, J.; JONAK, C. Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. *Journal of Experimental Botany*, 63:1593–1608, 2012.

KWAK, K. J.; KIM, J. Y.; KIM, Y.O.; KANG, H. Characterization of Transgenic *Arabidopsis* Plants Overexpressing High Mobility Group B Proteins under High Salinity, Drought or Cold Stress. *Plant Cell Physiology*, 48(2):221-231, 2006.

LEGNAIOLI, T.; CUEVAS, J.; MAS, P. TOC1 functions as a molecular switch connecting the circadian clock with plant responses to drought. *The EMBO Journal*. 28(23):3745–3757, 2009.

LIU, Y.; JI, X.; ZHENG, L.; NIE, X.; WANG, Y. Microarray Analysis of Transcriptional Responses to Abscisic Acid and Salt Stress in *Arabidopsis thaliana*. *International Journal of Molecular Sciences*, 14(5):9979-9998, 2013.

LU, M.; YING, S.; ZHANG, D. F.; SHI, Y. S.; SONG, Y. C.; WANG, T. Y.; LI, Y. A maize stress-responsive NAC transcription factor, ZmSNAC1, confers enhanced tolerance to dehydration in transgenic *Arabidopsis*. *Plant Cell Reporter*. 31(9):1701-1711, 2012.

MARUYAMA, K.; TODAKA, D.; MIZOI, J.; YOHISDA, T.; KIDOKORO, S.; MATSUKURA, S.; TAKASAKI, H.; SAKURAI, T.; YAMAMOTO, Y. Y.; YOSHIWARA, K.; KOJIMA, M.; SAKAKIBARA, H.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Identification of cisacting promoter elements in cold and dehydration-induced transcriptional pathways in *Arabidopsis*, rice, and soybean. *DNA Research*. 19:37–49, 2012.

MASTROENI, P.; JANIS, C.; GRANT, A. J.; MCKINLEY, T. J.; MORGAN, F. J. E.; JOHN, V. F.; DOUGAN, G. In vivo regulation of the VI antigen in salmonella and induction of immune responses with an in vivo-inducible promoter. *Infection and Immunity*, 79(6):2481-2488, 2011.

MEHROTRA, R.; PANWAR, J. Dimerization of GT element interferes negatively with gene activation, *Journal of Genetics*, 88(2):257–260, 2009.

MEHROTRA, R.; KIRAN, K.; CHATURVEDI ET AL, C. P. Effect of copy number and spacing of the ACGT and GT cis elements on transient expression of minimal promoter in plants. *Journal of Genetics*, 84(2):183–187, 2005.

MIZOI, J.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. AP2/ERF Family transcription factors in plant abiotic stress responses. *Biochimica et Biophysica Acta*, doi:10.1016/j.bbagr.2011.08.004, 2012

MORTAZAVI, A.; WILLIAMS, B. A.; MCCUE, K.; SCHAEFFER, L.; WOLD, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, 5(7):621-628, 2008.

MULLER, B.; PANTIN, F.; GÉNARD, M.; TURC, O. et al. Water deficits uncouple growth from photosynthesis, increase C content, and modify the relationships between C and growth in sink organs. *Journal of Experimental Botany*, 62:1715–1729, 2011.

NAKASHIMA, K.; TAKASAKI, H.; MIZOI, J.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. NAC transcription factors in plant abiotic stress responses. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, 1819(2), 97-103, 2012.

NAKASHIMA, K.; YAMAGUCHI-SHINOZAKI, K. Molecular Studies on Stress-Responsive Gene Expression in *Arabidopsis* and Improvement of Stress Tolerance in Crop Plants by Regulon Biotechnology. *Japan Agricultural Research Quarterly*, 39(4):221, 2005.

NOH, S. A.; LEE, H. S.; HUH, G. H.; OH, M. J.; PAEK, K. H.; SHIN, J. S.; BAE, J. M. A sweet potato SRD1 promoter confers strong root taproot and tuber specific expression in *Arabidopsis*, carrot and potato. *Transgenic Research*, 21:265–278, 2012.

OLIVEROS, J. C. VENNY. An interactive tool for comparing lists with Venn Diagrams. <http://bioinfogp.cnb.csic.es/tools/venny/index.html>, 2007.

OLVERA, C. Y.; CAMPOS, F.; LUIS, R. J.; GARCIARRUBIO, A.; COVARRUBIAS, A. A. Functional analysis of the group 4 late embryogenesis abundant proteins reveals their relevance in the adaptive response during water deficit in *Arabidopsis*. *Plant Physiology*, 154:373-390, 2010.

OYA, T.; NEPOMUCENO, A. L.; NEUMAIER, N.; FARIAS, J. R. B.; TOBITA, S.; ITO, O. Drought tolerance characteristics of brazilian soybean cultivars. *Plant Production Science* 7:129-137, 2004.

PETRONI, K.; KUMIMOTO, R. W.; GNESUTTA, N.; CALVENZANI, V.; FORNARI, M.; TONELLI, C.,...MANTOVANI, R. The promiscuous life of plant nuclear factor  $\gamma$  transcription factors. *The Plant Cell Online*, 24(12):4777-4792, 2012.

PINHEIRO, C.; CHAVES, M. M. Photosynthesis and drought: can we make metabolic connections from available data? *Journal of Experimental Botany*, 62(3):869-882, 2011.

QUACKENBUSH, J. Microarray data normalization and transformation. *Nature Genetics*, 32:496-501, 2002.

QIN, X.; ZEEVAART, J. A. D. The 9-cis-epoxycarotenoid cleavage reaction is the key regulatory step of abscisic acid biosynthesis in water-stressed bean. *Proceedings of the National Academy of sciences*, 96(26):15354-15361, 1999.

RAMEGOWDA, V.; SENTHIL-KUMAR, M.; NATARAJA, K. N.; REDDY, M. K.; MYSORE, K. S.; UDAYAKUMAR, M. Expression of a Finger Millet Transcription Factor, EcNAC1, in Tobacco Confers Abiotic Stress-Tolerance. *PLoS ONE* 7(7):e40397, 2012.

REGUERA, M.; PELEG, Z.; BLUMWALD, E. Targeting metabolic pathways for genetic engineering abiotic stress-tolerance in crops. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, 1819(2):186-194, 2012.

ROBINSON, M. D.; MCCARTHY, D. J.; SMYTH, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1):139-140, 2010.

RODRIGO, M-J.; ALQUEZAR, B.; ZACARIAS, L. Cloning and characterization of two 9-cis-epoxycarotenoid dioxygenase genes, differentially regulated during fruit maturation and under stress conditions, from orange (*Citrus sinensis* L. Osbeck). *Journal of Experimental Botany*, 57(3):633-643, 2006.

RODRIGUES, F. A.; MARCOLINO-GOMES, J.; CARVALHO, J. F. C.; NASCIMENTO, L. C.; NEUMAIER, N.; FARIAS, J. R. B.; CARAZZOLLE, M. F.; MARCELINO, F. C.; NEPOMUCENO, A. L. Subtractive libraries for prospecting

differentially expressed genes in the soybean under water deficit. *Genetics and Molecular Biology*, 35(10):304-314, 2012.

RODRÍGUEZ, M.; CANALES, E.; BORRÁS-HIDALGO, O. Molecular aspects of abiotic stress in plants. *Biotechnologia Aplicada*, La Habana, 22:1-10, 2005.

ROLLAND, F.; BAENA-GONZALEZ, E.; SHEEN J. Sugar sensing and signaling in plants: conserved and novel mechanisms. *Annual Review of Plant Biology*, 57: 675–709, 2006.

ROSA, M.; PRADO, C. PODAZZA, G.; INTERDONATO, R.; GONZÁLEZ, J. A., HILAL, M.; PRADO, F. E. Soluble sugars—Metabolism, sensing and abiotic stress: A complex network in the life of plants. *Plant Signaling & Behavior*, Bethesda, 4(5):388-393, 2009.

SAINT PIERRE, C.; CROSSA, J. L.; BONNETT, D.; YAMAGUCHI-SHINOZAKI, K.; REYNOLDS, M. P. Phenotyping transgenic wheat for drought resistance. *Journal of experimental botany*, 63(5):1799-1808, 2012.

SANCHEZ, A.; SHIN, J.; DAVI, S. J. Abiotic stress and the plant circadian clock *Plant Signaling & Behavior*, 6:2 223-231, 2011.

SCHNEIDERREIT, A.; IMLAU, A.; SAUER, N. Conserved cis-regulatory elements for DNA-binding-with-one-finger and homeo-domain-leucine-zipper transcription factors regulate companion cell-specific expression of the *Arabidopsis thaliana* SUCROSE TRANSPORTER 2 gene. *Planta*, 228:651-62, 2008.

SCHWARZ-SOMMER, Z.; HUIJSER, P.; NACKEN, W.; SAEDLER, H.; SOMMER, H. Genetic control of flower development by homeotic genes in *Antirrhinum Majus*. *Science*, 250:931-936, 1990.

SEKI, M.; UMEZAWA, T.; URANO, K.; SHINOZAKI, K. Regulatory metabolic networks in drought stress responses. *Current Opinion in Plant Biology*, 10(3):296-302, 2007.

SEKI, M.; NARUSAKA, M.; ISHIDA, J.; NANJO, T.; FUJITA, M.; OONO, Y.; KAMIYA, A.; NAKAJIMA, M.; ENJU, A.; SAKURAI, T.; SATOU, M.; AKIYAMA,

K.; TAJI, T.; YAMAGUCHI-SHINOZAKI, K.; CARNINCI, P.; KAWAI, J.; HAYASHIZAKI, Y.; SHINOZAKI, K. Monitoring the expression profiles of 7,000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. *Plant Journal*, 31:279-292, 2002.

SHAO, H. B.; GUO, Q. J.; CHU, L. Y.; ZHAO, X. N.; SU, Z. L.; HU, Y. C.; CHENG, J. F. Understanding molecular mechanism of higher plant plasticity under abiotic stress. *Colloids and Surfaces B: Biointerfaces*, 54(1):37-45, 2007.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Gene networks involved in drought stress response and tolerance. *Journal of Experimental Botany*, 58(2): 221-227, 2007.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Current Opinion in Plant Biology*, 3:217-223, 2000.

SIMPSON, S. D.; NAKASHIMA, K.; NARUSAKA, Y.; SEKI, M.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Two different novel cis-acting elements of *erd1*, a *clpA* homologous Arabidopsis gene function in induction by dehydration stress and dark-induced senescence. *Plant Journal*, 33:259-270, 2003.

SVENSSON, M. Evolution of a family of plant genes with regulatory functions in development; studies on *Picea abies* and *Lycopodium annotinum*. Doctoral thesis. Acta Universitatis Upsaliensis, Department of Evolutionary Biology. ISBN 91-554-4826-7, 2000.

TAIZ, L.; ZEIGER, E. *Fisiologia vegetal*. 4.ed. revista. e ampliada. Porto Alegre: Artmed, 739-774, 2010.

TITTARELLI, A.; SANTIAGO, M.; MORALES, A.; MEISEL, L.A.; SILVA, H. Isolation and functional characterization of cold-regulated promoters, by digitally identifying peach fruit cold-induced genes from a large EST dataset. *BMC Plant Biology*. 9:121, 2009.

TWYMAN, R. M. Growth and development: Control of gene expression, regulation of transcription. Encyclopedia of Applied Plant Sciences. London UK: Elsevier Science, 558-567, 2003.

UMEZAWA, T. Systems biology approaches to abscisic acid signaling. Journal of Plant Research, 124:539-548, 2011.

UMEZAWA, T.; FUJITA, M.; FUJITA, Y.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future. Current Opinion in Biotechnology 17:113-122, 2006.

USADEL, B.; NAGEL, A.; THIMM, O.; REDESTIG, H.; BLAESING, O. E.; PALACIOS-ROJAS, N.; SELBIG, J.; HANNEMANN, J.; PIQUES, M. C.; STEINHAUSER, D.; SCHEIBLE, W. R.; GIBON, Y.; MORCUENDE, R.; WEICHT, D.; MEYER, S.; STITT, M. Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. Plant Physiology, 138:1195-204, 2005.

VALERIO, C.; COSTA, A.; MARRI, L.; ISSAKIDIS-BOURGUET, E.; PUPILLO, P.; TROST, P.; SPARLA, F. Thioredoxin-regulated beta-amylase (BAM1) triggers diurnal starch degradation in guard cells, and in mesophyll cells under osmotic stress. Journal of Experimental Botany, 62:545–555, 2011.

VENU, R. C.; SREEREKHA, M. V.; MADHAV, M. S.; NOBUTA, K.; MOHAN, K. M.; CHEN, S.;...WANG, G. L. Deep transcriptome sequencing reveals the expression of key functional and regulatory genes involved in the abiotic stress signaling pathways in rice. Journal of Plant Biology, 56(4):216-231, 2013.

WAN, X.; LI, L. Molecular cloning and characterization of a dehydrated inducible cDNA encoding a putative 9-cis- epoxy carotenoid dioxygenase in *Arachis Hypogaeae* L., Mitochondrial DNA. 16:217-223, 2005.

WANG, C. T.; YANG, Q. Y.; YANG, Y. M. Characterization of the ZmDBP4 gene encoding a CRT/DRE-binding protein responsive to drought and cold stress in maize. Acta Physiologiae Plantarum, 33(2):575-583, 2011.

WANG, L.; LI, L.; XU, L.; ZHOU, J.; ZHUANG, H.; GONG, X.; WANG, M.; SUN, S. S. M.; ZHUGE, Q. Isolation and functional analysis of the Poplar RbcS gene promoter. *Plant Molecular Biology Reporter*, 31:120–127, 2013.

WILKINS, O.; BRÄUTIGAM, K.; CAMPBELL, M. M. Time of day shapes *Arabidopsis* drought transcriptomes. *Plant Journal*, 63:715–727, 2010.

WOLF, E.; AIGNER, B.; KLYMIUK, N. Transgenic pigs for xenotransplantation: selection of promoter sequences for reliable transgene expression. *Current Opinion in Organ Transplantation*, 15(2):201-206, 2010.

XIAO, X.; FAN, Y.; SHENG, Z.; HELENA, K.; CHUNYANG, L. Physiological and proteomic responses of two contrasting *Populus cathayana* populations to drought stress. *Physiologia Plantarum*, 136:150–168, 2009.

XU, L.; YE, R.; ZHENG, Y.; WANG, Z.; ZHOU, P.; LIN, Y.; LI, D. Isolation of the endosperm-specific LPAAT gene promoter from coconut (*Cocosnucifera* L.) and its functional analysis in transgenic rice plants. *Plant Cell Reports*, 29:1061–1068, 2010.

XU, Z. S.; CHEN, M.; LI, L. C.; MA, Y. Z. Functions of the ERF transcription factor family in plants. *Botany* 865:969–977, 2008.

YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annual. Review. Plant Biology*, 57:781–803, 2006.

YAMAMOTO, Y. Y.; YOSHIOKA, Y.; HYAKUMACHI, M., MARUYAMA, K.; YAMAGUCHI-SHINOZAKI, K.; TOKIZAWA, M.; KOYAMA, H. Prediction of transcriptional regulatory elements for plant hormone responses based on microarray data *BMC Plant Biology*, 11:39, 2011.

YI, N.; KIM, Y. S.; JEONG, M. H.; OH, S.J.; JEONG, J. S.; PARK, S. H.; JUNG, H.; CHOI, Y.D.; KIM, J. K. Functional analysis of six drought-inducible promoters in transgenic rice plants throughout all stages of plant growth. *Planta*, 232:743–754, 2010.

ZHANG, C.; WANG, J.; HUA, X.; FANG, J.; ZHU, H.; GAO, X. A mutation degree model for the identification of transcriptional regulatory elements. *BMC Bioinformatics*, 12:262, 2011.

ZHENG, X.; CHEN, B.; LU, G.; HAN, B. Overexpression of a NAC transcription factor enhances rice drought and salt tolerance. *Biochemical and Biophysical Research Communications*, 379(4):985-989, 2009.

ZHU, J. K. Salt and drought stress signal transduction in plants. *Annual Review. Plant Biology*, 53:247-273, 2002.

ZOU, C.; SUN, K.; MACKALUSO, J. D.; SEDDON, A. E.; JIN, R.; THOMASHOW, M. F.; SHIU, S. H. Cis-regulatory code of stress-responsive transcription in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences*, 108(36):14992-14997, 2011.

.

## 5.8. Material Suplementar

**Tabela 1.** Transcritos diferencialmente expressos comuns ao déficit hídrico progressivo e os três tempos do déficit hídrico de curto período.

Identificação	T25/50min	T75/100min	T125/150min	Prolongado	Descrição
Glyma01g19420.1	1,791	2	2	1,50	unknown protein
Glyma04g40700.1	3,345	5	5	1,40	unknown protein
Glyma05g08240.1	2,392	5	5	1,41	unknown protein
Glyma06g10580.1	1,718	3	3	1,01	unknown protein
Glyma06g14090.1	4,175	5	4	1,70	unknown protein
Glyma09g31740.1	3,897	9	11	7,69	unknown protein
Glyma11g12990.1	3,438	5	6	1,02	unknown protein
Glyma11g13290.1	2,109	2	2	1,12	unknown protein
Glyma11g36030.1	4,506	7	7	1,45	unknown protein
Glyma12g05080.1	2,049	3	3	1,48	unknown protein
Glyma12g05320.1	3,694	4	5	2,64	unknown protein
Glyma13g21580.1	4,897	7	6	2,64	unknown protein
Glyma13g22470.1	1,892	4	4	1,56	unknown protein
Glyma14g39950.1	3,245	4	5	1,37	unknown protein
Glyma17g11270.1	1,401	3	2	1,17	unknown protein
Glyma17g13720.1	5,438	8	8	7,62	unknown protein
Glyma20g26140.1	4,268	5	5	1,82	unknown protein
Glyma20g29770.1	3,940	8	10	6,94	unknown protein
Glyma13g31930.1	3,732	5	4	2,47	unknown protein
Glyma14g11150.1	1,737	4	5	1,78	unknown protein
Glyma08g02130.1	1,463	1	1	1,40	1-aminocyclopropane-1-carboxylic acid (acc) synthase 6
Glyma16g08470.1	2,234	3	3	2,38	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
Glyma02g13120.1	4,376	5	5	4,73	ABI five binding protein 3

Glyma04g04170.1	1,672	4	4	3,17	abscisic acid responsive elements-binding factor 2
Glyma18g50240.1	2,035	4	4	1,45	AFP3 (ABI FIVE BINDING PROTEIN 3)
Glyma07g09860.1	1,201	2	2	1,10	alpha/beta-Hydrolases superfamily protein
Glyma08g06110.1	3,477	6	6	8,19	alpha/beta-Hydrolases superfamily protein
Glyma10g01770.1	3,791	5	6	2,93	alpha/beta-Hydrolases superfamily protein
Glyma20g16910.1	1,616	2	3	1,69	AP2 domain-containing transcription factor, putative
Glyma06g00990.1	1,386	1	2	4,00	arginine decarboxylase 2
Glyma02g03890.1	5,101	5	4	4,05	ARM repeat superfamily protein
Glyma17g09850.1	3,116	4	4	3,83	ARM repeat superfamily protein
Glyma18g04770.1	3,123	3	4	2,26	ARM repeat superfamily protein
Glyma13g43800.1	1,264	1	2	1,39	AUX/IAA transcriptional regulator family protein
Glyma13g28630.1	3,553	4	4	1,61	beta-amylase 3
Glyma15g10480.1	2,226	3	4	2,28	beta-amylase 3
Glyma13g19560.1	3,940	4	4	2,87	C2H2 and C2HC zinc fingers superfamily protein
Glyma10g41840.1	3,023	4	4	1,11	Core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein
Glyma08g33510.1	2,197	5	5	2,19	CTP synthase family protein
Glyma11g33450.1	3,342	3	4	2,85	CYS, MET, PRO, and GLY protein 1
Glyma14g39300.1	3,175	3	3	1,07	CYS, MET, PRO, and GLY protein 1
Glyma13g27840.1	1,395	2	2	2,51	Eukaryotic aspartyl protease family protein
Glyma14g39350.1	3,376	5	4	3,70	Eukaryotic aspartyl protease family protein
Glyma08g07590.1	1,605	4	4	2,85	EXORDIUM like 2
Glyma10g04640.1	1,957	3	2	1,20	Exostosin family protein
Glyma13g43730.1	4,897	6	6	2,17	Galactose oxidase/kelch repeat superfamily protein
Glyma16g03190.1	1,259	2	2	2,12	G-box binding factor 3
Glyma18g06840.1	2,184	6	7	5,65	glutamine-dependent asparagine synthase 1
Glyma03g02230.1	1,355	2	3	1,16	glutaredoxin-related
Glyma01g39260.1	1,503	3	3	1,15	heat shock factor 4
Glyma04g05500.1	2,223	4	2	3,19	heat shock transcription factor A2

Glyma07g05800.1	4,023	7	8	7,96	homeobox 7
Glyma02g40800.1	3,138	4	2	4,80	Homeodomain-like superfamily protein
Glyma18g05050.1	3,312	4	5	2,55	Homeodomain-like superfamily protein
Glyma13g41460.1	1,983	3	2	2,24	MATE efflux family protein
Glyma15g03930.1	1,352	2	3	2,87	MATE efflux family protein
Glyma14g37500.1	3,746	6	6	3,49	mitogen-activated protein kinase kinase kinase 14
Glyma04g38560.1	2,231	3	4	1,22	NAC domain transcriptional regulator superfamily protein
Glyma06g38410.1	2,541	6	6	5,87	NAC domain transcriptional regulator superfamily protein
Glyma12g35000.1	2,553	5	6	3,69	NAC domain containing protein 3
Glyma13g35550.1	2,675	5	6	4,15	NAC domain containing protein 3
Glyma06g11970.1	1,246	2	2	1,91	NAC domain transcriptional regulator superfamily protein
Glyma12g22880.1	3,023	7	8	6,48	NAC domain transcriptional regulator superfamily protein
Glyma01g01400.1	1,487	1	1	1,24	NB-ARC domain-containing disease resistance protein
Glyma08g18800.1	3,865	5	6	4,79	NCED3 9-cis-epoxycarotenoid dioxygenase
Glyma15g40070.1	4,996	7	7	4,72	nine-cis-epoxycarotenoid dioxygenase 3
Glyma08g17630.1	1,831	3	2	1,15	nuclear factor Y, subunit C2
Glyma11g33040.1	4,180	5	6	2,89	oxidative stress 3
Glyma10g35550.1	1,689	4	4	2,04	PATATIN-like protein 6
Glyma05g30190.1	1,853	2	2	1,59	phospholipase D delta
Glyma02g11480.1	1,676	3	3	1,27	plant U-box 26
Glyma16g06410.1	1,736	2	1	1,16	P-loop containing nucleoside triphosphate hydrolases superfamily protein
Glyma07g32020.1	1,438	2	2	2,21	polyubiquitin 10
Glyma13g17820.1	1,129	2	2	2,10	polyubiquitin 10
Glyma13g24500.1	1,050	2	1	1,41	polyubiquitin 10
Glyma08g07820.1	1,416	2	2	2,66	predicted protein [Populus trichocarpa]
Glyma02g04000.1	3,312	4	4	2,21	Protein of unknown function (DUF1645)
Glyma16g10000.1	3,107	4	4	4,21	Protein of unknown function (DUF506)
Glyma08g11000.1	1,518	3	2	1,81	Protein of unknown function (DUF579)

Glyma02g01210.1	2,381	3	2	1,07	Protein phosphatase 2C family protein
Glyma11g34410.1	2,835	4	5	4,64	protein phosphatase 2CA
Glyma05g08950.1	2,360	4	4	5,24	Raffinose synthase family protein
Glyma11g34470.1	2,318	2	2	4,41	Regulator of chromosome condensation (RCC1) family protein
Glyma13g01930.1	1,980	2	1	1,15	related to AP2 4
Glyma14g34590.1	3,157	3	3	2,42	related to AP2 4
Glyma19g32280.1	1,800	3	3	3,81	remorin family protein
Glyma11g37780.1	2,701	4	4	5,29	RING/U-box superfamily protein
Glyma18g01720.1	2,618	4	4	6,82	RING/U-box superfamily protein
Glyma04g04760.1	1,903	2	3	3,10	salt tolerance zinc finger
Glyma06g04840.1	2,359	2	3	3,09	salt tolerance zinc finger
Glyma02g46610.1	2,656	1	1	1,08	salt-inducible zinc finger 1
Glyma10g02210.1	1,413	2	2	1,11	senescence-associated gene 21
Glyma04g16880.1	1,537	3	3	1,23	sigma factor binding protein 1
Glyma06g13100.1	1,492	2	2	4,24	sigma factor E
Glyma15g16360.1	1,821	1	2	1,83	soybean gene regulated by cold-2
Glyma03g27860.1	1,507	2	2	1,16	transcription factor bZIP38 [Glycine max]
Glyma04g11330.1	1,940	2	1	4,22	Transducin/WD40 repeat-like superfamily protein
Glyma17g07530.1	2,397	5	6	3,65	trehalose phosphatase/synthase 11
Glyma09g06980.1	1,838	2	1	1,52	WRKY DNA-binding protein 11
Glyma11g14580.1	1,889	1	1	2,60	zinc finger (C3HC4-type RING finger) family protein
Glyma12g06460.1	2,625	3	3	2,46	zinc finger (C3HC4-type RING finger) family protein
Glyma10g40400.1	2,367	4	5	2,42	zinc-finger protein 1
Glyma11g31330.1	-3,732	2	3	3,72	GLUTATHIONE S-TRANSFERASE TAU 19
Glyma19g22590.1	-3,022	1	3	3,48	amino acid transmembrane transporter
Glyma11g10340.1	-1,031	2	2	2,20	Calcium-binding EF-hand family protein
Glyma15g05060.1	-4,147	1	1	2,01	Protein kinase superfamily protein
Glyma19g03010.1	-1,147	2	3	1,93	UDP-glucuronosyl/UDP-glucosyl transferase family protein

Glyma08g18200.1	-1,870	-4	-5	-6,85	unknown protein
Glyma09g17260.1	-2,211	-3	-2	-2,86	unknown protein
Glyma02g08970.1	-1,131	-1	-2	-1,44	3-oxo-5-alpha-steroid 4-dehydrogenase family protein
Glyma06g19720.1	-2,147	-3	-4	-4,41	AT5G14920 gibberellin-regulated family protein
Glyma08g26090.1	-2,380	-2	-4	-1,03	C2 calcium/lipid-binding plant phosphoribosyltransferase family protein
Glyma05g32190.1	-1,483	-2	-2	-2,00	carbamoyl phosphate synthetase A
Glyma11g05710.1	-1,198	-1	-1	-1,86	Cystathionine beta-synthase (CBS) family protein
Glyma09g31810.1	-1,203	-2	-1	-1,03	cytochrome P450, family 71, subfamily A, polypeptide 22
Glyma06g07270.1	-1,835	-2	-2	-1,26	DnaJ/Hsp40 cysteine-rich domain superfamily protein
Glyma06g04740.1	-3,167	-5	-7	-6,50	Gibberellin-regulated family protein
Glyma13g17950.1	-1,062	-2	-2	-1,44	HAD-superfamily hydrolase, subfamily IG, 5'-nucleotidase
Glyma18g01830.1	-1,003	-3	-5	-1,86	homeobox protein 6
Glyma08g01060.1	-1,005	-2	-2	-1,83	magnesium-protoporphyrin IX methyltransferase
Glyma15g42860.1	-1,163	-2	-2	-1,08	Mog1/PsbP/DUF1795-like photosystem II reaction center PsbP family protein
Glyma16g26900.1	-1,682	-2	-3	-2,26	NDH-dependent cyclic electron flow 5
Glyma06g44640.1	-3,469	-2	-6	-1,09	O-acyltransferase (WSD1-like) family protein
Glyma09g36640.1	-2,422	-4	-6	-5,34	Plant invertase/pectin methylesterase inhibitor superfamily protein
Glyma08g22290.1	-2,671	-4	-3	-1,20	Protein of unknown function (DUF2921)
Glyma16g29640.1	-2,324	-2	-3	-1,92	Protein of unknown function (DUF3464)
Glyma01g00500.1	-1,147	-3	-3	-1,88	Protein of unknown function (DUF581)
Glyma15g10150.1	-1,597	-3	-5	-3,07	Protein of unknown function (DUF642)
Glyma06g38160.1	-2,727	-3	-3	-1,12	protochlorophyllide oxidoreductase A
Glyma05g22180.1	-1,827	-2	-2	-3,69	root hair specific 19
Glyma06g04810.1	-1,519	-4	-4	-1,44	Subtilase family protein
Glyma01g09520.1	-1,842	-2	-2	-1,04	TCP-1/cpn60 chaperonin family protein
Glyma10g39740.1	-2,873	-4	-5	-3,74	thiazole biosynthetic enzyme, chloroplast (ARA6) (THI1) (THI4)
Glyma06g01750.1	-1,165	-2	-3	-1,02	tonoplast monosaccharide transporter2

**Tabela 2.** Identificação dos *cis* elementos putativos encontrados nos 134 transcritos diferencialmente expressos identificados em comum a ambos os experimentos de déficit hídrico progressivo e de curta duração: sua abundância, respectivas sequências, identificação de família e matriz.

ID	Match Total	Sequência	Identificação da família	Identificação da matrix
P\$AHBP	1564	tgaATAAttgt	Arabidopsis homeobox protein	HDZip class I protein ATHB5
P\$GTBX	1014	tgtagaGTTAttgtttt	GT-box elements	GT1-Box binding factors with a trihelix DNA-binding domain
O\$PTBP	823	tcaaTATAtatattt	Plant TATA binding protein factor	Plant TATA box
P\$MYBL	657	agtgtagAGTTattgtt	MYB-like proteins	Myb domain protein 96 (MYBCOV1)
P\$L1BX	585	attCATTaatttctttt	L1 box, motif for L1 layer-specific expression	Arabidopsis thaliana meristem layer 1
P\$CCAF	534	acaatcacAATCttc	Circadian control factors	Circadian clock associated 1
P\$MADS	480	tggtcccagAAATagtaactc	MADS box proteins	MADS-box protein SQUAMOSA
P\$DOFF	396	aagtatttAAAGataaa	DNA binding with one finger (DOF)	Prolamin box, conserved in cereal seed storage protein gene promoters
P\$NACF	393	gagctgcgaaaagggCACGtaccatca	Plant specific NAC [NAM-ATAF172, CUC2 ] transcription factors	Transcription factor of rice and barley binding to the iron deficiency-responsive cis-acting element 2 (IDE2)
P\$MYBS	368	gagaATATgatggaagt	MYB proteins with single DNA binding repeat	Phosphate starvation response 1
P\$GBOX	322	taattattACGTcaagtagat	Plant G-box/C-box bZIP proteins	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)
P\$HMGF	311	aataTATTttttttg	High mobility group factors	High mobility group I/Y-like proteins
P\$IBOX	288	ttaaaGATAaatgcatt	Plant I-Box sites	Class I GATA factors
P\$MYCL	286	aaaacACAAatgttcaaga	Myc-like basic helix-loop-helix binding factors	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)
P\$NCS1	284	tAAAAGttaa	Nodulin consensus sequence 1	Nodulin consensus sequence 1

P\$OPAQ	273	tcaagTGAGtcatgcat	Opaque-2 like transcriptional activators	GCN4, conserved in cereal seed storage protein gene promoters, similar to yeast GCN4 and vertebrate AP-1
P\$SUCB	271	atAAATaattattttaa	Sucrose box	Sequence motif from the promoters of different sugar-responsive genes
P\$LREM	209	agATCTaatca	Light responsive element motif, not modulated by different light qualities	RAP2.2, involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes
P\$SPF1	201	taTACTatagt	Sweet potato DNA-binding factor with two WRKY-domains	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT) sequences of sporamin and beta-amylase genes
P\$ASRC	187	tagTTGAg	AS1/AS2 repressor complex	AS1/AS2 repressor complex binding motif II
P\$ABRE	185	caaagACACgtagcttt	ABA response elements	ABA (abscisic acid) inducible transcriptional activator
P\$SWNS	173	aatttatttaAAAGaaag	Secondary wall NACS	Secondary wall NAC binding elements
P\$HEAT	163	acaatGTTCaagaaat	Heat shock factors	Arabidopsis thaliana class A heat shock factor 1a
P\$STKM	163	ataTAAAtaattat	Storekeeper motif	Storekeeper (STK), plant specific DNA binding protein important for tuber-specific and sucrose-inducible gene expression
P\$WBXF	157	tttaTTGAcAAatatt	W Box family	Calmodulin binding WRKY transcription factor 11
P\$GAPB	149	acgtATGAacacaag	GAP-Box (light response elements)	Cis-element in the GAPDH promoters conferring light inducibility
P\$CGCG	127	ataCGCGtatgatttga	Calmodulin binding / CGCG box binding proteins	Arabidopsis thaliana signal-responsive gene1, Ca <sup>2+</sup> / calmodulin binding protein homolog to NtER1 (tobacco early ethylene-responsive gene)
P\$SEF4	121	agTTTTtatta	Soybean embryo factor 4	Soybean embryo factor 4

O\$INRE	114	ttTCATtttct	Core promoter initiator elements	Drosophila initiator motifs
P\$SBPD	114	ggcacGTACcatcatca	SBP-domain proteins	SQUA promoter binding proteins
P\$CAAT	105	atCCAAtaa	CCAAT binding factors	CCAAT-box in plant promoters
P\$MIIG	103	ctttGGTTggtttgg	MYB IIG-type binding sites	Maize activator P of flavonoid biosynthetic genes
P\$PSRE	99	tcccaGAAAtagtaact	Pollen-specific regulatory elements	GAAA motif involved in pollen specific transcriptional activation
P\$IDDF	97	agttTTGTcatta	ID domain factors	Maize INDETERMINATE1 zinc finger protein
P\$CNAC	90	atTTGCTTgtttccaaaccg	Calcium regulated NAC-factors	Calmodulin-binding NAC protein
P\$MSAE	88	tccctAACGgaccag	M-phase-specific activator elements	M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)
P\$GAGA	87	aagaagAGAGaaaaatagttgataa	GAGA elements	Basic pentacysteine proteins
P\$PSPE	85	ttcGAAGaaaaaatt	Protein secretory pathway element	Cis-element involved in SA (salicylic acid) induction of secretion-related genes via NPR1
P\$DREB	83	tattgataCCGAcacttcggt	Dehydration responsive element binding factors	C-repeat/dehydration response element
P\$CARM	73	ctccccaAACActtcaaag	CA-rich motif	CA-rich element
P\$TEFB	68	taAAGGaaattatagtaaag	TEF-box	TEF cis acting elements in both RNA polymerase II-dependent promoters and rDNA spacer sequences
P\$EINL	63	aTGCAtctt	Ethylen insensitive 3 like factors	TEIL (tobacco EIN3-like)
P\$BRRE	61	tcctCGTGcaaataca	Brassinosteroid (BR) response element	Brassinazole-resistant 1
P\$TERE	61	cttaAAAGaaa	Tracheary-element-regulating cis-elements, conferring TE-specific expression	Tracheary-element-regulating cis-element
P\$LEGB	56	aagtgagtCATGcatgcatgcatgcat	Legumin Box family	RY and Sph motifs conserved in seed-specific promoters
P\$DPBF	54	cACACgcgctcc	Dc3 promoter binding factors	bZIP factors DPBF-1 and 2 (Dc3 promoter binding factor-1 and 2)

P\$ROOT	51	tgacaaacgcgtcagCACGtatgaa	Root hair-specific cis-elements in angiosperms	Root hair-specific element with a 2-nucleotid spacer between left part (LP) and right part (RP)
P\$NCS2	50	ttttatCTCTttctg	Nodulin consensus sequence 2	Nodulin consensus sequence 2
P\$TCPF	45	gcgggaCCCAcct	DNA-binding proteins with the plant specific TCP-domain	TCP class I transcription factor
P\$SALT	40	gatgggcGGTGgttt	Salt/drought responsive elements	Zinc-finger protein in alfalfa roots, regulates salt tolerance
P\$AREF	36	aagTGTCacactt	Auxin response element	Auxin Response Element
P\$ERSE	35	ccacttctcccgcgcCACG	ER stress-response elements	ERSE I (ER stress-response element I)-like motif
P\$LFYB	34	aTCCAgtgtt	LFY binding site	Plant specific floral meristem identity gene LEAFY (LFY)
P\$RAV5	33	gacAACAtaaa	5'-part of bipartite RAV1 binding site	5'-part of bipartite RAV1 binding site, interacting with AP2 domain
P\$AGP1	32	ggaGATCtaat	Plant GATA-type zinc finger protein	AG-motif binding protein 1
P\$SEF3	31	acatgACCCaaccca	Soybean embryo factor 3	SEF3, Soybean embryo factor 3
P\$CE3S	27	tgtcgcCGCGtgtcaaac	Coupling element 3 sequence	Coupling element 3 (CE3), non-ACGT ABRE
P\$MYCS	27	ttTCTTtttct	Cis-regulatory elements involved in the transcriptional activation of the arbuscular mycorrhizal (AM)-mediated inorganic phosphate transporter genes	Mycorrhiza transcription factor
P\$TDTF	26	aacaCACGcgtccgtcc	Transposase-derived transcription factors	FAR-RED ELONGATED HYPOCOTYLS 3 and FAR-RED IMPAIRED RESPONSE 1
P\$TELO	26	tctaaCCCTaatcct	Telo box (plant interstitial telomere motifs)	Ribosomal protein box, appears unique to plant RP genes and genes associated with gene expression
P\$GARP	26	AGATtctca	Myb-related DNA binding proteins (Golden2, ARR, Psr)	Type-B response regulator (ARR10), member of the GARP-family of plant myb-related DNA binding motifs

P\$CE1F	25	tttcCACctgtat	Coupling element 1 binding factors	Sugar and ABA responsive element conserved in several rbcS promoters
P\$E2FF	24	ctctTTCctctcta	E2F-homolog cell cycle regulators	E2F class I sites
O\$MTEN	23	caaatcgATCGaccgggtacc	Core promoter motif ten elements	Drosophila motif ten element
O\$TF2D	20	ctagtcctaatttcttttattccaaaattGACGttgaa	General transcription factor IID, GTF2D	Initiator (INR) and downstream promoter element (DPE) with strictly maintained spacing
P\$GCCF	19	aacttcCGCCata	GCC box family	Ethylene-responsive elements (ERE) and jasmonate- and elicitor-responsive elements (JERE)
P\$FORC	19	ttgtttGGGCcaaccct	Fungal and oomycete pathogen response cluster - promoter motif	Promoter motif enriched in a cluster of Arabidopsis genes co-expressed during immune response
P\$EREF	19	aGCGGaagcaccgaagaag	Ethylene response element factors	WRINKLED 1
P\$SLIM	18	tgaaaCATTgaatataaatac	Sulphur limitation, elements found in genes inducible during sulphur deprivation	Binding sites for AtSLIM1 and NtEIL2
P\$PALA	16	gtacattCCGTcagaaact	Conserved box A in PAL and 4CL gene promoters	Putative cis-acting element on various PAL and 4CL gene promoters
P\$PNRE	14	acgagtgtgttttaagtagggAAGAggtgaat	Plant nitrate-responsive cis-elements	Nitrate-responsive element
P\$PREM	14	tccgCGACTtcagccacctcttcagtagctt	Motifs of plastid response elements	Promoter elements involved in MgProto (Mg-protoporphyrin IX) and light-mediated induction
P\$URNA	14	gaggtcCCACagtgcag	Upstream sequence element of U-snRNA genes	Upstream sequence elements in the promoters of U-snRNA genes of higher plants
P\$VRES	14	aaacaGCTAtt	VIP1 responsive elements	VirE2-interacting protein 1
P\$NCS3	13	gaCACcctaac	Nodulin consensus sequence 3	Nodulin consensus sequence 3

O\$XCPE	12	gaGCGGgaccc	Activator-, mediator- and TBP-dependent core promoter element for RNA polymerase II transcription from TATA-less promoters	X gene core promoter element 1
P\$IDRS	8	actccacgtgaggtCCACcccgaatgca	Iron-dependent regulatory sequences	Iron-dependent regulatory sequence
P\$TALE	8	aaaGACAggttgt	TALE (3-aa acid loop extension) class homeodomain proteins	KNOTTED1 (KN1) and KNOTTED interacting protein (KIP) are TALE class homeodomain proteins. The KN1-KIP complex binds this DNA motif with high affinity.
P\$EPFF	8	aactaaagaggCAGTtaaaaaaa	EPF-type zinc finger factors, two canonical Cys2/His2 zinc finger motifs separated by spacers of various length	Member of the EPF family of zinc finger transcription factors
P\$SURE	8	taataAAAAaaaa	Sugar response elements, SURE and SURE-like sequences in plant promoters	Sugar responsive elements
P\$HOCT	5	aagagccATCCacgtgt	Octamer motif of Histone H3, H4 promoters	Octamer motif found in plant histone H3 and H4 genes
P\$CDC5	4	agcttAGCGtt	Arabidopsis CDC5 homolog	Myb related transcriptional activator, direct ortholog of S.pombe CDC5 protein
O\$TF3C	4	GGTTagagact	General transcription factor IIIC, GTF3C	RNA polymerase III transcription initiation factor complex (TFIIIC)
O\$TF2B	2	ccgCGCC	RNA polymerase II transcription factor II B	Transcription factor II B (TFIIB) recognition element

---