UNIVERSIDADE
ESTADUAL DE LONDRINA

JULIANA MARCOLINO GOMES

“IDENTIFICAÇÃO E ANÁLISE DA EXPRESSÃO DE GENES DE SOJA RESPONSIVOS AO DÉFICIT HÍDRICO E AO CICLO CIRCADIANO”

Londrina
2015

Universidade Estadual de Londrina
Centro de Ciências Biológicas
Departamento de Biologia Geral

Programa de
Pós-graduação em
Genética e Biologia Molecular

JULIANA MARCOLINO GOMES

“IDENTIFICAÇÃO E ANÁLISE DA EXPRESSÃO DE GENES DE SOJA RESPONSIVOS AO DÉFICIT HÍDRICO E AO CICLO CIRCADIANO”

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

Orientador: Dr. Alexandre Lima Nepomuceno

Co-orientadores:
Dr. Frank G. Harmon
Dra. Fabiana Aparecida Rodrigues

Londrina
2015

JULIANA MARCOLINO GOMES

“IDENTIFICAÇÃO E ANÁLISE DA EXPRESSÃO DE GENES DE SOJA RESPONSIVOS AO DÉFICIT HÍDRICO E AO CICLO CIRCADIANO”

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

BANCA EXAMINADORA

____
Prof° Dr. Alexandre Lima Nepomuceno
Empresa Brasileira de Pesquisa Agropecuária
EMBRAPA

____
Prof° Dr. Ricardo Vilela Abdelnoor
Empresa Brasileira de Pesquisa Agropecuária
EMBRAPA

____
Prof° Dr. Douglas Silva Domingues
Universidade Estadual Paulista "Júlio de
Mesquita Filho" - UNESP (Rio Claro)

____
Prof° Dr. Luiz Gonzaga Esteves Vieira
Universidade do Oeste Paulista - UNOESTE
(Presidente Prudente)

____
Dr. Hugo Bruno Correa Molinari
Empresa Brasileira de Pesquisa Agropecuária
EMBRAPA – Agroenergia

Londrina, março de 2015

Dedico este trabalho aos meus Pais, que me tornaram a pessoa que sou hoje; e ao meu Esposo, pela amizade e amor incondicionais.

AGRADECIMENTOS

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

Aos meus pais, os pilares da minha vida; meus exemplos de honestidade e perseverança; aqueles que me ensinaram a superar todos os desafios e me deram a certeza de que eu poderia ser qualquer coisa que eu sonhasse;

Ao meu esposo, Douglas, por ter me ajudado em todos os momentos da minha vida, mesmo nos mais difíceis; por ter suportado a distância e nunca ter me deixado desistir; por ter preenchendo minha vida com amor, alegria e esperança; por ter me ensinado a ser uma pessoa melhor a cada dia; sem você nada disso teria se tornado realidade;

Ao meu orientador, Dr. Alexandre Lima Nepomuceno, à quem eu admiro não só na vida acadêmica, mas como ser humano; por ter me ensinado O Caminho, e por ter estado sempre presente ao longo da Caminhada, me ajudando a superar meus próprios limites;

À minha co-orientadora Dr. Fabiana A. Rodrigues à Dr. Renata Fuganti-Pagliarini, por terem me disponibilizado seu tempo e atenção sempre que eu precisei;

Aos docentes e funcionários da Universidade Estadual de Londrina e do Programa de Pós-graduação em Genética e Biologia Molecular, principalmente à Dr. Ana Lúcia Dias, pelo apoio e incentivo que me acompanham desde a Graduação;

Ao *Plant Gene Expression Center* (PGEC) – USDA/ARS, onde tive a oportunidade de realizar Doutorado Sanduíche no Exterior (SWE), sob a co-orientação do Dr. Frank G. Harmon, sem o qual esse trabalho não teria sido possível;

Às instituições CNPq e CAPES pelo apoio financeiro à pesquisa e pela oportunidade de aprendizado;

À todos os funcionários da Embrapa Soja, pelo apoio e auxílio na realização desse trabalho e;

Aos meus amigos do Laboratório de Biotecnologia Vegetal – Embrapa Soja, por tantas vezes terem me ajudado a superar as dificuldades, compartilhado suas experiências e me feito sorrir mesmo quando nada parecia dar certo: vocês estão no meu coração.

“Dê o primeiro passo na fé. Você não tem que ver a escada toda, só precisa dar o primeiro passo.”

Martin Luther King

GOMES, Juliana Marcolino. **“Identificação e análise da expressão de genes de soja responsivos ao déficit hídrico e ao ciclo circadiano”.** 2015. Teses (Doutor em Genética e Biologia Molecular). Londrina, 2015.

RESUMO

Em todo o mundo, o déficit hídrico é o fator abiótico com mais impacto na produção de grandes culturas, como a soja. O ciclo circadiano desempenha um papel crítico na resposta e adaptação das plantas às diferentes condições ambientais. Recentemente, as oscilações circadianas foram associadas às respostas ao déficit hídrico, em nível transcricional. Todavia, pouco se sabe sobre a relação entre o ciclo circadiano e as respostas das plantas ao déficit hídrico, principalmente em culturas como a soja. Neste estudo, examinamos como o déficit hídrico interfere na oscilação diurna dos principais genes do ciclo circadiano e de resposta ao déficit hídrico em soja. Nossos resultados mostraram que o déficit hídrico causa mudanças marcantes na expressão gênica de componentes do ciclo circadiano, como *LCL1-*, *GmELF4-* e *PRR-like*, reduzindo sua expressão. As mesmas condições resultaram no avanço de fase na expressão dos genes *GmTOC1-like*, *GmLUX-like* e *GmPRR7-like*. O déficit hídrico também resultou em alterações significativas no padrão rítmico de expressão dos genes seca-responsivos tais como *DREB-*, *bZIP-*, *GOLS-*, *RAB18-* e *Remorin-like*. A fim de selecionar genes referência para a normalização da expressão gênica relativa em estudos sobre o déficit hídrico e as oscilações diurnas, nós validamos experimentalmente um conjunto inédito de genes referência de soja. Os genes *Fyve, NUDIX* e *Golgin-84* apresentaram expressão mais estável do que os genes normalizadores, convencionalmente utilizados, *ELF1-β* e *β-actina*, e foram considerados adequados para a normalização de dados de expressão em estudos sobre as respostas das plantas ao déficit hídrico em diferentes períodos do dia. Nossos resultados indicaram a existência de conexão entre as respostas ao déficit hídrico e o ciclo circadiano de soja, uma vez que (i) o estresse afetou a expressão de componentes do ciclo circadiano e que (ii) genes responsivos ao estresse apresentaram oscilação diurna. Esses resultados certamente irão incentivar estudos futuros sobre as implicações dessa conexão nas respostas das plantas a estas condições.

Palavras-chave: Seca. Expressão gênica. RT-qPCR. Transcriptoma. Genes referência. Genes normalizadores. Oscilações diurnas. Estresse hídrico. *Glycine max*.

GOMES, Juliana Marcolino. **“Identificação e análise da expressão de genes de soja responsivos ao déficit hídrico e ao ciclo circadiano”.** 2015. Teses (Doutor em Genética e Biologia Molecular). Londrina, 2015.

ABSTRACT

Drought is the most impacting abiotic stress influencing soybean crop production worldwide. Rhythms produced by the endogenous circadian clock play a critical role in allowing plants to respond and adapt to the environment. Recently, the circadian clock oscillations have been associated with the plant responses to water deficit at the transcription level. However, little is known about the regulatory link between the circadian clock and the plant responses to drought, mainly in crop species like soybean. In this study we investigated how drought impacts on diurnal oscillation of both drought responsive and circadian clock soybean genes. Our results show that drought induced marked changes in gene expression of several circadian clock-like components, such as *LCL1-*, *GmELF4-* and *PRR-like* genes, which had reduced expression in stressed plants. The same conditions produced a phase advance of expression for the circadian clock genes *GmTOC1-like*, *GmLUX-like* and *GmPRR7-like*. Similarly, the rhythmic expression pattern of the soybean drought-responsive genes *DREB-*, *bZIP-*, *GOLS-*, *RAB18-* and *Remorin-like* changed significantly after exposure to drought. In order to provide stable reference genes for accurately quantify relative gene expression in studies combining drought and diurnal oscillations, a novel set of soybean genes was validated. The genes *FYVE*, *NUDIX* and *Golgin-84* showed greater expression stability when compared to the conventionally used housekeeping genes *ELF1-β* and *β-actin*, and they are considered suitable for gene expression normalization in studies about plant responses to drought in different times of day. Our results indicated that some connection between the drought response and the circadian clock may exist in soybean since (i) drought stress affected gene expression of circadian clock components and (ii) several stress responsive genes displayed diurnal oscillation in soybeans. These results will certainly encourage further studies about this connection in plant response to these conditions.

Key-words: Drought. Gene expression. RT-qPCR. Transcriptome. Reference genes. Normalizers. Diurnal oscillations. Water stress. *Glycine max*.

LISTA DE FIGURAS

Figuras da Revisão Bibliográfica:

Figuras do Artigo 1

Figuras do Artigo 2

LISTA DE TABELAS

Tabelas do Artigo 1

Tabelas do Artigo 2

SUMÁRIO

1. INTRODUÇÃO

O déficit hídrico constitui um dos principais fatores abióticos que prejudica o desenvolvimento e a produtividade de culturas economicamente importantes, como a soja. Recentemente no Brasil, a ocorrência de um prolongado período de estiagem ocasionou a perda de milhões de toneladas de grãos na safra 2013/2014 (CONAB, 2012). Só no estado de SP, as adversidades climáticas provocaram forte redução na produtividade (30,2%). Este cenário poderá ainda ser agravado se consideradas as projeções de mudanças climáticas globais que preveem frequentes eventos de estiagem para as próximas décadas (IPCC, 2013). Diversas estratégias têm sido utilizadas, por meio da biotecnologia, para o desenvolvimento de plantas mais tolerantes ao déficit hídrico. Neste contexto, o estudo de genes relacionados aos mecanismos de resposta e tolerância ao déficit hídrico em plantas tem se destacado no melhoramento genético de soja (MOROZOVA; HIRST; MARRA, 2009; SHANKER et al., 2014; SHARONI et al., 2011; SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2007).

O mecanismo de resposta ao déficit hídrico utiliza rotas metabólicas que são reguladas pelo ácido abscísico (ABA-dependentes) e outras que atuam de forma independente do ABA. Em ambas as rotas a transdução do sinal de estresse culmina na ativação de fatores de transcrição, estes induzem a expressão dos genes de resposta ao déficit hídrico (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2007), os quais, por sua vez, codificam proteínas de diversas classes tais como as proteínas transdutoras de sinais (quinases, fosfatases) (MELCHER et al., 2009); os fatores reguladores da transcrição DREB (*Dehydration Responsive Element Binding*) (MIZOI; SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2012) e AREB (ABA - *Responsive Element Binding*) (YOSHIDA et al., 2010); as proteínas osmoprotetoras LEA (*Late Embryogenesis Abundant proteins*; (SAVITRI et al., 2013); proteínas estruturais de membrana (Remorinas) (RAFFAELE et al., 2007) e enzimas detoxificadoras de espécies reativas de oxigênio (peroxidases, catalases) (KHANNA-CHOPRA; SELOTE, 2007).

O ciclo circadiano controla processos chave para o crescimento, desenvolvimento das plantas, assim como a aclimatação às oscilações ambientais. O ciclo circadiano é um oscilador, composto por *loops* de *feedback* interconectados, que age como um “marca-passo” interno. Os principais componentes do oscilador central do ciclo circadiano de plantas são *TOC1/PRR1* (*Timing of Cab Expression1/Pseudo-Response Regulator1*), *LHY* (*Late Elongated Hypocotyl*), *CCA1* (*Circadian Clock Associated1*), *PRR3*, *5*, *7* e *9* (*Pseudo-Response Regulator* 3, 5, 7 e 9), *ELF3* e *4* (*Early Flowering* 3 e 4), *LUX* (*Lux Arrhythmo*), *ZTL* (*Zeitlupe*) e *GI* (*Gigantea*) (POKHILKO et al., 2012). Todavia, pouco se sabe sobre os componentes do ciclo circadiano em soja (HUDSON, 2010; LIU et al., 2009; QUECINI et al., 2007), tornando necessários estudos que visem sua melhor caracterização.

Embora o ciclo circadiano seja um sistema endógeno das plantas ele pode também ser modulado por sinais ambientais, dentre eles luz e temperatura, os quais participam de sua sincronização com o ambiente e proporcionam a expressão apropriada de seus componentes (DING et al., 2007; HUANG et al., 2012; SEO et al., 2012). A capacidade de perceber alterações ambientais e ajustar as respostas biológicas sugere a participação do ciclo circadiano nas respostas das plantas a estresses abióticos. Estudos com *Arabidopsis thaliana* sugerem que o ciclo circadiano possui a capacidade de modular os perfis transcricionais em resposta ao déficit hídrico (WILKINS; BRÄUTIGAM; CAMPBELL, 2010) e às baixas temperaturas (BIENIAWSKA et al., 2008). LEGNAIOLI e colaboradores (2009) também propuseram mecanismos de relação entre o gene do ciclo circadiano *TOC1* e as respostas ao déficit hídrico em *Arabidopsis thaliana* (LEGNAIOLI; CUEVAS; MAS, 2009).

Neste contexto, o presente estudo tratou da avaliação da expressão de genes de soja homólogos aos principais genes reguladores do ciclo circadiano em *Arabidopsis thaliana* (*TOC1, CCA1/LHY, LUX, ELF4, CHE, GI, PRR, ZTL* e *Jumonji*), assim como da análise da expressão de genes de resposta ao déficit hídrico pertencentes às vias ABA-dependente (*ABAR, Remorina, GOLS, RAB18, bZIP, PP2C* e *SnRK2*) e ABA-independente (*DREB1, DREB2, DREB3*), ao longo do dia. Detalhes sobre os resultados dessas análises encontram-se descritos no manuscrito “*Diurnal Oscillations of Soybean Circadian Clock and Drought Responsive Genes*”, publicado no periódico PLOS One (doi:10.1371/journal.pone.0086402). Em suma, os resultados obtidos revelam que o déficit hídrico afeta a expressão de genes componentes do ciclo circadiano em soja, e que vários genes responsivos ao déficit hídrico apresentam oscilação ao longo do dia, indicando a existência de conexão entre as respostas à seca e o relógio circadiano de soja.

Considerando ainda a importância do processo de normalização para a acurácia de dados de expressão de quantificação gênica relativa, também foram realizados estudos de genes de referência para o estudo de expressão via RT-qPCR em resposta ao ciclo circadiano e ao déficit hídrico em soja. Os dados aqui apresentados contribuirão para a melhor compreensão do ciclo circadiano de soja e de sua participação nas respostas das plantas ao déficit hídrico.

2.OBJETIVOS

Avaliar a expressão de genes do ciclo circadiano e de resposta ao déficit hídrico em diferentes condições de disponibilidade hídrica e em resposta à aplicação de ABA ao longo do dia.

2.1 Objetivos específicos

- Identificar em soja genes homólogos aos genes do ciclo circadiano em *Arabidopsis thaliana*, *TOC1, CCA1/LHY, LUX, ELF4, CHE, GI, PRR, ZTL* e *Jumonji*;
- Obter o perfil de expressão dos genes do ciclo circadiano de soja em resposta a diferentes condições de disponibilidade hídrica e à aplicação exógena de ABA ao longo de diferentes períodos do dia;
- Obter o perfil de expressão dos genes de resposta ao déficit hídrico *ABAR, Remorin, GOLS, DREB1, DREB2, DREB3, RAB18, bZIP, PP2C* e *SnRK2*) em soja submetida a diferentes condições de disponibilidade hídrica e à aplicação exógena de ABA ao longo de diferentes períodos do dia;
- Identificar e validar genes de soja para serem utilizados como referência em estudos de expressão gênica em resposta ao déficit hídrico e às oscilações ao longo do dia;

3. REVISÃO BIBLIOGRÁFICA

3.1. A Soja

3.1.1. Origem e Dispersão

Diferentes autores discordam com relação ao centro de origem específico da soja, entretanto todos concordam que a área mais provável se localiza na região leste da Ásia. O cultivo deste grão permaneceu restrito ao continente asiático por milhares de anos, somente sendo introduzido na Europa no final do século XV, como curiosidade nos jardins botânicos da Inglaterra, França e Alemanha. Na segunda década do século XX, o teor de óleo e proteína do grão começou a despertar o interesse das indústrias mundiais. Nos Estados Unidos a produção tornou-se intensa após a Primeira Guerra Mundial (EMBRAPA, 2004; SHURTLEFF; AOYAGI, 2009).

No Brasil, a soja foi introduzida na Bahia no ano de 1882. Nas décadas seguintes, ela foi utilizada como cultura experimental em algumas instituições de pesquisa e como planta hortícula entre os descendentes de imigrantes japoneses. Na década de 60, a produção de soja teve um crescimento expressivo, pois o grão surgiu como uma opção de verão, em sucessão ao trigo, principal cultura do Sul do Brasil na época. A explosão do preço da soja no mercado mundial, em meados de 1970, despertou ainda mais o interesse dos agricultores e do próprio governo brasileiro para o plantio desta cultura.

Atualmente, a soja cultivada apresenta um grande número de cultivares adaptadas a diferentes regiões produtoras, fruto do intenso processo de melhoramento genético (EMBRAPA, 2004; SHURTLEFF; AOYAGI, 2009). Todavia, o processo de domesticação com a utilização de um pequeno grupo de ancestrais e o subsequente melhoramento resultaram no estreitamento de sua base genética, impactando negativamente sua capacidade de responder às situações adversas, como os estresses bióticos e abióticos. As cultivares modernas de soja cultivadas nos EUA, por exemplo, são provenientes de uma dúzia de linhagens originárias de uma pequena região do nordeste da China (TANKSLEY, 1997). No Brasil, o estudo sobre a base genética de 444 cultivares de soja brasileiras revelou a existência de 60 ancestrais. Entretanto, desse total, um grupo de apenas 14 ancestrais é responsável por 92,4% de toda a base genética, sendo os demais 46 ancestrais responsáveis por apenas 7,59% (WYSMIERSKI; VELLO, 2013). Esses dados indicam que a base genética atual da soja brasileira é bastante estreita, sendo muito similar à da soja dos EUA, com a qual compartilha seis dos principais ancestrais.

3.1.2. Importância econômica

A soja é uma cultura muito importante economicamente, pois é utilizada na alimentação humana e, principalmente, na alimentação animal, e tem sido cada vez mais aplicada na produção de biocombustíveis. O óleo extraído do grão é um dos produtos mais utilizados na alimentação humana e sua participação no mercado mundial de óleos vegetais comestíveis é de 27,5%. Já a proteína de soja é abundantemente empregada como complemento nutricional em diferentes formas de alimento com o objetivo de melhorar sua qualidade proteica e com a vantagem de ser mais barata que as proteínas de origem animal (SEDIYAMA, 2009). Além das funções nutricionais básicas, atualmente a soja tem sido apontada por seus efeitos benéficos à saúde humana em casos de diabetes (KWON et al., 2010), hipertensão (NAKAHARA et al., 2010), sintomas do climatério e câncer de mama (BARNES, 2010).

Atualmente o Brasil é o segundo maior produtor de soja, sendo prevista a produção recorde de 95,9 milhões de toneladas para a safra 2014/2015, segundo a Companhia Nacional de Abastecimento (Conab). O país ocupa também a posição de segundo maior exportador do produto, sendo que em 2014 foram exportados 38,8 milhões de toneladas de grãos, provenientes da safra 2013/2014 (CONAB, 2014).

3.1.3. O genoma da soja

A soja é uma planta herbácea, pertencente à família das leguminosas. O genoma diplóide, que derivou de um ancestral alotetraplóide, é um dos maiores genomas completos de planta já sequenciados, consistindo em 978 megabases (Mb). A última versão do genoma (Wm82.a2.v1) integra resultados de ~1.6 mi de *Expressed sequence tags* (*EST*s) provenientes de sequenciamento de última geração. Este genoma está organizado em 20 cromossomos, de acordo com os 20 grupos de ligação de soja. Possui 88,647 transcritos e 56,044 genes preditos como codificadores de proteína. [http://www.phytozome.net/soybean_er.php]. O genoma da soja apresenta uma estrutura gênica exon–intron altamente conservada com *Populus trichocarpa* e *Vitis vinifera* (SCHMUTZ et al., 2010).

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

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

A disponibilidade do genoma completo da soja, aliado às ferramentas de bioinformática e genômica funcional trazem novas oportunidades para a pesquisa da tolerância ao déficit hídrico em soja (GOODSTEIN et al., 2012; KULCHESKI et al., 2011; MANAVALAN et al., 2009; NASCIMENTO et al., 2012; ZHANG et al., 2008).

3.2 Déficit hídrico

3.2.1 Visão geral e Impactos econômicos

As plantas estão sujeitas às oscilações ambientais diurnas, resultantes do movimento do planeta em torno de seu eixo, gerando variações de temperatura e luminosidade. Entretanto, além das variações ambientais geradas por esse movimento, as plantas também estão sujeitas a outras variações ambientais, como a ocorrência de estresses bióticos e abióticos. Dentre os estresses abióticos, o déficit hídrico, em conjunto com o aumento da temperatura e radiação, torna-se a maior adversidade ambiental para a sobrevivência das plantas e para a produtividade agrícola (SHANKER et al., 2014).

A soja é considerada uma planta sensível ao déficit hídrico e seus períodos críticos de sensibilidade a falta de água encontram-se principalmente durante a emergência, floração e enchimento de vagens, ocasiões em que a ocorrência de déficit hídrico pode levar a drásticas diminuições na produtividade da cultura (PEDERSEN, 2003). Na safra de 2011/2012, períodos prolongados de estiagem, causados pelo fenômeno "La Niña", levaram à perda de 8,94 milhões toneladas de soja, correspondendo a um decréscimo de 11% em relação à safra anterior (CONAB, 2012). Nesse mesmo ano, o Departamento de Agricultura dos Estados Unidos (*United States Department of Agriculture* - USDA) relatou a estiagem mais severa e prolongada dos últimos 25 anos no país, na qual aproximadamente 88% das áreas produtoras de soja nos EUA foram afetadas, impactando severamente a produção e, por conseguinte, os preços dos alimentos (USDA, 2013).

Apesar das projeções realizadas no início do safra 2013/2014 terem apontado uma produção recorde de soja no Brasil, a qual elevaria o país como novo líder na produção mundial, a ocorrência de condições climáticas adversas, destacando-se a estiagem combinada à altas temperaturas, manteve o país na segunda posição do *ranking*, com uma produção de 86,3 milhões de toneladas (CONAB, 2014). Segundo levantamento da CONAB, em 2014 as perdas de produtividade da soja em decorrência de longos períodos de estiagem chegaram a 27,9% em alguns estados brasileiros, apontando o déficit hídrico como maior fator limitante da produção brasileira de soja (CONAB, 2014).

Segundo a FAO (*Food and Agriculture Organization* - Organização das Nações Unidas), o aumento da produtividade agrícola é fundamental para o crescimento econômico e segurança alimentar. Neste panorama, estudos no intuito de melhorar a produtividade da cultura sob condições ambientais desfavoráveis, como o déficit hídrico, tornam-se de suma importância.

3.2.2 Mecanismos anatômicos e morfo-fisiológicos de resposta ao déficit hídrico

Embora seja considerada sensível ao déficit hídrico, a soja é capaz de desenvolver mecanismos de aclimatação que a permitem tolerar o déficit hídrico e consequentemente sobreviver a determinados períodos de estresse. Diversas características têm sido associadas à capacidade de adaptação das plantas a ambientes de baixa disponibilidade de água, dentre elas, redução no tamanho da planta, maturidade precoce e sistema radicular mais profundo (CATTIVELLI et al., 2008).

Os mecanismos das plantas para tolerar o déficit hídrico podem ser divididos em 3 principais estratégias: (i) o retardo da dessecação, (ii) a evitação à seca, atualmente denominada como tolerância ao déficit hídrico, e (iii) o escape ao déficit hídrico. A estratégia de retardo da dessecação envolve mecanismos que permitam a manutenção da hidratação do tecido, como a diminuição da condutância estomática, o depósito de cera e espessamento da cutícula e o ajuste osmótico. Já a tolerância ao déficit hídrico envolve mecanismos que protegem as células e tecidos e permitem o seu funcionamento mesmo quando o conteúdo de água estiver baixo nas células. A estratégia de escape é adotada por diversas espécies de plantas anuais e consiste na conclusão de seu ciclo de vida em períodos de boa disponibilidade de água no ambiente, antes da época de seca (TAIZ; ZEIGER, 2013a). Embora haja uma divisão didática entre as estratégias das plantas para tolerar o déficit hídrico, as respostas a esse tipo de condição são complexas e na maioria das vezes envolvem a combinação de mais de uma dessas estratégias, tornando difícil estabelecer fronteiras entre elas.

O ajuste osmótico, por exemplo, contribui tanto para o retardo da dessecação quanto para a capacidade da planta de tolerar o déficit hídrico. Esse mecanismo envolve o acúmulo ativo de solutos na célula vegetal, que resulta na diminuição do potencial osmótico e, por conseguinte, do potencial hídrico, o que resulta na retenção de água na planta (BABITA et al., 2010; OGAWA; YAMAUCHI, 2006; TURNER et al., 2007). O ajuste osmótico ocorre tanto nas raízes quanto na parte aérea da planta. Nas raízes, ele possibilita a manutenção da absorção de água e a manutenção da pressão de turgor celular, que permite o crescimento celular nas raízes por expansão. Nas folhas, a manutenção da pressão de turgor celular possibilita a sustentação das trocas gasosas e, por conseguinte, da taxa fotossintética (SANDERS; ARNDT, 2012).

Além de contribuir para o turgor celular e absorção de água, diversas moléculas envolvidas no ajuste osmótico participam na proteção e estabilização das membranas celulares e dos complexos proteicos, sendo denominadas moléculas osmoprotetoras. Essas moléculas mitigam os danos causados pelo estresse osmótico, permitindo que a maquinaria metabólica continue funcionando, o que faz do ajuste osmótico não apenas um mecanismo de retardo da dessecação como também um mecanismo de tolerância ao déficit hídrico. A variação da capacidade de ajuste osmótico entre diferentes culturas pode ser um indício da habilidade de suportar períodos de déficit hídrico. A soja normalmente apresenta baixa capacidade de ajuste osmótico, entretanto, um estudo identificou genótipos com menor declínio de teor relativo de água (TRA) em função do ajuste osmótico, evidenciando uma maior habilidade destes genótipos em manter o turgor celular durante o déficit hídrico (JAMES; LAWN; COOPER, 2008).

O controle da condutância estomática também contribui para a manutenção do conteúdo de água em condições de déficit hídrico, uma vez que permite a redução da perda pela abertura estomática (TAIZ; ZEIGER, 2013a). A redução da condutância estomática pode ocorrer por meio de sinais hidráulicos (diminuição do turgor celular), químicos (ação do hormônio ácido abscísico – ABA) ou pela interação de ambos os mecanismos. Liu e colaboradores (2005) relataram um decréscimo na condutância estomática coincidente com o aumento de ABA no xilema, mesmo antes de ocorrerem mudanças significativas no turgor celular, indicando que o ABA produzido nas raízes ascendia via xilema levando ao fechamento estomático na parte aérea (LIU et al., 2005). Este mecanismo reduz a perda de água, contribuindo para a manutenção do conteúdo de água, mas limita a difusão de dióxido de carbono (CO_2) para os cloroplastos, levando à diminuição da atividade fotossintética. Desta maneira o controle da abertura estomática é essencial para manter um balanço ótimo entre o potencial hídrico adequado e as funções energéticas da planta (FLEXAS et al., 2006).

Outra importante característica adaptativa ao déficit hídrico é a presença de tricomas nas folhas, constantemente relacionados às plantas xerófilas. Os tricomas restringem a perda de água por transpiração, através da manutenção da camada limítrofe de ar nas folhas, contribuindo para o retardo da dessecação. Além de diminuir a perda de água, a presença de tricomas aumenta a reflectância da folha e reduz a temperatura foliar, fatores que em conjunto contribuem para a manutenção da fotossíntese (DU; FU; YU, 2009). Mecanismos que contribuem para a redução da temperatura foliar, como a presença de tricomas e mudanças no posicionamento das folhas, assumem importante papel em condições de déficit hídrico visto que o principal mecanismo de resfriamento das folhas, a evapotranspiração, encontra-se comprometido devido à redução na condutância estomática. Logo, o aumento na densidade de tricomas é considerado como uma das respostas ao déficit hídrico (FU et al., 2013).

Além da presença de tricomas, a deposição de cera e o aumento na espessura da parede celular são modificações adaptativas para a tolerância ao déficit hídrico, e contribuem para a diminuição da condutância epidermal. Sabe-se que a condutância total de uma planta é resultado da soma das condutâncias estomática e cuticular. Na parte aérea, quando os estômatos estão abertos, a condutância cuticular é uma fração irrelevante de condutância total, mas quando o déficit hídrico leva ao fechamento estomático, o componente cuticular da condutância epidermal pode exceder a condutância estomática. A manipulação do gene *Wax Production1* (*WXP1*), envolvido na produção de cera, resultou em plantas mais tolerantes ao déficit hídrico em *Arabidopsis thaliana*, alfafa e *Trifolium repens* (JIANG et al., 2010). Em soja, genótipos com baixa condutância epidermal têm mostrado maior tolerância à deficiência hídrica, sendo que genótipos de soja adaptados a ambientes tropicais possuem menor condutância epidermal quando comparados a genótipos de regiões temperadas, sugerindo que esta pode ser uma adaptação dos genótipos tropicais a ambientes mais secos (JAMES; LAWN; COOPER, 2008). Adicionalmente, a ocorrência de raízes fibrosas limita a perda de água por transpiração durante os períodos de maior demanda evaporativa do solo, além de possibilitar o crescimento em solos mais compactados, como descrito para a cultivar de soja PI416937, que apresenta maior tolerância ao déficit hídrico quando comparada a outras cultivares (ABDEL-HALEEM; LEE; BOERMA, 2011).

Nesse contexto, outro fator crítico para a adaptação das plantas ao déficit hídrico envolve a plasticidade das raízes. Isto porque em condições de déficit hídrico a água disponível encontra-se em camadas mais profundas do solo, tornando o desenvolvimento de raízes mais profundas uma importante característica. Em soja, assim como em diversas espécies vegetais, existe uma correlação positiva significativa entre a tolerância ao déficit hídrico e fatores relacionados à raiz, como peso seco, comprimento total, volume e número de raízes laterais (VÉGH, 2013). Conhecer os genes que codificam para essas características é essencial para o melhoramento genético de culturas de importância econômica. Em arroz, por exemplo, a manipulação do gene *Deeper Rooting 1* (*DRO1*), responsável pela arquitetura das raízes, possibilitou o desenvolvimento de plantas com maior tolerância ao déficit hídrico (UGA et al., 2013).

Por fim, o resultado da interação de diversos mecanismos morfo-fisiológicos de resposta e adaptação ao déficit hídrico, presentes tanto nas raízes como na parte aérea, pode ser traduzido em maiores valores de eficiência do uso de água (EUA), definida como a quantidade de biomassa acumulada por unidade de água utilizada. Associações positivas entre a EUA e a biomassa total têm sido estabelecidas, sugerindo que aumentos na EUA resultem em uma melhor produtividade. Entretanto, argumenta-se que a seleção ou melhoramento genético para aumento do EUA, pode levar à redução da produtividade e tolerância ao déficit hídrico. Isto visto que a bioquímica da fotossíntese não poderia ser melhorada geneticamente, e maiores taxas de EUA podem ser alcançados apenas através da redução da transpiração e uso de água da planta, processos cruciais para manter a produtividade sob condições de déficit hídrico (BLUM, 2009). Todavia, um estudo recente demonstrou que a eficiência da fotossíntese pode ser melhorada geneticamente em plantas através da expressão do gene *Se7942*, da cianobactéria *Synechococcus elongatus* PCC7942, que codifica uma enzima Rubisco mais eficiente na fixação de carbono (LIN et al., 2014).

Além disso, existem evidências de que a ocorrência de déficit hídrico moderado em diferentes estádios do desenvolvimento pode aumentar a capacidade da planta em tolerar futuros eventos de estresse (KRON; SOUZA; RIBEIRO, 2008). Segundo os autores, a ocorrência de déficit hídrico em soja nos estádios vegetativos anterior ao florescimento e enchimento de grãos, permitiu que as plantas desenvolvessem mecanismos de aclimatação que resultaram na maior tolerância a um segundo período de falta de água.

Compreender os mecanismos anatômico e morfo-fisiológicos envolvidos na tolerância ao déficit hídrico, assim como os genes e rotas metabólicas relacionados a eles fornecerá subsídios tanto para a seleção assistida por marcadores quanto para o desenvolvimento de estratégias de engenharia genética para o desenvolvimento de cultivares mais tolerantes (MANAVALAN et al., 2009).

3.2.3 Respostas moleculares ao déficit hídrico

A capacidade de tolerar condições ambientais adversas envolve a participação de diversos mecanismos de defesa, os quais iniciam em nível molecular e celular e culminam na fisiologia da planta como um todo. De acordo com Shanker e colaboradores, o déficit hídrico pode induzir três tipos principais de respostas nas plantas: (a) alterações na expressão de genes envolvidos na manutenção de processos celulares vitais; (b) alterações na dinâmica de produção e degradação de proteínas com função de mitigar os danos causados pelo estresse e (c) modificações no *pool* metabólico para levar à produção de novas moléculas relacionadas à resposta ao estresse (SHANKER et al., 2014).

A primeira etapa na resposta molecular ao déficit hídrico é a percepção do estresse via receptores específicos. Uma vez ativados, estes receptores ativam (ou suprimem) uma cascata de eventos moleculares para transmitir a informação através de complexas vias de transdução de sinais (SEKI et al., 2007). A perda de água causa uma diminuição no turgor celular, levando à percepção mecânica do déficit hídrico (KALEFETOGLU; EKMEKÇI, 2005). Em *Arabidopsis thaliana* foram identificadas proteínas receptoras histidina quinase, codificadas pelos genes *Histidina Kinase 1, 2* e *3* (*AHK1/ATHK1, AHK2* e *AHK3*), e receptores de citocininas, como o codificado pelo gene *Cytokinin Response1* (*CRE1*). Estas proteínas possuem domínios receptores e domínios transmembranas que permitem detectar as mudanças no potencial osmótico da célula em resposta ao déficit hídrico e osmótico. Estudos de repressão ou superexpressão dos genes *AHK1* demonstram que ele é um regulador positivo de genes que atuam em resposta ao déficit hídrico, estresse salino e ao ABA (TRAN et al., 2007). O gene *NtC7* (*Nicotiana tabacum clone 7*) codifica uma proteína similar a um receptor de membrana em tabaco. Esse gene tem sido descrito como sensível às alterações osmóticas em plantas e sua superexpressão levou ao aumento da tolerância ao estresse osmótico, sugerindo um importante papel na sinalização deste tipo de estresse (NAKAMURA; SANO, 2009).

O próximo passo na transdução do sinal, após a percepção por osmosensores, é a ativação de cascatas de fosforilação, envolvendo a participação de MAPKs (*Mitogen-Activated Protein Kinase*). A rota das MAPKs consiste em uma cascata de proteínas quinases que fosforilam de maneira ordenada, a partir de uma *MAPK kinase kinase* (MAPKKK) para uma *MAPK kinase* (MAPKK), e, então, para uma MAPK (ZHI-MING et al., 2006). Em plantas, um grande número de MAPKs tem sido identificado. Sabe-se que o genoma de *Arabidopsis thaliana* codifica aproximadamente 60 MAPKKKs, 10 MAPKKs e 20 MAPKs. Em soja, foram identificados 150 genes para MAPKKKs, 11 para MAPKKs e 38 para MAPKs. Os genes de soja duplicados para GmMAPK e GmMAPKKK apresentam relações evolutivas complexas e divergência funcional quando comparados aos genes de *Arabidopsis thaliana* (NEUPANE et al., 2013). Embora a função da cascata de fosforilação em resposta a estresses ainda não esteja totalmente esclarecida, sabe-se que os membros da cascata de MAPKs são ativados por mais de um tipo de estresse, o que sugere que esta via atue como ponto de convergência na sinalização de diferentes estresses (YOSHIOKA; SHINOZAKI, 2009). Uma MAPK de soja (*GmK1*) é induzida em resposta a estresse salino e mostrou-se regulada tanto pela ação de peroxido de hidrogênio (H_20_2) quanto de ácido fosfatídico (*phosphatidic acid*-PA) (IM et al., 2012).

Os derivados de fosfolipídeos constituem outra importante classe de sinalizadores de estresse. Sabe-se que a estrutura fluida da membrana plasmática é consequência do ambiente aquoso da célula, onde as caudas hidrofóbicas dos fosfolipídeos são repelidas pela água formando a bicamada lipídica. Quando a água deixa a célula, esta estrutura é alterada (KALEFETOGLU; EKMEKÇI, 2005). Hipóteses indicam que os fosfolipídeos presentes em membranas celulares sob condições de estresses são clivados por fosfolipases C e D, produzindo derivados que agem como mensageiros secundários. Os principais derivados de fosfolipídeos relacionados ao estresse osmótico são o inositol 1,4,5-trifosfato (IP3), diacilglicerol

(DAG) e ácido fosfatídico (PA) (MUÑOZ-SÁNCHEZ; ALTÚZAR-MOLINA; HÉRNANDEZ-SOTOMAYOR, 2012).

Adicionalmente, o processo de sinalização do estresse em plantas tem sido dividido em vias dependentes e independentes de ácido abscísico (ABA). Sabe-se que em condições de estresse, como o déficit hídrico, ocorre naturalmente o acúmulo de ABA. Na sinalização dependente de ABA este hormônio liga-se a proteínas START, que são receptores proteicos, denominados PYR/PYL/RCARs. Esta classe de proteínas START inibe a ação de fosfatases, permitindo a ativação de SnRK2 quinases. As quinases ativadas nessa via fosforilam efetores *downstream* tanto no citoplasma quanto no núcleo. No citoplasma essas quinases agirão na fosforilação dos canais iônicos e canais de ânions, nas membranas das células guarda, levando à despolarização que intermediará o fechamento estomático induzido por ABA (MELCHER et al., 2009). No núcleo, as quinases fosforilam fatores de transcrição conhecidos como *ABA-responsive element binding factors* (ABFs/AREBs). Fatores de transcrição são proteínas que se ligam a sequências específicas de DNA, os *cis*-elementos, presentes na região promotora dos genes alvo, modulando sua expressão. Os fatores de transcrição AREB ligam-se aos cis-elementos *ABA-responsive element* (ABRE) e induzem a expressão de diversos genes alvo em resposta ao déficit hídrico (RAGHAVENDRA et al., 2010).

Na via de sinalização ABA dependente, o cálcio citosólico (Ca^{+2}) atua como um dos principais mensageiros secundários. Esta via ocorre mediante rápidas mudanças na concentração do cálcio citosólico, desencadeadas pela participação dos canais de cálcio, que regulam os níveis deste elemento. Como molécula sinalizadora, o cálcio está estreitamente relacionado com diversos estresses abióticos, dentre eles o déficit hídrico (KIM et al., 2009). A importância dessa via de sinalização para a resposta a estresses abióticos é demonstrada por estudos como o de Cheong e colaboradores, no qual a superexpressão de uma proteína sensora de cálcio em *Arabidopsis thaliana* conferiu maior tolerância ao estresse osmótico e ao déficit hídrico (CHEONG et al., 2010).

A rede de sinalização via cálcio é regulada por proteínas decodificadoras de sinais de cálcio. Três classes de moléculas sensíveis a alterações na concentração de cálcio foram caracterizadas em plantas: calmodulinas, *Calcium-Dependent*

Protein Kinases (CDPKs) e *Calcineurin B-like Proteins* (CBLs) (BATISTIC; KUDLA, 2009). CDPKs são proteínas quinases que possuem um domínio auto-regulatório no qual se liga o cálcio. Estas proteínas são amplamente descritas na sinalização ABA-dependente e, em alguns casos, genes codificadores de CDPKs possuem elementos ABA-responsivos em seus promotores, mostrando-se induzidos em nível transcricional na presença de ABA. Sabe-se que as proteínas CBL, inicialmente descritas em *Arabidopsis thaliana,* interagem principalmente com quinases SnRK3, formando uma rede interligada de sinalização (COELLO; HEY; HALFORD, 2011; HEY; BYRNE; HALFORD, 2010).

Entretanto, existem também vias de sinalização para ativação da expressão gênica em resposta ao déficit hídrico que são ativadas independente do ABA, as quais envolvem principalmente membros da família *APETALA 2/Ethylene-Responsive Element Binding Factor* (AP2/ERF). Esta família compreende um grande grupo de fatores de transcrição específicos de plantas, incluindo as subfamílias *Apetala* 2 (AP2), *Related to ABI3/VP1* (RAV), *Ethylene Response Factors* (ERF) e *C-repeat/DRE-Binding Factor / Dehydration Responsive Element Binding* (CBF/DREB) (MIZOI; SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2012). Dentro da família AP2/ERF, as subfamílias ERF e DREB são descritas na resposta a estresses, sendo que os membros da subfamília ERF ligam-se a *cis*-elementos de resposta ao etileno (*Ethylene Responsive Element* - ERE), enquanto os da subfamília DREB possuem maior afinidade com *cis*-elementos de resposta ao déficit hídrico (*Dehydration Responsive Elements* - DRE).

O elemento responsivo à desidratação (DRE) foi identificado como um importante *cis*-elemento atuante na regulação da expressão gênica em resposta ao déficit hídrico, alta salinidade e frio (LIU et al., 1998). Durante estresses abióticos, a interação entre os fatores de transcrição DREB e os elementos *cis*-atuantes DRE resulta na ativação de genes estresse-induzidos, importantes na resposta de defesa da planta (MARUYAMA et al., 2012; SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2007). Em uma análise *in silico*, Wang e seus colaboradores (2009) identificaram 160 genes de *Arabidopsis thaliana*, responsivos a estresses abióticos, alvos dos fatores de transcrição DREB (WANG et al., 2009). Em soja, análises de *Expressed Sequence Tags* (ESTs) indicam a presença de 148 membros da superfamília AP2/ERF, dos quais 36 genes pertencem à subfamília DREB (ZHANG et al., 2008).

Dentro desta subfamília, os genes *DREB1* e *DREB2* assumem destaque na via de resposta independente de ABA. Os genes *DREB1* foram inicialmente identificados na resposta ao frio (LIU et al., 1998), sendo mais tarde também descritos na resposta a alta salinidade (QIN et al., 2004), estresse térmico e déficit hídrico (KIDOKORO et al., 2014). Já os genes *DREB2* foram inicialmente descritos principalmente na resposta ao déficit hídrico (SAKUMA et al., 2002, 2006a), sendo mais tarde associados também ao estresse osmótico (KIM et al., 2011) e térmico (SAKUMA et al., 2006b; SCHRAMM et al., 2008). A superexpressão dos genes *DREB* em diferentes espécies de plantas como soja (CHEN et al., 2007), milho (QIN et al., 2004), girassol (ALMOGUERA et al., 2009), arroz (BIHANI; CHAR; BHARGAVA, 2010), cana-de-açúcar (REIS et al., 2014) e amendoim (SARKAR et al., 2014), aumentou a tolerância das plantas transgênicas a estresses abióticos. Embora os genes DREB sejam considerados parte da via de sinalização independente de ABA, análises de expressão gênica e do acúmulo de proteínas em plantas transgênicas para o gene *ABA responsive AP2-like*1 (*ARAG1*) de arroz indicam o envolvimento desse gene na sinalização via ABA, sugerindo sua participação na justaposição entre as respostas dependente e independente de ABA (ZHAO et al., 2010).

A transdução do sinal de déficit hídrico e a ativação de fatores de transcrição culmina na expressão de proteínas funcionais *downstream* envolvidas na resposta ao estresse. Essas proteínas pertencem às mais variadas classes, como chaperonas, proteínas de canal de água, enzimas detoxificadoras e enzimas para a biossíntese de solutos orgânicos e inorgânicos (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2007). A Figura 1 ilustra os principais eventos moleculares de resposta ao déficit hídrico em plantas.

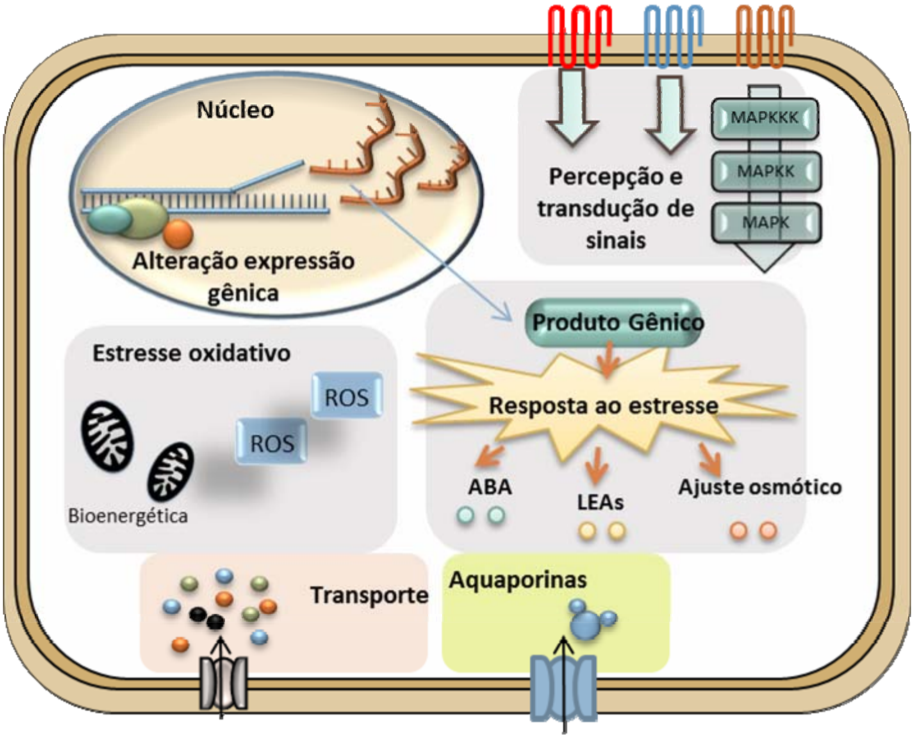


Figura 1. Principais respostas moleculares ao déficit hídrico na célula vegetal.

As proteínas chaperonas participam no dobramento e na estabilização de outras proteínas, mantendo-as em sua conformação funcional. Isso permite que enzimas e proteínas chave do metabolismo continuem desempenhando seu papel mesmo em condições adversas. As proteínas chaperonas da família Hsp70 (*Heat shock protein 70 kilodalton*) têm sido amplamente descritas na resposta a estresses como o déficit hídrico e o estresse térmico. Sua atividade baseia-se na propriedade de interagir, de maneira ATP-dependente, com seguimentos de peptídeos hidrofóbicos presentes nas proteínas-alvo (MAYER; BUKAU, 2005). Já as chaperonas Hsp40, também conhecidas como DNAJ, agem como co-chaperonas, induzindo e estabilizando a interação entre as Hsp70 e seus alvos (MAYER; BUKAU, 2005). Estudos demonstram que a superexpressão de chaperonas da família Hsp70 (AUGUSTINE et al., 2015) e Hsp40 (XIA et al., 2014) resultou no aumento da tolerância ao déficit hídrico em cana-de-açúcar e *Arabidopsis*, respectivamente.

Outro importante grupo de proteínas de proteção contra o déficit hídrico é composto por proteínas da família *Late Embryogenic Abundant protein* (LEA). Essas proteínas, inicialmente descritas em sementes, são proteínas hidrofóbicas que se acumulam em tecidos tolerantes à dessecação (DURE; GREENWAY; GALAU, 1981). Mais tarde, transcritos de diversos genes para LEAs foram detectados em diferentes órgãos vegetais em resposta à estresses abióticos, incluindo o déficit hídrico (HUNDERTMARK; HINCHA, 2008). Embora seu mecanismo de ação não seja totalmente compreendido, até o presente momento nenhuma atividade enzimática foi reportada para essas proteínas, o que sugere seu papel de proteção física de membranas e biomoléculas durante a desidratação. Estudos recentes apontam sua participação na estabilização de membranas, na prevenção da agregação de proteínas e no preenchimento do espaço intracelular, formando uma estrutura conhecida como “escudo celular” (CANDAT et al., 2014; CHAKRABORTEE et al., 2012).

Também faz parte da proteção contra a desidratação a indução da expressão de genes relacionados com a biossíntese e transporte de moléculas envolvidas no ajuste osmótico, como aminoácidos e derivados (prolina, ectoína, glicina betaína), polióis (manitol, sorbitol e D-ononitol), açúcares (trehalose) e íons K^+, Na^+ e Cl^- (Silva et al. 2010). Como abordado anteriormente, a síntese e acúmulo dessas moléculas propicia a diminuição do potencial osmótico e hídrico, o que resulta na retenção de água e na manutenção da pressão de turgor necessários para manutenção do crescimento e da abertura estomática (BABITA et al., 2010; MOLINARI et al., 2007; OGAWA; YAMAUCHI, 2006; TURNER et al., 2007).

O déficit hídrico também modifica a expressão de proteínas de canal de água, denominadas aquaporinas. Em condições normais de hidratação, a água é absorvida nas raízes pelas rotas simplástica e apoplástica (TAIZ; ZEIGER, 2013a). No entanto, em condições de déficit hídrico, a rota simplástica assume maior importância, e as aquaporinas assumem um papel importante como canais de controle do fluxo de água entre membranas. Estudos demonstraram que genes que codificam aquaporinas, como *Plasma Membrane Intrinsic Proteins 1* (*PIP1*) e *PIP2,* são induzidos pelo déficit hídrico. A superexpressão desses genes resultou no aumento da tolerância ao déficit hídrico em plantas transgênicas de banana devido à diminuição de danos à membrana e melhor controle do potencial osmótico (XU et al., 2014), e em plantas de tabaco devido à ativação de mecanismos antioxidantes (ZHOU et al., 2012).

Mecanismos de proteção contra a oxidação são extremamente importantes na resposta das plantas às condições limitantes de água, uma vez que o aumento da geração de espécies reativas de oxigênio (ROS - do inglês *Reactive Oxygen Species*) é uma das consequências do déficit hídrico. Durante o déficit hídrico, ROS são geradas principalmente nos cloroplastos, pela transfêrencia de elétrons para o O_2, e nos peroxissomos, devido ao aumento da fotorrespiração (CARVALHO, 2008; SUZUKI et al., 2012). Quando em níveis adequados, as ROS possuem um papel importante na manutenção do fluxo de energia nos sistemas de oxido-redução e na sinalização do estresse (BAXTER; MITTLER; SUZUKI, 2014; SUZUKI et al., 2012). Todavia, níveis elevados de ROS causam danos às membranas e outros componentes celulares, podendo resultar na morte celular. Por esse motivo, as plantas desenvolveram um sistema composto por enzimas detoxificadoras de ROS, composto por catalases, superóxido-desmutases e peroxidases. O estudo de diferentes cultivares de cana-de-açúcar sugere que a presença de um sistema mais eficaz para a detoxificação de espécies reativas de oxigênio é o responsável pelo maior nível de tolerância ao déficit hídrico em alguns genótipos (BOARETTO et al., 2014).

3.3 O ciclo circadiano e as plantas

3.3.1 Visão geral

.As plantas como organismos sésseis fotossintéticos evoluíram ao longo dos anos para se adaptar a estas oscilações diurnas, desenvolvendo um sistema endógeno complexo que otimiza os diversos processos metabólicos em resposta às variações do dia, denominado ciclo circadiano. A observação de que os organismos possuem ritmos diários têm intrigado pesquisadores, incluindo o naturalista Charles Darwin, que descreve o movimento foliar de plantas em seu livro "*The Power of Movement in Plants*", publicado em 1880. Entretanto foi apenas a partir de 1930 que o ciclo circadiano passou a ser melhor compreendido, através da publicação de estudos em realizados em *Drosophila melanogaster* e *Phaseolus*, onde o ritmo circadiano foi descrito como fruto de um mecanismo endógeno dos organismos, o ciclo circadiano (PITTENDRIGH, 1993). Em 1980 a descoberta do acúmulo circadiano de três transcritos nucleares, codificantes da proteína *Light-Harvesting Chlorophyll a/b Binding protein* (LHCB), da subunidade pequena da RuBisCO, e de uma proteína induzida pela luz (*Early Light-induced Protein* - ELIP) (KLOPPSTECH, 1985), marcou o início do estudo aprofundado dos fatores genéticos e moleculares que permeiam o ciclo circadiano em plantas. A partir da década de 90, diversos estudos estabeleceram que a expressão de vários genes encontra-se sob controle do ciclo circadiano (MCCLUNG, 2006).

Atualmente, diversos componentes moleculares do ciclo circadiano têm sido identificados em uma vasta gama de organismos, desde os mais complexos, como vegetais superiores (GALEOU; PROMBONA, 2012; HUDSON, 2010; LIU et al., 2009; MCCLUNG, 2013; RAMOS et al., 2005), até organismos de organização celular mínima, como *Ostreococcus tauri*, uma alga fotossintetizante descrita como o menor organismo eucarioto vivo (CORELLOU et al., 2009). Embora a importância do ciclo circadiano para os organismos seja indicada por sua presença ubíqua entre os eucariotos, ainda há discordância sobre a sua origem e evolução. Uma das hipóteses vigentes sobre sua origem é o surgimento desse sistema para modular a replicação do DNA para o período do escuro, o que protegeria essa molécula dos danos causados pela luz, possibilitando que a informação genética fosse melhor preservada (PITTENDRIGH, 1993). No que diz respeito à sua evolução, a hipótese mais aceita, e mais parcimoniosa, é a de que esse sistema tenha surgido apenas uma vez, antes da divergência dos eucariotos, evoluindo e divergindo para originar os demais sistemas observados atualmente (MCCLUNG, 2013).

3.3.2 Componentes e Características

Em plantas, o relógio circadiano mais bem estudado e compreendido é o de *Arabidopsis thaliana*. Nesse organismo, o modelo proposto para o ciclo circadiano é compostos por três partes principais: as rotas de *input*, o oscilador central e as rotas de *output*. As rotas de *input* são as vias de entrada para a informação ambiental, permitindo que o ambiente seja capaz de modular os ciclos circadianos. O oscilador central age como um relógio circadiano responsável por coordenar os ritmos metabólicos e fisiológicos que resultarão nos *outputs*. O resultado da interação entre os sinais de *input*, *output* e o oscilador central é a geração de fenômenos cíclicos denominados ritmos circadianos. O oscilador circadiano e seus principais *inputs* e *outputs* encontram-se ilustrados na Figura 2.

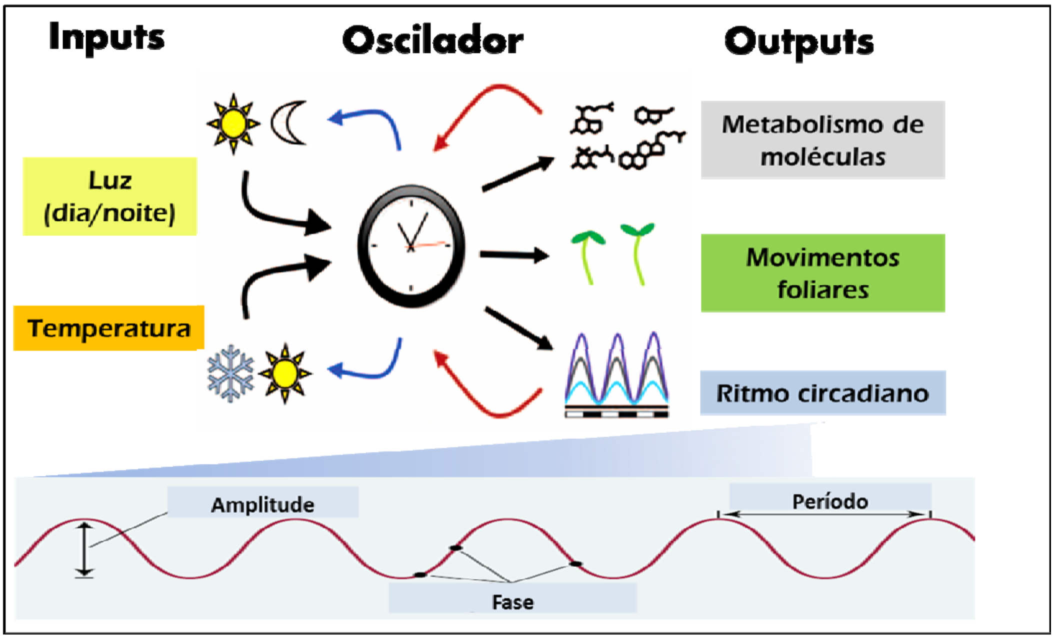


Figura 2. O ciclo circadiano: interações entre o oscilador, seus *inputs* e *outputs*. O oscilador central funciona como um relógio circadiano que é sincronizado pelos sinais ambientais de luz e temperatura, os quais agem como *inputs* do sistema. Por sua vez, o relógio circadiano controla múltiplas respostas de saída, os *outputs*, alguns dos quais possuem *feedback* para o relógio (setas vermelhas), portanto, também servindo como *inputs*. O relógio possui também a capacidade de interferir na percepção e na resposta aos *inputs* (setas azuis). O detalhe ao final da figura ilustra o Ritmo circadiano onde “Período**”** é o tempo entre pontos comparáveis no ciclo circadiano; “Fase” é qualquer ponto no ciclo reconhecível por seu relacionamento com o resto do ciclo e “Amplitude” é a distância entre um pico e um vale. Figura adaptada de (SANCHEZ; SHIN; DAVIS, 2011) e (TAIZ; ZEIGER, 2013b)**.**

3.3.3 Rotas de *input* do ciclo circadiano

Embora seja um sistema endógeno das plantas, os componentes do ciclo circadiano podem ser modulados por sinais ambientais, chamados *Zeitgeber* (em alemão “marcadores de tempo”), em um processo conhecido como *entrainment,* cuja tradução mais aproximada seria “sincronização*”*. Os principais sinais ambientais que modulam o ciclo circadiano são a luz e a temperatura (RUGNONE et al., 2013; THINES; HARMON, 2010) (Figura 2).

Existem três classes conhecidas de moléculas receptoras de luz em plantas: as fototropinas, fitocromos e criptocromos, sendo os fitocromos e criptocromos os mais importantes para a regulação do ciclo circadiano. Os fitocromos são compostos pela associação de um cromóforo a uma apoproteína e possuem a capacidade de interconversão em formas que absorvem diferentes comprimentos de luz, como vermelho (Pr), e vermelho distante (Prf). Uma vez ativados, possuem a capacidade de alterar os potenciais de membrana e o fluxo de íons que participam da rota de transdução de sinais ativada pela luz (TEPPERMAN et al., 2004).

Além disso, estudos demonstram que os fitocromos interagem com fatores de transcrição capazes de ativar ou reprimir a expressão gênica no núcleo. A luz vermelha ativa a forma Pfr do fitocromo B (PHYB), que é transferido para o núcleo, onde interage com o fator de transcrição *Phytochrome-Interacting Factor* (PIF3). Esse fator de transcrição liga-se a *cis*-elementos G box (CACGTG) presentes na região promotora de diversos genes controlados pela luz, dentre eles os genes do ciclo circadiano *CCA1* e *LHY*, os quais, por sua vez, codificam fatores de transcrição que se ligam a *cis*-elementos conhecidos como *CCA1-binding sites* (CBS), e induzem a expressão do gene *Chlorophyl a/b binding* (*CAB*) (QUAIL, 2002).

O componente do ciclo circadiano *Early Flowering 3* (ELF3) também parece ter papel fundamental na sincronização do ciclo circadiano pela luz. Estudos em levedura demonstram que ELF3 interage diretamente com o PHYB *in vivo*. Adicionalmente, dados obtidos a partir de mutantes *elf3/phyb* revelaram que a perda de função desses genes resulta no florescimento precoce em *Arabidopsis thaliana* e influencia diversos processos controlados pelo ciclo circadiano (LIU et al., 2001).

Estudos mais recentes identificaram uma nova família de genes induzidos pela luz e regulados pelo ciclo circadiano, a família *Night light–inducible and clock-regulated* (LNK) (RUGNONE et al., 2013). Segundo os autores, os genes dessa família conectam a regulação da expressão gênica pela luz e o controle dos ritmos circadianos diários e sazonais. Isso é possível através da ação dos genes homólogos *LNK1* e *LNK2*, que integram os sinais de luz no início do dia e as informações temporais providas pelo oscilador central do ciclo circadiano, a fim de controlar a expressão dos genes expressos durante a tarde, o que permite que as plantas adaptem-se às mudanças sazonais do comprimento do dia. Embora os mecanismos de interação entre os genes *LNK1* e *LNK2* e o ciclo circadiano não sejam inteiramente conhecidos, sabe-se que esses genes são diretamente reprimidos por membros da família TOC/PRR (*Timing Of Cab1 Expression/Pseudo*

Response Regulator), que possui diversos genes importantes do ciclo circadiano, e que LNK é um importante ativador transcricional de TOC1 e PRR5 (XIE et al., 2014).

Nesse contexto, a capacidade das plantas de percepção da duração da luz possibilita o monitoramento da duração do comprimento do dia, processo conhecido como fotoperiodismo. A percepção do fotoperíodo é crucial para a sincronização de processos importantes para as plantas, como o florescimento em determinadas estações do ano, e está diretamente ligado ao ciclo circadiano. A temperatura é outro importante sinal ambiental que funciona como *input*, influenciando no comprimento do período do ciclo circadiano. Quando cultivadas sob temperaturas entre 12°C e 32°C, por exemplo, o período do ciclo de plantas de *Arabidopsis* varia cerca de 2 h e 30 min. Entretanto, plantas mutantes para um gene chave do ciclo circadiano (*toc1*) apresentaram períodos de ciclo mais curtos em resposta a diferentes temperaturas, quando comparadas com plantas não transformadas, o que evidencia a interação entre os *inputs* de temperatura e o oscilador central (SOMERS et al., 1998).

Além de seu importante papel na resposta à luz, a participação do componente do ciclo circadiano *ELF3* também tem sido apontada na resposta à temperatura. Em plantas mutantes para esse gene, a aplicação de pulsos de temperatura falhou em ativar a expressão de genes conhecidamente responsivos à temperatura, tornando as plantas incapazes de responder a esse sinal ambiental. Esses resultados demonstram que a expressão do gene *ELF3* é requerida para o funcionamento adequado do ciclo circadiano e de sua capacidade de sincronização com as mudanças de temperatura que, assim como as alterações de luz, indicam a progressão de dia para noite (THINES; HARMON, 2010).

3.3.4 O oscilador central

A característica mais marcante do ciclo circadiano é sua ritmicidade, característica que permite a antecipação e sincronização das respostas (fisiológicas e moleculares) das plantas às oscilações que ocorrem ao longo do dia. Um atributo interessante do ritmo circadiano é seu caráter endógeno e autossustentável, refletido pela capacidade de manutenção da ritmicidade por períodos de aproximadamente

24h, mesmo na ausência de estímulos ambientais, como luz e temperatura (THINES; HARMON, 2011). Essa ritmicidade se dá pela oscilação na abundância dos componentes/efetores que compõe o oscilador central do ciclo circadiano (Figura 2).

Em *Arabidopsis thaliana,* o núcleo do oscilador central é composto pelos genes *Late Elongated Hypocotyl* (*LHY*) e *Circadian Clock Associated 1* (*CCA1*), que codificam fatores de transcrição da superfamília Myb, e pelo gene *Timing of CAB Expression 1* (*TOC1*), membro da família de proteínas *Pseudo-response Regulator* (PRR), também conhecido como *PRR1*. Os genes *LHY* e *CCA1* são expressos durante a manhã, enquanto *TOC1* possui pico de expressão à noite (ALABADÍ et al., 2001). Uma característica importante para a função dos fatores *LHY* e *CCA1* é sua característica de reconhecer e ligar-se a uma sequencia de 9 bp denominada *Evening Element* (EE; AAATATCT), presente na região promotora de genes da noite, incluindo *TOC1*, reprimindo sua expressão (ALABADÍ et al., 2001; HARMER et al., 2000).

Além de *TOC1* (*PRR1*), a família *PRR* inclui os genes do ciclo circadiano *PRR3*, *PRR5*, *PRR7* e *PRR9*, que apresentam expressão sequencial ao longo do dia, na ordem *PRR9- PRR7- PRR5- PRR3*, e agem como repressores de genes importantes do relógio circadiano, como *CCA1* e *LHY*, sendo necessários para a oscilação do relógio circadiano (NAKAMICHI et al., 2010).

Também fazem parte do oscilador central os genes expressos durante a noite *Early Flowering 3* (*ELF3*), *Early Flowering* (*ELF4*) e *Lux Arrythmo* (*LUX*). Juntos, esses fatores formam um complexo denominado *Evening Complex* (EC), que se liga ao promotor de genes do relógio circadiano agindo como repressor de sua expressão, e é essencial para a manutenção da ritmicidade (POKHILKO et al., 2012).

A proteína F-box *Zeitlupe* (ZTL) e o fator *Gigantea* (GI) também fazem parte do oscilador e desempenham papel importante na percepção da luz azul pelo ciclo circadiano e na regulação pós-traducional dos componentes do ciclo. Essas duas proteínas interagem diretamente através do domínio *Light, Oxygen or Voltage* (LOV) presente em *ZTL*, garantindo, assim, a oscilação dos níveis dessa proteína (KIM et al., 2007).

3.3.5 *Loops* de regulação transcricional e pós-traducional do oscilador central

O oscilador central é o responsável por gerar e manter a oscilação e o período circadiano. Para isso ele conta com uma série de *loops* de *feedback* negativo interconectados de forma que seus componentes variem em abundância de transcritos e proteínas (POKHILKO et al., 2012). Em *Arabidopsis thaliana*, foram descritos 3 *loops* principais:

(I) o *loop* central é composto pelos componentes centrais do oscilador: *TOC1*, *CCA1* e *LHY*. Nesse *loop*, as proteínas CCA1 e LHY ligam-se ao *cis*-elemento EE do gene *TOC1*, inibindo sua expressão. Adicionalmente, *CCA1* e LHY também inibem a expressão de *GI* durante o dia;

(II) o *loop* da noite é composto principalmente pelo complexo de proteínas ELF3–ELF4–LUX, conhecido como *Evening Complex* (EC). Esse complexo reprime a expressão dos genes *TOC1 e GI,* assim como dos próprios genes *LUX, ELF4 e ELF3*. O *loop* da noite conecta-se com o loop da manhã através da inibição dos genes *PRR* causada pela interação com o EC, que resulta no aumento da expressão de *LHY* e *CCA1* pela manhã;

(III) o *loop* da manhã é composto por *LHY*, *CCA1* e pelos genes *PRR* (5,7 e 9). Nesse *loop*, *LHY* e *CCA1* ativam a transcrição dos genes *PRR5, PRR7* e *PRR9* pela manhã. Todavia, o produto dos genes *PRR* atua como repressor da transcrição dos genes *LHY* e *CCA1*, e, neste mecanismo de *feedback* negativo, à medida que as quantidade das proteínas PRR aumentam, as quantidades dos fatores CCA1 e LHY diminuem. O *loop* da manhã conecta-se com o *loop* da noite pela ação dos fatores LHY e CCA1, que suprimem a transcrição dos genes da noite *TOC1, LUX, ELF4, ELF3* e *GI.*

Por muitos anos acreditou-se que TOC1 agiria como ativador de *CCA1* e *LHY* em um loop que envolveria a participação de um elemento “y”. Atualmente sabe-se que a ação de TOC1 sobre *CCA1* e *LHY* é de repressão, e que o EC é o responsável pela ativação indireta da expressão de *CCA1* e *LHY*, através da supressão da expressão de seus repressores (genes *PRR*s) (GENDRON et al., 2012; POKHILKO et al., 2012). A Figura 3 mostra a versão atualizada dos principais *loops* de regulação do oscilador circadiano.

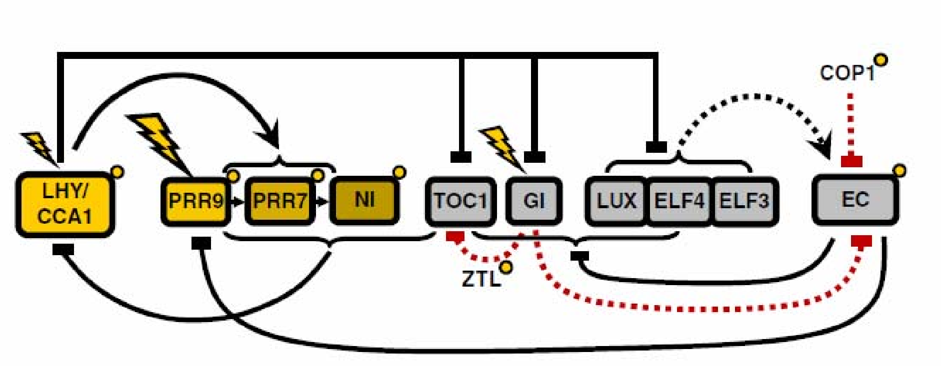


Figura 3. Versão atualizada e simplificada do oscilador circadiano de *Arabidopsis*. Os elementos dos loops da manhã e da noite são representados, em amarelo e cinza, respectivamente. Apenas as proteínas do EC, ZTL e COP1 são mostradas, para simplificar a representação. Linhas sólidas representam regulação transcricional. Linhas pontilhadas pretas representam a formação o EC. Linhas vermelhas representam a regulação pós-traducional. As respostas à luz são representadas por "raios" e círculos amarelos para os eventos de transcrição e tradução, respectivamente. Figura adaptada de (POKHILKO et al., 2012).

Além de *loops* de controle transcricional, há ainda a participação de *loops* pós-traducionais na modulação dos componentes do relógio circadiano. O componente do EC ELF3, por exemplo, sofre regulação pós-traducional pela degradação via *Constitutive Photomorphogenic 1* (COP1) ubiquitina ligase E3, induzida pela luz. Essa degradação permite a liberação da repressão de genes sob controle do EC, como o gene *TOC1* (POKHILKO et al., 2012).

Adicionalmente, a proteína F-box ZTL marca o fator TOC1 para degradação no proteassoma. Todavia, o componente GI interage fisicamente com ZTL, e essa interação (dependente de luz) estabiliza TOC1 durante o dia, inibindo sua degradação (KIM et al., 2007). A regulação da degradação do CONSTANS via COP1, que demarca essa proteína para degradação no proteassoma também possui influência no controle circadiano do florescimento em resposta à duração do dia (JANG et al., 2008).

Em conjunto, estes *loops* interconectados proporcionam a ocorrência de ritmos precisos e robustos, que modulam as respostas biológicas das plantas ao longo do dia.

3.3.6 Rotas de *output* do ciclo circadiano

O ciclo circadiano controla uma série de aspectos fundamentais para o desenvolvimento e metabolismo das plantas, tais como posicionamento foliar, condutância estomática, crescimento e reprodução (MCCLUNG, 2006) e é considerado uma importante característica adaptativa (YERUSHALMI; GREEN, 2009). O controle desses processos ocorre pelas rotas de *output* do ciclo circadiano, que são controladas pelo oscilador central, que por sua vez é modulado pelos sinais ambientais (*inputs*) (Figura 2). O estudo de mutantes para componentes do ciclo circadiano revelou que a oscilação circadiana intermedeia a sincronização entre a assimilação de energia luminosa e os períodos de luminosidade do dia levando à otimização da fotossíntese, ressaltando o papel do ciclo circadiano na sincronização dos processos fisiológicos e o ambiente em que a planta está inserida. Plantas com funcionamento normal dos genes do ciclo circadiano apresentaram maior assimilação de carbono, e, por conseguinte, maior acúmulo de biomassa, culminando em maior crescimento e maior vantagem adaptativa em relação às plantas mutantes (YERUSHALMI; GREEN, 2009).

O ciclo circadiano regula o crescimento e o desenvolvimento nas plantas principalmente através do controle de fitormônios. Análises de expressão gênica apontam que uma quantidade significativa de transcritos envolvidos no metabolismo, percepção e sinalização hormonal é regulada pelo ciclo circadiano, o qual possui uma interação significativa com os fitormônios ABA, auxina, citocininas e etileno. Essas interações sugerem que é através do controle das oscilações hormonais que o ciclo circadiano modula processos importantes para as plantas, como os movimentos estomáticos e o crescimento (ROBERTSON et al., 2009).

O fitormônio ácido abscísico (ABA) está entre os mais bem estudados fitormônios regulados pelo ciclo circadiano. O ABA modula processos essenciais para a planta que vão desde o crescimento e desenvolvimento até as respostas a estresses, como o déficit hídrico. Nestas condições, ocorre o acúmulo de ABA, que age principalmente no fechamento estomático (MELCHER et al., 2009) e na ativação de genes responsivos ao estresse (YOSHIDA et al., 2010).

O ciclo circadiano também controla o florescimento para que este ocorra em períodos que apresentem condições ambientais ótimas para a produção de sementes, visto que o controle do florescimento é um dos principais fatores para o sucesso reprodutivo. O ciclo circadiano possui a capacidade de mensurar a duração do dia, através da percepção das transições entre os períodos de luz/escuro, que correspondem às transições dia/noite e ao fotoperíodo (NIWA et al., 2007). Em nível molecular, o controle do florescimento ocorre pela ação de dois genes principais: *CO* (*CONSTANS*) e *FT* (*Flowering locus T*), sendo a expressão de *CO* controlada pelo ciclo circadiano, principalmente pelo fator GI. Em plantas de dia longo (DL), como *Arabidopsis thaliana*, GI forma um complexo proteico com o fator *Flavinbinding,Kelch Repeat*, F-Box 1 (FKF1) ao final da tarde. Esse complexo age no promotor de *CO,* liberando-o do efeito repressor de *Cycling Dof Factor* 1 (CDF1). Uma vez expresso em RNA e proteína, CO age como fator de transcrição ativando o gene *FT* no floema foliar. O produto da expressão de *FT* desloca-se ao meristema apical, onde inicia a cascata transcricional que induz ao florescimento (IMAIZUMI, 2010). Nesse mecanismo, CO age como ponto de integração entre o controle interno da planta (o ciclo circadiano), e os sinais ambientais externos (a duração do período de luz).

Em plantas de dia curto, como arroz e soja, as bases moleculares do controle do florescimento ainda estão sendo identificadas e caracterizadas. Em arroz, os genes *Heading Date 1* (*Hd1*) e *Hd3* foram identificados como ortólogos aos genes *CO* e *FT* de *Arabidopsis*, respectivamente, e são os responsáveis pelo florescimento. Em dias curtos, o gene *GI* de arroz ativa a expressão de *Hd1*, que por sua vez induz a expressão de *Hd3*a, promovendo o florescimento (TSUJI; TAOKA; SHIMAMOTO, 2011). Em soja, foram identificados 26 genes ortólogos ao gene *CO* de *Arabidopsis*, denominados *Glycine max CONSTANS- Like* (*GmCOL*). Análises da expressão desses genes revelaram que, em dias curtos, o pico de expressão de *GmCOL1a* e *GmCOL1b* ocorre ao final da madrugada, e coincide com o pico de expressão de *GmFT5a*, o gene ortólogo a *FT* de *Arabidopsis*, induzindo o florescimento. Por outro lado, em dias longos, a expressão de *GmCOL1a* e *GmCOL1b* ocorre no meio da noite e declina antes da madrugada, e nessas condições a expressão de *GmFT5a* não é detectada (WU et al., 2014).

O controle rítmico da expressão gênica também é um importante *output* do ciclo circadiano. Em *Arabidopsis thaliana*, análises de expressão gênica identificaram que 9 a 12% do genoma apresenta oscilação circadiana significativa (COVINGTON et al., 2008). Todavia, algumas predições apontam que até 36% dos genes de *Arabidopsis* podem estar sob controle circadiano (MICHAEL; MCCLUNG, 2003). Diversos componentes do oscilador central do ciclo circadiano são fatores de transcrição, podendo controlar a expressão de diversos genes. Desse modo, o controle temporal da expressão gênica via ciclo circadiano ocorre pela interação direta entre fatores de transcrição do oscilador central e as regiões reguladoras nos promotores de genes alvo. O gene *CCA1*, por exemplo, codifica um fator de transcrição que atinge níveis máximo pela manhã, e liga-se aos *evening elements* (EE) e *CCA1 binding sites* (CBS) presentes nos genes relacionados ao crescimento e síntese de açúcar e amido, reprimindo-os (THINES; HARMON, 2011).

3.4 O ciclo circadiano e as respostas a estresses abióticos

Como visto anteriormente, embora seja um sistema endógeno, os componentes do ciclo circadiano podem ser modulados por sinais ambientais, sendo a luz e a temperatura os principais (RUGNONE et al., 2013; THINES; HARMON, 2010). Entretanto, além das variações ambientais de luz e temperatura geradas pela alternação dia/noite e pela sazonalidade, as plantas também estão sujeitas à ocorrência de estresses abióticos como o frio e o déficit hídrico.

Nesse contexto, estudos recentes têm buscado compreender as relações entre o ciclo circadiano e as respostas das plantas a condições de estresses abióticos. Em 2005 Fowler e colaboradores demonstraram que a indução dos fatores de transcrição CBF/BREB em resposta à baixa temperatura é regulada pelo ciclo circadiano (FOWLER; COOK; THOMASHOW, 2005). Em um estudo posterior, a imposição deste mesmo tipo de estresse levou à ruptura do ciclo circadiano. Esta ruptura é apontada como responsável pela grande variação do transcriptoma de plantas sob baixas temperaturas (BIENIAWSKA et al., 2008). Adicionalmente, descobriu-se que a região promotora de diversos genes de resposta à baixa temperatura possui *cis*-elementos para modulação via ciclo circadiano, como EE, associados a elementos de resposta ao ABA, como *ABA-responsive elements* (ABRE), sugerindo que ambas as respostas estão integradas (MIKKELSEN; THOMASHOW, 2009).

Além das respostas ao estresse de baixa temperatura, o ciclo circadiano também tem sido associado às respostas ao déficit hídrico. Umas das primeiras evidências moleculares dessa interação foi a identificação 127 genes ABA-regulados com oscilação circadiana, dentre os quais estão genes de resposta ao déficit hídrico, como dehidrinas, *RD29* e *GOLS* (MIZUNO; YAMASHINO, 2008). Posteriormente Legnaioli e colaboradores (2009) estabeleceram uma relação entre o elemento chave do ciclo circadiano, *TOC1*, e o ABA, demonstrando que *TOC1* é induzido por ABA e liga-se ao promotor do gene ABA-relacionado *ABAR/CHLH/GUN5*, controlando sua expressão de maneira circadiana (LEGNAIOLI; CUEVAS; MAS, 2009). Estudos globais de expressão demonstraram ainda que o transcriptoma de *Arabidopsis thaliana* em resposta à seca varia significativamente ao longo do dia, demontrando que a magnitude das respostas ao déficit hídrico varia não apenas em função da duração do período de déficit hídrico, mas pela ação do ciclo circadiano (WILKINS; BRÄUTIGAM; CAMPBELL, 2010).

Recentemente, descobriu-se que PRR7, um componente central do ciclo circadiano, reprime a expressão de genes de resposta ao frio e ao déficit hídrico. Essa repressão ocorre pela interação direta entre o fator de transcrição PRR7 e a região promotora dos genes alvo, em um local próximo ao sítio de início da transcrição. Análises de imunoprecipitação da cromatina e sequenciamento apontaram que 80% dos genes alvo de PRR7 são responsivos ao déficit hídrico, destacando a importância desse fator de transcrição no controle das respostas a esse tipo de estresse (LIU et al., 2013). A integração entre os sinais ambientais de estresse e a regulação via PRR7 foi estabelecida em *Arabidopsis* pela descoberta de um novo repressor de PRR7, o fator de transcrição Heat Shock Factor B2b (*HsfB2b*). *HsfB2b* é induzido em resposta ao calor, e liga-se à região promotora de PRR7, reprimindo-o (KOLMOS et al., 2014). Tendo em vista o efeito repressor de PRR7 sobre genes de resposta a estresses abióticos, a repressão de PRR7 por *HsfB2b* pode constituir um evento chave para a ativação das respostas ao estresse.

Adicionalmente, a superexpressão do gene *Cold-circadian rhythm-RNA binding like* (*CCL*) em algodão resultou no aumento da tolerância ao déficit hídrico e estresse salino. Esse gene é responsivo a diversos estresses abióticos, dentre eles o estresse salino, manitol, estresse térmico e déficit hídrico, e está envolvido no aumento da estabilidade do mRNA. A análise da região promotora desse gene identificou diversos *cis*-elementos de resposta à luz, estresses abióticos e hormônios (DHANDAPANI et al., 2014). Com bases nos dados disponíveis atualmente, entende-se que o ciclo circadiano e os sinais ambientais, incluindo a ocorrência de estresses, interagem para regular os processos biológicos das plantas.

3.5. Bibliografia

ABDEL-HALEEM, H.; LEE, G.-J.; BOERMA, R. H. Identification of QTL for increased fibrous roots in soybean. **TAG. Theoretical and Applied Genetics.**, v. 122, n. 5, p. 935–46, mar. 2011.

ALABADÍ, D. et al. Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock. **Science**, v. 293, n. 5531, p. 880–3, 3 ago. 2001.

ALMOGUERA, C. et al. The HaDREB2 transcription factor enhances basal thermotolerance and longevity of seeds through functional interaction with HaHSFA9. **BMC Plant Biology**, v. 9, p. 75, jan. 2009.

AUGUSTINE, S. M. et al. Erianthus arundinaceus HSP70 (EaHSP70) overexpression increases drought and salinity tolerance in sugarcane (Saccharum spp. hybrid). **Plant Science**, v. 232, p. 23–34, mar. 2015.

BABITA, M. et al. Osmotic adjustment, drought tolerance and yield in castor (Ricinus communis L.) hybrids. **Environmental and Experimental Botany**, v. 69, n. 3, p. 243–249, dez. 2010.

BARNES, S. The biochemistry, chemistry and physiology of the isoflavones in soybeans and their food products. **Lymphatic Research and Biology**, v. 8, n. 1, p. 89–98, mar. 2010.

BATISTIC, O.; KUDLA, J. Plant calcineurin B-like proteins and their interacting protein kinases. **Biochimica et Biophysica Acta**, v. 1793, n. 6, p. 985–92, jun. 2009.

BAXTER, A.; MITTLER, R.; SUZUKI, N. ROS as key players in plant stress signalling. **Journal of Experimental Botany**, v. 65, n. 5, p. 1229–40, mar. 2014.

BIENIAWSKA, Z. et al. Disruption of the Arabidopsis circadian clock is responsible for extensive variation in the cold-responsive transcriptome. **Plant physiology**, v. 147, n. 1, p. 263–279, maio 2008.

BIHANI, P.; CHAR, B.; BHARGAVA, S. Transgenic expression of sorghum DREB2 in rice improves tolerance and yield under water limitation. **The Journal of Agricultural Science**, v. 149, n. 01, p. 95–101, 8 nov. 2010.

BLUM, A. Effective use of water (EUW) and not water-use efficiency (WUE) is the target of crop yield improvement under drought stress. **Field Crops Research**, v. 112, n. 2-3, p. 119–123, jun. 2009.

BOARETTO, L. F. et al. Water stress reveals differential antioxidant responses of tolerant and non-tolerant sugarcane genotypes. **Plant physiology and biochemistry**, v. 74, p. 165–75, jan. 2014.

CANDAT, A. et al. The ubiquitous distribution of late embryogenesis abundant proteins across cell compartments in Arabidopsis offers tailored protection against abiotic stress. **The Plant cell**, v. 26, n. 7, p. 3148–66, jul. 2014.

CARVALHO, M. H. C. Drought stress and reactive oxygen species. **Plant Signaling and Behaviour**, v. 3, n. 3, p. 156–165, 2008.

CATTIVELLI, L. et al. Drought tolerance improvement in crop plants: An integrated view from breeding to genomics. **Field Crops Research**, v. 105, n. 1-2, p. 1–14, jan. 2008.

CHAKRABORTEE, S. et al. Intrinsically disordered proteins as molecular shields. **Molecular bioSystems**, v. 8, n. 1, p. 210–9, jan. 2012.

CHEN, M. et al. GmDREB2, a soybean DRE-binding transcription factor, conferred drought and high-salt tolerance in transgenic plants. **Biochemical and biophysical research communications**, v. 353, n. 2, p. 299–305, 9 fev. 2007.

CHEONG, Y. H. et al. Constitutive overexpression of the calcium sensor CBL5 confers osmotic or drought stress tolerance in Arabidopsis. **Molecules and cells**, v. 29, n. 2, p. 159–65, 28 fev. 2010.

COELLO, P.; HEY, S. J.; HALFORD, N. G. The sucrose non-fermenting-1-related (SnRK) family of protein kinases: Potential for manipulation to improve stress tolerance and increase yield. **Journal of Experimental Botany**, v. 62, p. 883–893, 2011.

CONAB. **12º Levantamento de Grãos- Safra 2011/12**. . Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/12_09_06_09_18_33_boletim_graos_-_setembro_2012.pdf>.

CONAB. **10º Levantamento de Grãos- Safra 2013/14**. Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/14_07_09_09_36_57_10_levantamento_de_graos_julho_2014.pdf>. Acesso em: 31 jul. 2013.

CORELLOU, F. et al. Clocks in the green lineage: comparative functional analysis of the circadian architecture of the picoeukaryote ostreococcus. **The Plant cell**, v. 21, n. 11, p. 3436–49, nov. 2009.

COVINGTON, M. F. et al. Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development. **Genome biology**, v. 9, n. 8, p. R130, jan. 2008.

DHANDAPANI, G. et al. A Gene Encoding Cold-Circadian Rhythm-RNA Binding-Like Protein (CCR-Like) from Upland Cotton (Gossypium hirsutum L.) Confers Tolerance to Abiotic Stresses in Transgenic Tobacco. **Plant Molecular Biology Reporter**, 4 maio 2014.

DING, Z. et al. A complex genetic interaction between Arabidopsis thaliana TOC1 and CCA1/LHY in driving the circadian clock and in output regulation. **Genetics**, v. 176, n. 3, p. 1501–10, jul. 2007.

DU, W. J.; FU, S. X.; YU, D. Y. Genetic analysis for the leaf pubescence density and water status traits in soybean [Glycine max (L.) Merr.]. **Plant Breeding**, v. 128, n. 3, p. 259–265, jun. 2009.

DURE, L.; GREENWAY, S. C.; GALAU, G. A. Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by in vitro and in vivo protein synthesis. **Biochemistry**, v. 20, n. 14, p. 4162–4168, jul. 1981.

EMBRAPA. **Tecnologias de Produção de Soja Paraná 2004**. Disponível em: <http://www.cnpso.embrapa.br/producaosojaPR/SojanoBrasil.htm>. Acesso em: 6 fev. 2014.

FLEXAS, J. et al. Keeping a positive carbon balance under adverse conditions : responses of photosynthesis and respiration to water stress. **Physiologia Plantarumyiolo**, v. 127, p. 343–352, 2006.

FOWLER, S. G.; COOK, D.; THOMASHOW, M. F. Low temperature induction of Arabidopsis CBF1, 2, and 3 is gated by the circadian clock. **Plant physiology**, v. 137, n. 3, p. 961–8, mar. 2005.

FU, Q. S. et al. Leaf morphological and ultrastructural performance of eggplant (Solanum melongena L.) in response to water stress. **Photosynthetica**, v. 51, n. 1, p. 109–114, 24 jan. 2013.

GALEOU, A.; PROMBONA, A. Light at night resynchronizes the evening-phased rhythms of TOC1 and ELF4 in Phaseolus vulgaris. **Plant Science**, v. 184, p. 141–147, mar. 2012.

GENDRON, J. M. et al. Arabidopsis circadian clock protein, TOC1, is a DNA-binding transcription factor. **Proceedings of the National Academy of Sciences of the United States of America**, v. 109, n. 8, p. 3167–72, 21 fev. 2012.

GOODSTEIN, D. M. et al. Phytozome: a comparative platform for green plant genomics. **Nucleic acids research**, v. 40, n. Database issue, p. D1178–86, jan. 2012.

HARMER, S. L. et al. Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. **Science**, v. 290, n. 5499, p. 2110–2113, 15 dez. 2000.

HEY, S. J.; BYRNE, E.; HALFORD, N. G. The interface between metabolic and stress signalling. **Annals of botany**, v. 105, n. 2, p. 197–203, fev. 2010.

HUANG, W. et al. Mapping the core of the Arabidopsis circadian clock defines the network structure of the oscillator. **Science**, v. 336, n. 6077, p. 75–9, 6 abr. 2012.

HUDSON, K. A. The Circadian Clock-controlled Transcriptome of Developing Soybean Seeds. **The Plant Genome Journal**, v. 3, n. 1, p. 3, 2010.

HUNDERTMARK, M.; HINCHA, D. K. LEA (late embryogenesis abundant) proteins and their encoding genes in Arabidopsis thaliana. **BMC genomics**, v. 9, p. 118, jan. 2008.

IM, J. H. et al. Soybean MAPK, GMK1 is dually regulated by phosphatidic acid and hydrogen peroxide and translocated to nucleus during salt stress. **Molecules and cells**, v. 34, n. 3, p. 271–8, set. 2012.

IMAIZUMI, T. Arabidopsis circadian clock and photoperiodism: time to think about location. **Current opinion in plant biology**, v. 13, n. 1, p. 83–9, fev. 2010.

IPCC. **Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change**. Cambridge: Cambridge University Press, 2013. p. 1552

JAMES, A. T.; LAWN, R. J.; COOPER, M. Genotypic variation for drought stress response traits in soybean. I. Variation in soybean and wild Glycine spp. for epidermal conductance, osmotic potential, and relative water content. **Australian Journal of Agricultural Research**, v. 59, n. 7, p. 656, 2008.

JANG, S. et al. Arabidopsis COP1 shapes the temporal pattern of CO accumulation conferring a photoperiodic flowering response. **The EMBO journal**, v. 27, n. 8, p. 1277–88, 23 abr. 2008.

JIANG, Q. et al. Improvement of drought tolerance in white clover (Trifolium repens) by transgenic expression of a transcription factor gene WXP1. **Functional Plant Biology**, v. 37, n. 2, p. 157–165, 2010.

KALEFETOGLU, T.; EKMEKÇI, Y. The effects of drought on plants and tolerance mechanisms. **G.U. Journal of Science**, v. 18, n. 4, p. 723–740, 2005.

KHANNA-CHOPRA, R.; SELOTE, D. S. Acclimation to drought stress generates oxidative stress tolerance in drought-resistant than -susceptible wheat cultivar under field conditions. **Environmental and Experimental Botany**, v. 60, n. 2, p. 276–283, jun. 2007.

KIDOKORO, S. et al. Soybean DREB1/CBF-type transcription factors function in heat and drought as well as cold stress-responsive gene expression. **The Plant Journal**, v. 81, n. 3, p. 505–518, 2014.

KIM, J.-S. et al. An ABRE promoter sequence is involved in osmotic stress-responsive expression of the DREB2A gene, which encodes a transcription factor regulating drought-inducible genes in Arabidopsis. **Plant & cell physiology**, v. 52, n. 12, p. 2136–46, dez. 2011.

KIM, M. C. et al. Calcium and calmodulin-mediated regulation of gene expression in plants. **Molecular plant**, v. 2, n. 1, p. 13–21, jan. 2009.

KIM, W.-Y. et al. ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. **Nature**, v. 449, n. 7160, p. 356–60, 20 set. 2007.

KLOPPSTECH, K. Diurnal and circadian rhythmicity in the expression of light-induced plant nuclear messenger RNAs. **Planta**, v. 165, p. 502–506, 1985.

KOLMOS, E. et al. HsfB2b-mediated repression of PRR7 directs abiotic stress responses of the circadian clock. **Proceedings of the National Academy of Sciences of the United States of America**, v. 111, n. 45, 28 out. 2014.

KRON, A. P.; SOUZA, G. M.; RIBEIRO, R. V. Water deficiency at different developmental stages of glycine max can improve drought tolerance. **Bragantia**, v. 67, n. 1, p. 43–49, 2008.

KULCHESKI, F. R. et al. Identification of novel soybean microRNAs involved in abiotic and biotic stresses. **BMC genomics**, v. 12, n. 1, p. 307, jan. 2011.

KWON, D. Y. et al. Antidiabetic effects of fermented soybean products on type 2 diabetes. **Nutrition research (New York, N.Y.)**, v. 30, n. 1, p. 1–13, jan. 2010.

LEGNAIOLI, T.; CUEVAS, J.; MAS, P. TOC1 functions as a molecular switch connecting the circadian clock with plant responses to drought. **The EMBO journal**, v. 28, n. 23, p. 3745–57, 2 dez. 2009.

LIN, M. T. et al. A faster Rubisco with potential to increase photosynthesis in crops. **Nature**, v. 513, p. 547–550, 2014.

LIU, F. et al. ABA regulated stomatal control and photosynthetic water use efficiency of potato (Solanum tuberosum L.) during progressive soil drying. **Plant Science**, v. 168, n. 3, p. 831–836, mar. 2005.

LIU, H. et al. Analysis of clock gene homologs using unifoliolates as target organs in soybean (Glycine max). **Journal of plant physiology**, v. 166, n. 3, p. 278–89, 15 fev. 2009.

LIU, Q. et al. Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. **The Plant cell**, v. 10, n. 8, p. 1391–406, ago. 1998.

LIU, T. et al. Direct regulation of abiotic responses by the Arabidopsis circadian clock component PRR7. **The Plant journal**, v. 76, n. 1, p. 101–14, out. 2013.

LIU, X. L. et al. ELF3 Encodes a Circadian Clock – Regulated Nuclear Protein That Functions in an Arabidopsis PHYB Signal Transduction Pathway. v. 13, n. June, p. 1293–1304, 2001.

MANAVALAN, L. P. et al. Physiological and molecular approaches to improve drought resistance in soybean. **Plant & cell physiology**, v. 50, n. 7, p. 1260–76, jul. 2009.

MARUYAMA, K. et al. Identification of cis-acting promoter elements in cold- and dehydration-induced transcriptional pathways in Arabidopsis, rice, and soybean. **DNA Research**, v. 19, n. 1, p. 37–49, jan. 2012.

MAYER, M. P.; BUKAU, B. Hsp70 chaperones: cellular functions and molecular mechanism. **Cellular and molecular life sciences : CMLS**, v. 62, n. 6, p. 670–84, mar. 2005.

MCCLEAN, P. E. et al. Synteny mapping between common bean and soybean reveals extensive blocks of shared loci. **BMC genomics**, v. 11, p. 184, jan. 2010.

MCCLUNG, C. R. Plant circadian rhythms. **The Plant cell**, v. 18, n. 4, p. 792–803, abr. 2006.

MCCLUNG, C. R. **Beyond Arabidopsis: The circadian clock in non-model plant speciesSeminars in Cell and Developmental Biology**Elsevier Ltd, , maio 2013. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/23466287>. Acesso em: 28 set. 2013

MELCHER, K. et al. A gate-latch-lock mechanism for hormone signalling by abscisic acid receptors. **Nature**, v. 462, n. 7273, p. 602–8, 3 dez. 2009.

MICHAEL, T. P.; MCCLUNG, C. R. Enhancer trapping reveals widespread circadian clock transcriptional control in Arabidopsis. **Plant physiology**, v. 132, n. 2, p. 629–39, jun. 2003.

MIKKELSEN, M. D.; THOMASHOW, M. F. A role for circadian evening elements in cold-regulated gene expression in Arabidopsis. **Plant Journal**, v. 60, n. 2, p. 328–339, out. 2009.

MIZOI, J.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. AP2/ERF family transcription factors in plant abiotic stress responses. **Biochimica et biophysica acta**, v. 1819, n. 2, p. 86–96, fev. 2012.

MIZUNO, T.; YAMASHINO, T. Comparative transcriptome of diurnally oscillating genes and hormone-responsive genes in Arabidopsis thaliana: insight into circadian clock-controlled daily responses to common ambient stresses in plants. **Plant & cell physiology**, v. 49, n. 3, p. 481–7, mar. 2008.

MOLINARI, H. B. C. et al. Evaluation of the stress-inducible production of proline in transgenic sugarcane (Saccharum spp.): osmotic adjustment, chlorophyll fluorescence and oxidative stress. **Physiologia Plantarum**, v. 130, n. 2, p. 218–229, jun. 2007.

MOROZOVA, O.; HIRST, M.; MARRA, M. A. Applications of new sequencing technologies for transcriptome analysis. **Annual review of genomics and human genetics**, v. 10, p. 135–51, jan. 2009.

MUÑOZ-SÁNCHEZ, J. A.; ALTÚZAR-MOLINA, A.; HÉRNANDEZ-SOTOMAYOR, S. M. T. Phospholipase signaling is modified differentially by phytoregulators in Capsicum chinense J. cells. **Plant signaling & behavior**, v. 7, n. 9, p. 1103–5, 1 set. 2012.

NAKAHARA, T. et al. Antihypertensive effect of peptide-enriched soy sauce-like seasoning and identification of its angiotensin I-converting enzyme inhibitory substances. **Journal of agricultural and food chemistry**, v. 58, n. 2, p. 821–7, 27 jan. 2010.

NAKAMICHI, N. et al. PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the Arabidopsis circadian clock. **The Plant cell**, v. 22, n. 3, p. 594–605, mar. 2010.

NAKAMURA, K.; SANO, H. A plasma-membrane linker for the phosphoinositide-specific phospholipase C in tobacco plants. **Plant Signaling & Behavior**, v. 4, n. 1, p. 26–29, fev. 2009.

NASCIMENTO, L. C. et al. A web-based bioinformatics interface applied to the GENOSOJA Project: Databases and pipelines. **Genetics and molecular biology**, v. 35, n. 1 (suppl), p. 203–11, jun. 2012.

NEUPANE, A. et al. Identification, nomenclature, and evolutionary relationships of mitogen-activated protein kinase (MAPK) genes in soybean. **Evolutionary bioinformatics online**, v. 9, p. 363–86, jan. 2013.

NIWA, Y. et al. Genetic linkages of the circadian clock-associated genes, TOC1, CCA1 and LHY, in the photoperiodic control of flowering time in Arabidopsis thaliana. **Plant & cell physiology**, v. 48, n. 7, p. 925–37, jul. 2007.

OGAWA, A.; YAMAUCHI, A. Root Osmotic Adjustment under Osmotic Stress in Maize Seedlings 2 . Mode of Accumulation of Several Solutes for Osmotic Adjustment in the Root. **Plant Production Science**, v. 9, n. 1, p. 39–46, 2006.

PEDERSEN, P. **Crop Physiology: Yield, Maturity Groups, and Growth Stages**. Ames: Iowa State University, 2003. p. 28

PITTENDRIGH, C. S. Temporal organization: reflections of a Darwinian clock-watcher. **Annual review of physiology**, v. 55, p. 16–54, jan. 1993.

POKHILKO, A. et al. The clock gene circuit in Arabidopsis includes a repressilator with additional feedback loops. **Molecular systems biology**, v. 8, n. 574, p. 574, jan. 2012.

QIN, F. et al. Cloning and functional analysis of a novel DREB1/CBF transcription factor involved in cold-responsive gene expression in Zea mays L. **Plant & cell physiology**, v. 45, n. 8, p. 1042–52, ago. 2004.

QUAIL, P. H. Photosensory perception and signalling in plant cells: new paradigms? **Current Opinion in Cell Biology**, v. 14, n. 2, p. 180–188, abr. 2002.

QUECINI, V. et al. Identification of Soybean Genes Involved in Circadian Clock Mechanism and Photoperiodic Control of Flowering Time by In Silico Analyses. **Journal of Integrative Plant Biology**, v. 49, n. 11, p. 1640–1653, nov. 2007.

RAFFAELE, S. et al. Genome-wide annotation of remorins, a plant-specific protein family: evolutionary and functional perspectives. **Plant physiology**, v. 145, n. 3, p. 593–600, nov. 2007.

RAGHAVENDRA, A. S. et al. ABA perception and signalling. **Trends in plant science**, v. 15, n. 7, p. 395–401, jul. 2010.

RAMOS, A. et al. Winter disruption of the circadian clock in chestnut. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 19, p. 7037–42, 10 maio 2005.

REIS, R. R. et al. Induced over-expression of AtDREB2A CA improves drought tolerance in sugarcane. **Plant Science**, v. 221-222, p. 59–68, maio 2014.

ROBERTSON, F. C. et al. Interactions between circadian and hormonal signalling in plants. **Plant molecular biology**, v. 69, n. 4, p. 419–27, mar. 2009.

RUGNONE, M. L. et al. LNK genes integrate light and clock signaling networks at the core of the Arabidopsis oscillator. **Proceedings of the National Academy of Sciences of the United States of America**, v. 110, n. 29, p. 12120–5, 16 jul. 2013.

SAKUMA, Y. et al. DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. **Biochemical and biophysical research communications**, v. 290, n. 3, p. 998–1009, 25 jan. 2002.

SAKUMA, Y. et al. Functional Analysis of an Arabidopsis Transcription Factor , DREB2A , Involved in Drought-Responsive Gene Expression. **Plant Cell**, v. 18, n. May, p. 1292–1309, 2006a.

SAKUMA, Y. et al. Dual function of an Arabidopsis transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression. **Proceedings of the National Academy of Sciences of the United States of America**, v. 103, n. 49, p. 18822–7, 5 dez. 2006b.

SANCHEZ, A.; SHIN, J.; DAVIS, S. J. Abiotic stress and the plant circadian clock. **Plant Signaling & Behavior**, v. 6, n. 2, p. 223–231, 1 fev. 2011.

SANDERS, G. J.; ARNDT, S. K. Osmotic Adjustment Under Drought Conditions. In: AROCA, R. (Ed.). . **Plant Responses to Drought Stress**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. p. 199–229.

SARKAR, T. et al. Heterologous Expression of the AtDREB1A Gene in Transgenic Peanut-Conferred Tolerance to Drought and Salinity Stresses. **PloS one**, v. 9, n. 12, p. e110507, 2014.

SAVITRI, E. S. et al. Identification and characterization drought tolerance of gene LEA-D11 soybean (glycine max L . Merr) based on PCR-sequencing. **American Journal of Molecular Biology**, v. 3, p. 32–37, 2013.

SCHMUTZ, J. et al. Genome sequence of the palaeopolyploid soybean. **Nature**, v. 463, n. 7278, p. 178–83, 14 jan. 2010.

SCHRAMM, F. et al. A cascade of transcription factor DREB2A and heat stress transcription factor HsfA3 regulates the heat stress response of Arabidopsis. **The Plant journal : for cell and molecular biology**, v. 53, n. 2, p. 264–74, jan. 2008.

SEDIYAMA, T. **Tecnologias de Produção e Usos da Soja**. [s.l.] Mecenas, 2009. p. 314

SEKI, M. et al. Regulatory metabolic networks in drought stress responses. **Current opinion in plant biology**, v. 10, n. 3, p. 296–302, jun. 2007.

SEO, P. J. et al. A Self-Regulatory Circuit of CIRCADIAN CLOCK-ASSOCIATED1 Underlies the Circadian Clock Regulation of Temperature Responses in Arabidopsis. **The Plant cell**, v. 24, n. June, p. 2427–2442, 19 jun. 2012.

SHANKER, A. K. et al. Drought stress responses in crops. **Functional & integrative genomics**, v. 14, p. 11–22, 10 jan. 2014.

SHARONI, A. M. et al. Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. **Plant & cell physiology**, v. 52, n. 2, p. 344–60, fev. 2011.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Gene networks involved in drought stress response and tolerance. **Journal of experimental botany**, v. 58, n. 2, p. 221–7, jan. 2007.

SHURTLEFF, W.; AOYAGI, A. **History of soybeans and soyfoods in south america (1882-2009): extensively annotated bibliography and sourcebook**. Lafayette: Soyinfo Center, 2009.

SOMERS, D. E. et al. The short-period mutant, toc1-1, alters circadian clock regulation of multiple outputs throughout development in Arabidopsis thaliana. **Development**, v. 125, n. 3, p. 485–94, fev. 1998.

SUZUKI, N. et al. ROS and redox signalling in the response of plants to abiotic stress. **Plant, cell & environment**, v. 35, n. 2, p. 259–70, fev. 2012.

TAIZ, L.; ZEIGER, E. Respostas e Adaptações ao Estresse Abiótico. In: **Fisiologia Vegetal**. 5^{a}. ed. [s.l.] Artmed, 2013a. p. 954.

TAIZ, L.; ZEIGER, E. O controle do Florescimento. In: **Fisiologia Vegetal**. 5. ed. [s.l.] Artmed, 2013b. p. 954.

TANKSLEY, S. D. Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild. **Science**, v. 277, n. 5329, p. 1063–1066, 22 ago. 1997.

TEPPERMAN, J. M. et al. Expression profiling of phyB mutant demonstrates substantial contribution of other phytochromes to red-light-regulated gene expression during seedling de-etiolation. **The Plant journal : for cell and molecular biology**, v. 38, n. 5, p. 725–39, jun. 2004.

THINES, B.; HARMON, F. G. Ambient temperature response establishes ELF3 as a required component of the core Arabidopsis circadian clock. **Proceedings of the National Academy of Sciences of the United States of America**, v. 107, n. 7, p. 3257–62, 16 fev. 2010.

THINES, B.; HARMON, F. G. Four easy pieces: mechanisms underlying circadian regulation of growth and development. **Current opinion in plant biology**, v. 14, n. 1, p. 31–7, fev. 2011.

TRAN, L.-S. P. et al. Functional analysis of AHK1/ATHK1 and cytokinin receptor histidine kinases in response to abscisic acid, drought, and salt stress in Arabidopsis. **Proceedings of the National Academy of Sciences of the United States of America**, v. 104, n. 51, p. 20623–8, 18 dez. 2007.

TSUJI, H.; TAOKA, K.; SHIMAMOTO, K. Regulation of flowering in rice: two florigen genes, a complex gene network, and natural variation. **Current opinion in plant biology**, v. 14, n. 1, p. 45–52, fev. 2011.

TURNER, N. C. et al. Osmotic adjustment in chickpea (Cicer arietinum L.) results in no yield benefit under terminal drought. **Journal of experimental botany**, v. 58, n. 2, p. 187–94, jan. 2007.

UGA, Y. et al. Control of root system architecture by DEEPER ROOTING 1 increases rice yield under drought conditions. **Nature genetics**, v. 45, n. 9, p. 1097–102, set. 2013.

USDA. **Drought 2012 : Farm and Food Impacts**. Disponível em: <http://www.ers.usda.gov/topics/in-the-news/us-drought-2012-farm-and-food-impacts.aspx#.Uv4M2YZJNqh>.

VÉGH, K. R. Root and leaf traits, water use and drought tolerance of maize genotypes. **Biologia**, v. 68, n. 6, p. 1123–1127, 20 out. 2013.

WANG, S. et al. An in silico strategy identified the target gene candidates regulated by dehydration responsive element binding proteins (DREBs) in Arabidopsis genome. **Plant molecular biology**, v. 69, n. 1-2, p. 167–78, jan. 2009.

WILKINS, O.; BRÄUTIGAM, K.; CAMPBELL, M. M. Time of day shapes Arabidopsis drought transcriptomes. **Plant Journal**, v. 63, n. 5, p. 715–727, set. 2010.

WU, F. et al. Functional and evolutionary characterization of the CONSTANS gene family in short-day photoperiodic flowering in soybean. **PloS one**, v. 9, n. 1, p. e85754, jan. 2014.

WYSMIERSKI, P. T.; VELLO, N. A. The genetic base of Brazilian soybean cultivars : evolution over time and breeding implications. **Genetics and Molecular Biology**, v. 36, n. 4, p. 547–555, 2013.

XIA, Z. et al. Plant Physiology and Biochemistry Overexpression of a tobacco J-domain protein enhances drought tolerance in transgenic Arabidopsis. **Plant Physiology et Biochemistry**, v. 83, p. 100–106, 2014.

XIE, Q. et al. LNK1 and LNK2 Are Transcriptional Coactivators in the Arabidopsis Circadian Oscillator. **The Plant cell**, p. 1–16, 2014.

XU, Y. et al. A banana aquaporin gene, MaPIP1;1, is involved in tolerance to drought and salt stresses. **BMC plant biology**, v. 14, p. 59, jan. 2014.

YERUSHALMI, S.; GREEN, R. M. Evidence for the adaptive significance of circadian rhythms. **Ecology letters**, v. 12, n. 9, p. 970–81, set. 2009.

YOSHIDA, T. et al. 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. **Plant Journal**, v. 61, n. 4, p. 672–685, fev. 2010.

YOSHIOKA, K.; SHINOZAKI, K. **Signal Crosstalk in Plant Stress Responses**. Oxford, UK: Wiley-Blackwell, 2009. p. 1–210

ZHANG, G. et al. Phylogeny, gene structures, and expression patterns of the ERF gene family in soybean (Glycine max L.). **Journal of experimental botany**, v. 59, n. 15, p. 4095–107, jan. 2008.

ZHAO, L. et al. ARAG1, an ABA-responsive DREB gene, plays a role in seed germination and drought tolerance of rice. **Annals of botany**, v. 105, n. 2008, p. 401–409, 2010.

ZHOU, S. et al. Overexpression of the wheat aquaporin gene, TaAQP7, enhances drought tolerance in transgenic tobacco. **PloS one**, v. 7, n. 12, p. e52439, jan. 2012.

4. ARTIGO 1: “Diurnal oscillations of soybean circadian clock and drought responsive genes”

Artigo publicado no periódico PLOS ONE, Volume 9, Issue 1: e86402, Janeiro de 2014 (doi:10.1371/journal.pone.0086402).

Diurnal oscillations of soybean circadian clock and drought responsive genes

Juliana Marcolino-Gomes[a,b], Fabiana Aparecida Rodrigues[a], Renata Fuganti-Pagliarini[a], Claire Bendix[c], Thiago Jonas Nakayama[d], Brandon Celaya[c], Hugo Bruno Correa Molinari[e], Maria Cristina Neves de Oliveira[a], Frank G. Harmon[c] and Alexandre Nepomuceno[a,e]

Institutions

[a]Embrapa Soybean, Brazilian Agricultural Research Corporation , Londrina, Paraná, Brazil

[b]Department of Biology, State University of Londrina, Londrina, Paraná, Brazil

[c]Plant Gene Expression Center, ARS/USDA, Albany, California, USA and Department of Plant and Microbial Biology, University of California-Berkeley, Berkeley, California, USA

[d]Department of Crop Science, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil

[e]Embrapa LABEX US Plant Biotechnology, Plant Gene Expression Center-ARS/USDA, Albany, California, USA

ABSTRACT

Rhythms produced by the endogenous circadian clock play a critical role in allowing plants to respond and adapt to the environment. While there is a well-established regulatory link between the circadian clock and responses to abiotic stress in model plants, little is known of the circadian system in crop species like soybean. This study examines how drought impacts on diurnal oscillation of both drought responsive and circadian clock genes in soybean. Drought stress induced marked changes in gene expression of several circadian clock-like components, such as *LCL1*-, *GmELF4*- and *PRR*-like genes, which had reduced expression in stressed plants. The same conditions produced a phase advance of expression for the *GmTOC1*-like, *GmLUX*-like and *GmPRR7*-like genes. Similarly, the rhythmic expression pattern of the soybean drought-responsive genes *DREB*-, *bZIP*-, *GOLS*-, *RAB18*- and *Remorin*-like changed significantly after plant exposure to drought. *In silico* analysis of promoter regions of these genes revealed the presence of cis-elements associated both with stress and circadian clock regulation. Furthermore, some soybean genes with upstream ABRE elements were responsive to abscisic acid treatment. Our results indicates that some connection between the drought response and the circadian clock may exists in soybean since (i) drought stress affects gene expression of circadian clock components and (ii) several stress responsive genes display diurnal oscillation in soybeans.

Keywords: gene expression, water deficit, dehydration, circadian rhythms, diurnal oscillation.

Introduction

Plants are subjected to diurnal oscillations due the planet movement around its axis, which generates light and temperature variations. In addition to the normal day/night variations, plants are subject to other environmental variations via biotic and abiotic stresses. Drought-induced water deficit greatly affects plant development, and in crop species this is damaging for agronomic productivity. Drought stress leads to a number of molecular and physiological changes in plants that protect against water deficit. Signal transduction molecules play important roles in this process by mediating the transmission of the stress signals via complex signal transduction pathways. In *Arabidopsis*, the molecular drought response mechanism can be divided into abscisic acid (ABA)-dependent and ABA-independent pathways [1]. The ABA-dependent signal transduction pathway comprises the ABA-bound pyrabactin resistance/regulatory component of ABA receptor (PYR/RCAR) proteins, type 2C protein phosphatases (PP2C), and SNF1-related kinases (SnRK2) [2]. Additionally, the H subunit of the magnesium-protoporphyrin IX chelatase protein (ABAR/CHLH/GUN5) has also been described as ABA receptor in *Arabidopsis* under stress conditions [3]. ABA-mediated signal transduction leads to the activation of transcription factors, such as basic leucine zipper (bZIP) proteins [4]. In contrast, the ABA-independent pathway involves ethylene signaling and the participation of transcription factors, primarily from the ethylene-responsive factor (ERF) and C-repeat-binding factor/dehydration-responsive element-binding (CBF/DREB) subfamilies [1,5]. In both the ABA-dependent and ABA-independent pathways, transcription factors bind to specific *cis*-elements and induce several stress-responsive genes that encode different protein classes, including galactinol synthase (a key enzyme for the biosynthesis of the osmoprotectant molecule raffinose) [6], RAB18 (a protein involved in membrane vesicle transport) [7], Remorins (membrane structural proteins) [8], and peroxidases (ROS-scavenging proteins) [9].

The circadian clock is an endogenous timer that plays a key role in the coordination of plant biological activities with diurnal variations, conferring adaptive advantages to organisms. In this system, environmental cues like light and temperature play important role in entrain the circadian clock responses. Recent studies have demonstrated a correlation between the circadian clock and plant responses to drought, suggesting a close connection between both pathways [10,11].

The more recent existing model of the circadian clock in plants comprises interlocked feedback loops, which includes a ring of three sequential negative steps: (a)the inhibition of evening complex (EC) genes (ELF3, LUX, and ELF4) by the rise of LHY/CCA1 in the late night, (b) the inhibition of PRR genes by EC in the early night, and (c) the inhibition of LHY/CCA1 by PRRs in the day [12]. The evening complex is suggested to represent the structure of the evening loop [12]. The two partially redundant morning Myb-like transcription factors CCA1 and LHY regulate the expression of TOC1 and GI. TOC1/PRR1 is a member of the clock-specific transcription factor family of pseudo-response regulator (PRR) proteins, and GI is a vascular plant-specific protein with a poorly understood molecular function. CCA1 and LHY reach peak levels in the morning to repress the daytime expression of TOC1 and GI. CCA1 and LHY directly bind to the TOC1 promoter at a 9 bp target sequence referred to as the evening element (EE; AAATATCT), and this interaction suppresses transcription at this promoter [13,14]. Recently, Pokhilko and colleagues extended the inhibitory action of LHY/CCA1 to all evening genes (TOC1, LUX, ELF4, ELF3 and GI) [12]. TOC1 appears to repress CCA1 and LHY expression inside the morning loop [12,15], while the EC indirectly activate CCA1/LHY expression, by suppressing the expression of CCA1/LHY’s repressors [12,16–18]. Two additional loops maintain the adequate expression of CCA1/LHY and TOC1. In the morning, PRR5, PRR7 and PRR9 are activated through the activity of CCA1/LHY, and in turn PRR5, PRR7 and PRR9 feedback repress the expression of CCA1/LHY [19,20]. In a posttranslational loop, the F-box protein ZTL targets TOC1 for degradation [21]. GI controls the activity of ZTL, and the physical interaction between GI and ZTL (light dependent) stabilizes both ZTL and TOC1 expression during the day [21]. According to the most recent circadian clock model in Arabidopsis, GI increases TOC1 expression by the inhibition of the EC, which is a negative regulator of TOC1 expression [12]. Together, the three interlocking feedback loops ensure that the clock produces robust and accurate rhythms.

Although our present knowledge of the circadian clock suggests conservation among plant species, studies of clock function and molecular architecture in plants other than *Arabidopsis* are limited [22]. Soybean is one of the most well-studied crops in terms of its genetics and related molecular behavior under many circumstances; however, there is a lack of information concerning the behavior of the soybean circadian oscillator. Some orthologs to the *Arabidopsis* circadian clock genes, have been identified in the soybean genome and have been shown to oscillate in a manner similar to that in *Arabidopsis* in controlled situations [23–26]. Previous studies in Arabidopsis have shown connections between the plant responses to abiotic stresses (eg. heat, cold and drought) and the diurnal oscillations or the circadian clock [10,11,27,28]. However, to date the behavior of the soybean clock components in response to environmental stresses, like drought, was not investigated. Environmental cues like light, temperature and abiotic stresses can act as inputs that modulate the circadian clock, ensuring the precise synchronization of important plant molecular processes [29]. For crop species, such as soybean, flowering is a key component of productivity, whereby precise synchronization using environmental cues maximizes the number of flowers and pods produced and, consequently, increases the yield. Understand the impact of drought on the circadian clock components is of great interest, once the drought imposition can act as an environmental cue to the clock and the processes it controls.

Material and methods

1. Plant material, growth conditions, and treatment application

The seeds from plants of the BR16 genotype, which exhibit drought-sensitive characteristics [30], were cultivated in peat pots (Jiffy) with Supersoil® (Scotts Miracle-Gro Company, Marysville, Ohio, USA). The plants were grown in growth chambers set to simulate environmental conditions: 14 h light/10 h night cycles, with 500 μmol $m^{-2}s^{-1}$ of white light (provided by cool white fluorescent bulbs), with 28°C/20°C temperature cycles during the light and the dark period, respectively. Fifteen days after germination, when the plants reached the V2 developmental stage [31], water was withheld in the stress treatments to induce a water deficit. The soil moisture was calculated by the gravimetric humidity (GH), which corresponds to the percentage of water in the soil in relation to the dry weight of the soil. The volume of irrigation was adjusted to 70% (GH) (near field capacity) for the unstressed treatment, 30% GH for the moderate stress treatment, and 15% GH for the severe stress treatment. The pots were weighed twice a day, and water was added to maintain the treatments at the desired GH values. Sampling was initiated when 30% and 15% GH were obtained for the moderate and severe stress treatments, after 3 and 5 days of water withholding, respectively. To overcome the differences due developmental stage differences between plants from the 3rd and 5th days of water withholding, control plants (maintained at 70% GH) were collect for each day. In other words, control and stressed plants were age-matched. Fully expanded V1 leaves were collected from the 3 plants in each treatment at 4 h intervals from the time the lights came on (Zeitgeiber Time (ZT) 0), and were immediately frozen in liquid N_2 and stored at −80°C until further use. The samples obtained in the dark were collected with the aid of a small green LED light (PhotonLight.com).

For the ABA treatment, the BR16 plants were grown under the same conditions as described above. Fifteen days after germination (V2 developmental stage, [31]), ABA (MP Biomedicals; Santa Ana, California, USA) was sprayed directly onto the plants using 4.5 mL of a mixture of 100 mM ABA and 0.01% (v/v) Triton X-100 in water, in accordance with the methods of Legnaioli et al. (2009) [10]. The untreated plants received a mock treatment of 0.01% (v/v) Triton X-100 in water. The ABA solution was applied at ZT3 (11:00 am), and the samples were collected at the indicated time points and analyzed.

2. Real-time qPCR analysis of gene expression

The gene expression was monitored for the samples subjected to moderate and severe stress (30% and 15%GH, respectively) compared to unstressed treatment situation (70% GH), using quantitative real-time PCR (qPCR). A similar approach was used to compare samples subjected to 100 mM ABA treatment and the control plants (not treated with ABA). All of the experiments were completed with three biological replicates, consisting of two plants collected together and pooled. Each replicate tissue set was ground to a fine powder in liquid nitrogen, and the total RNA was isolated using the Concert Plant RNA Reagent (Life Technologies, Grand Island, NY, USA) according to the manufacturer's instructions. The contaminating DNA in the total RNA was removed using the Turbo DNA-free kit according to the manufacturer's instructions (Ambion by Life Technologies, Grand Island, NY, USA). The high-quality total RNA was used to analyze the transcripts in each treatment at 4 h intervals from the time the lights came on (Zeitgeiber Time (ZT) 0). The first-strand cDNA was generated using the Maxima Universal first-strand cDNA synthesis kit (Fermentas/ Thermo Fisher Scientific Inc., Waltham, MA, USA), according to the manufacturer's instructions and diluted five times with water, and 2 µL was used for the qPCR using a CFX Real-Time PCR Detection System (Bio-Rad, Hercules, CA, USA), with in two technical replicates. The qPCR reactions contained 1X EvaGreen Dye (Biotium, Hayward, CA, USA), 1X ExTaq buffer (Clontech, Mountain View, CA, USA), 200 µM each dNTP (New England Biolabs, Ipswich, MA, USA), 300 µM each primer, 0.05 mg/mL BSA (New England Biolabs, Ipswich, MA, USA), 0.1% Tween-20, and 5% DMSO mix.

The normalized cycle threshold (Ct) values were calculated based on the geometric average of two endogenous genes (Elongation factor 1-β, Glyma13g04050, and β-actin, Glyma15g05570) previously identified as stably expressed under many developmental stages and stress situations [32,33]. The normalized expression (NE) was calculated using the following formula: $NE=2^{\wedge -(\Delta Ct)}$, where ΔCt= Ct experimental – Ct normalizer (n). The primers for the *Glycine max*-like clock genes were designed using soybean Glyma ortholog sequences from Phytozome *G. max* v.1.0 (see Table 1). The soybean drought-responsive genes were chosen according to previous work confirming the upregulation of these genes under conditions of dehydration [34,35]. The primer sequences (Table 1) were designed using the PrimerQuest tool (Integrated DNA Technologies, Coralville, IA, USA), employing the sequence from the 3' untranslated region with the default settings. Gene expression was evaluated in 3 independent biological replicates, each one evaluated in 2 technical replicates. To compare gene expression between control and stressed/treated plants along a time course we performed statistical analysis using ΔCt values, as described by Yuan and colleges [36]. After performing descriptive and exploratory data analysis, the effects of water regimes (control, moderate and severe drought) or ABA treatment, day time (ZT0 to ZT20), and their interactions on determined variables were analyzed using ANOVA. When results from the overall significance test led to rejection of the null hypothesis, Duncan's multiple range test for multiple comparisons among groups tests (5%) was performed. The above analyses were conducted using SAS 9.2 software (SAS Institute, Cary, NC).

3. Gene expression analysis by RNA-seq

The soybean transcriptome was analyzed by RNA-seq in samples subjected to moderate stress (30% GH) compared to unstressed treatment situation (70% GH). After DNase treatment (Life Technologies, Grand Island, NY, USA), high-quality total RNA was used to analyze the transcripts for each time point: 8, 12, 16, 20, 24, and 4 h. Bulks of leaves from two plants were used in the RNA extraction to compose one replication. Three replications for each time point/treatment were sequenced. The RNA-seq libraries were built using the Nugen-Ovation® kit according to the manufacturer's instructions (NuGEN Technologies Inc., San Carlos, CA, USA). The libraries obtained were subjected to sequencing by Illumina HiSeq2000 (Illumina, San Diego, CA, USA). Mapping of the reads was performed with the Soybean genome (Phytosome Glycine max v1.1) using the GeneSifter platform (http://www.geospiza.com/Products/AnalysisEdition.shtml). To compare gene expression between different times and conditions, we $\log_2$-transformed the normalized reads per mapped million (RPM) value. Data were analyzed using ANOVA to evaluate the effects of water regimes (control and moderate drought), time point (ZT0 to ZT20), and their interactions. We performed Tukey's HSD multiple comparison tests (95% family-wise confidence level) to show the interactions between water regimes (control and moderate drought) and time points, whenever these interactions were significant in ANOVA analyses. The RNA-seq analyses were performed using the GeneSifter Analysis Edition platform (GSAE; a registered trademark of Geospiza, Inc.) [37].

4. Identification of soybean circadian clock genes

To identify homologs of the *Arabidopsis* circadian clock genes in the soybean genome, the amino acid sequences of the corresponding *Arabidopsis* proteins were used as queries in BLAST searches (TBLASTN tool) [38] in the *G. max* genome v1.0 using the Phytozome database (http://phytozome.net/soybean). The most similar sequences were selected on the basis of whether they had alignment e-values close to 0, using the cutoff e-value 1e-20. The phylogenetic tree construction was based on the alignment of the amino acid sequences using ClustalW and the generation of trees using the Neighbor-Joining method [39] in the MEGA 5.0 software [40]. The percentage of replicate trees in which the associated taxa were clustered together in the bootstrap test (1000 replicates) [41] are shown next to the branches. The tree was drawn to scale, with the branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [42] and are presented as the number of amino acid substitutions per site. All of the ambiguous positions were removed for each sequence pair (pairwise deletion) [43].

5. Identification of soybean circadian clock genes and cis-elements in the promoter regions of circadian clock and drought-responsive genes

The putative *cis*-regulatory elements in the promoter regions of the soybean genes were identified using a suite of Genomatix programs (http://www.genomatix.de, Genomatix Software, Munich, Germany) [44] (Tables S1 and S2). First, the Gene2promotor component was used to define the promoter region, which encompassed 500 bp upstream of the transcriptional start site and 100 bp downstream of the transcriptional start site into the presumed 5' untranslated region. Second, each 601 bp sequence was examined for putative target sequences of well-established plant transcription factors using MatInspector Version 4.3 [44]. The MatInspector program search for *cis*-elements was based on a “Matrix” consisting of a position weight matrix, a conservation profile (conservation index vector), and a core region for a set of training sequences. The “matrix similarity score” reflects the similarity between the obtained sequence and the matrix sequence, whereby a value of one corresponds to sequences with the most conserved nucleotide at each position in the matrix. Mismatches in the highly conserved positions of the matrix decrease the matrix similarity more than mismatches in the less-conserved regions [44,45]. We selected a threshold matrix similarity of >0.80 based on previous studies using sequence-training sets, which indicates that a consistent match to the matrix has a similarity of >0.80 [44].

Results and Discussion

1. Drought stress induces significant changes in the diurnal expression of putative circadian clock genes

To characterize the effects of drought stress on the transcriptional regulatory networks of the soybean circadian clock effectors, genes encoding probable circadian clock components were identified through in-house alignments of the amino acid sequences of core *Arabidopsis* circadian clock proteins to the *G. max* genome. We identified potential soybean homologs of TOC1, CCA1/LHY, LUX, ELF4, CHE, GI, ZTL, and several PRR-like proteins (Table 1).

The soybean genome encodes at least two *LHY* and *CCA1*-like homologs named *LCL1* and *LCL2* (*LHY/CCA1*-like 1 and 2), however it is not possible distinguish which gene (LCL1 or LCL2) corresponds to the Arabidopsis LHY or CCA1 [23]. In this study we evaluated specifically *GmLCL1* gene expression. The PRR proteins are highly similar in amino acid sequence, and the soybean genome is highly redundant [46]; therefore, a phylogenetic analysis was used to identify the most probable homologs for each PRR class (Fig. S1). In total, 11 genes encoding potential circadian clock genes were identified for our analysis (Table 1).

Table 1. *Arabidopsis* and soybean ortholog genes.

Gene name	Arabidopsis gene	Soybean ortholog gene	Soybean name	TBLASTN (e-value)	Identity (%)	Forward primer (5'->3')	Reverse primer (5'->3')
TOC1	AT5G61380	Glyma04g33110	GmTOC1-like	5.2e-52	40.3	TGACATAAGGATGAAGGGCCAACC	TGAGGGCGCATATTGGATCAACAC
PRR7	AT5G02810	Glyma10g05520	GmPRR7-like	8.3e-38	59.1	GGCAACAATTCTGGCACCACCTAA	GCAGCTGATGCTTCATGTTGTCAG
PRR9	AT2G46790	Glyma06g14150	GmPRR9-like	2.6e-36	67.7	CCCGAATCCTTAAATACCAGAAGCAC	CACGATCTACAGAAAGGGCAAATG
PRR3	AT5G60100	Glyma11g15580	GmPRR3-like	4.3e-36	89.1	TGATGTCATCTCATGATTCTATGGGT	ACTCACACTGTGGCATCTTCTCCA
CCA1	AT2G46830	Glyma07g05410	GmLCL1-like	1.6e-21	40.3	ACCATAGGGCTTGGACAAGGAAAG	ACCTTGATTGTTGCTCGCTCCAAC
LHY	AT1G01060						
ZTL	AT5G57360	Glyma09g06220	GmZTL-like	0	88.7	GCATGCTGTAGCAAGGGAAATGCT	CTGACCAGAGCAATACTCGTCAAG
GI	AT1G22770	Glyma20g30980	GmGI-like	0	82.2	GTGGCAGATGGCCTTTCAAACCTT	CGGACATGTGCACTTGGATGAGAA
LUX	AT3G46640	Glyma12g06410	GmLUX-like	5.6e-37	90.2	GAACCTAAGGTCAGCAGCAATCAC	TCAATTCGATCTCCTGCCAAATGC
ELF4	AT2G40080	Glyma18g03130	GmELF4-like	4.1e-26	73.1	ATTCAGCAGGTGAACGAGAACCAG	ACAACCTTGGAGATGTTGCCGTTG
CHE	AT5G08330	Glyma20g00350	GmCHE-like	3.2e-28	59.9	TATTGTGTTTGTCGGTGGGTGGGT	AGTCCTTCTCCTTGTCCACACACA
JUMONJI	AT3G20810	Glyma11g13910	GmJumonji-like	1.3e-36	58.5	TTTGGCACTCGTTGTCACTACACG	TACTGTTCGGGACTGCGTTTCACA
ABAR	AT5G13630	Glyma19g32070	GmABAR-like	0	88.7	AGAGAAGAGCAGCATCCTTCA	TTCAGAACTGCACAAACGAGA
REMORIN	AT2G41870	Glyma19g32280	GmRemorin-like	8.6e-33	47.1	TGGATTGCAGTAAGCAGCAC	AGCGTGACACCACTTATCACA
GOLS	AT2G47180	Glyma19g40680	GmGOLS-like	2.0e-73	79.4	ACGGGGAAGGAAGAGAACAT	TGCACTCATCAATGGCTTGT
DREB1	AT1G46768	Glyma14g09320	GmDREB1-like	5.5e-39	55.9	GATGATGATGCCTCGGAGTTG	CGGAAAAACAAGAAAAGGGATATATC
DREB2	AT4G39780	Glyma05g31370	GmDREB2-like	2.0e-46	47.5	GGCTGCTTCTGCAATGGATT	GACCACTACGACCCTCTCTGATTT
DREB3	AT1G22190	Glyma13g01930	GmDREB3-like	1.2e-41	44.3	TTGCTTATTGGCTATTCGATGGT	TCCATGGCCAAGCAAGAAA
RAB18	AT5G66400	Glyma09g31740	GmRAB18-like	1.0e-12	36	CAACTGGTGGCACTGGTTATGG	TGGTCATGCTGACGATGTTCCT
bZIP	AT3G19290	Glyma02g14880	GmbZIP	5.4e-45	42.8	TAATGGGAATGGGAATTTGGG	GTTGGTGTTGGTGTTGGTGTTGTG
PP2C	AT1G07160	Glyma14g37480	GmPP2C-like	2.7e-80	70.4	GCTATGTTGATTTATGCCGTGGTG	ACTTTGGTCTCAGGCTCTGCTGTCA
SnRK2	AT4G33950	Glyma02g15330	GmSnRK2-like	6.7e-158	92	CAAAGTGATCTCATGGATGGGA	TGCTATCTAAGTCAAGGTCAGGATC

The **Arabidopsis gene** identification using the TAIR database and the **soybean orthologous** gene identification using the Phytozome database are shown.

The **forward** and **reverse primer** sequences correspond to the oligonucleotides used to amplify the soybean orthologs.

The TBLASTN e-value and **Identity** correrspond to the local alignament between the Arabidospsis protein and the soybean translated genome at Phytozome database

The great plasticity exhibited by plant responses to drought calls for studies that simulate field conditions in order to obtain information on the plants responses under more natural conditions. A recent study showed that molecular analyses under constant environmental conditions (eg. constant light, temperature) may give partial or misleading indications of the plant responses to the natural fluctuating conditions [47]. Therefore, studies showed that light and temperature cycles entrain the clock and thereby ensure appropriate phasing of circadian rhythms [18,48]. In this context, to assess the impact of drought stress on the soybean circadian clock components, the expression profiles for each candidate soybean circadian clock gene were determined in V2-stage plants under simulated field light (14h light/10h dark) and temperature (28ºC/20ºC) cycles, over a 24 h time course under well-watered conditions (control) or one of two water-limited (stress) conditions. Prior to the implementation of the stress treatments, all of the plants were maintained under well-watered conditions, which corresponded to soil at 70% GH. After reaching the V2 stage of development, in accordance with the definition of previous study [31], irrigation was removed for the water-limited plants. The moderate stress condition was obtained when the soil achieved 30% GH, corresponding to 3 days of water withholding, whereas the severe stress treatment corresponded to a GH of 15%, achieved with 5 days of water withholding. To overcome possible differences in gene expression due disparity in plants developmental stages we compared age-matched stress and control plants, since authors have demonstrated changes in gene expression between plants in different developmental stages [49,50]. As we can see in Figures 1 and 2, for some genes we observed differences in the expression of control plants from the moderate and severe stress, which points the importance of using age-matched plants for comparisons between control and stress.

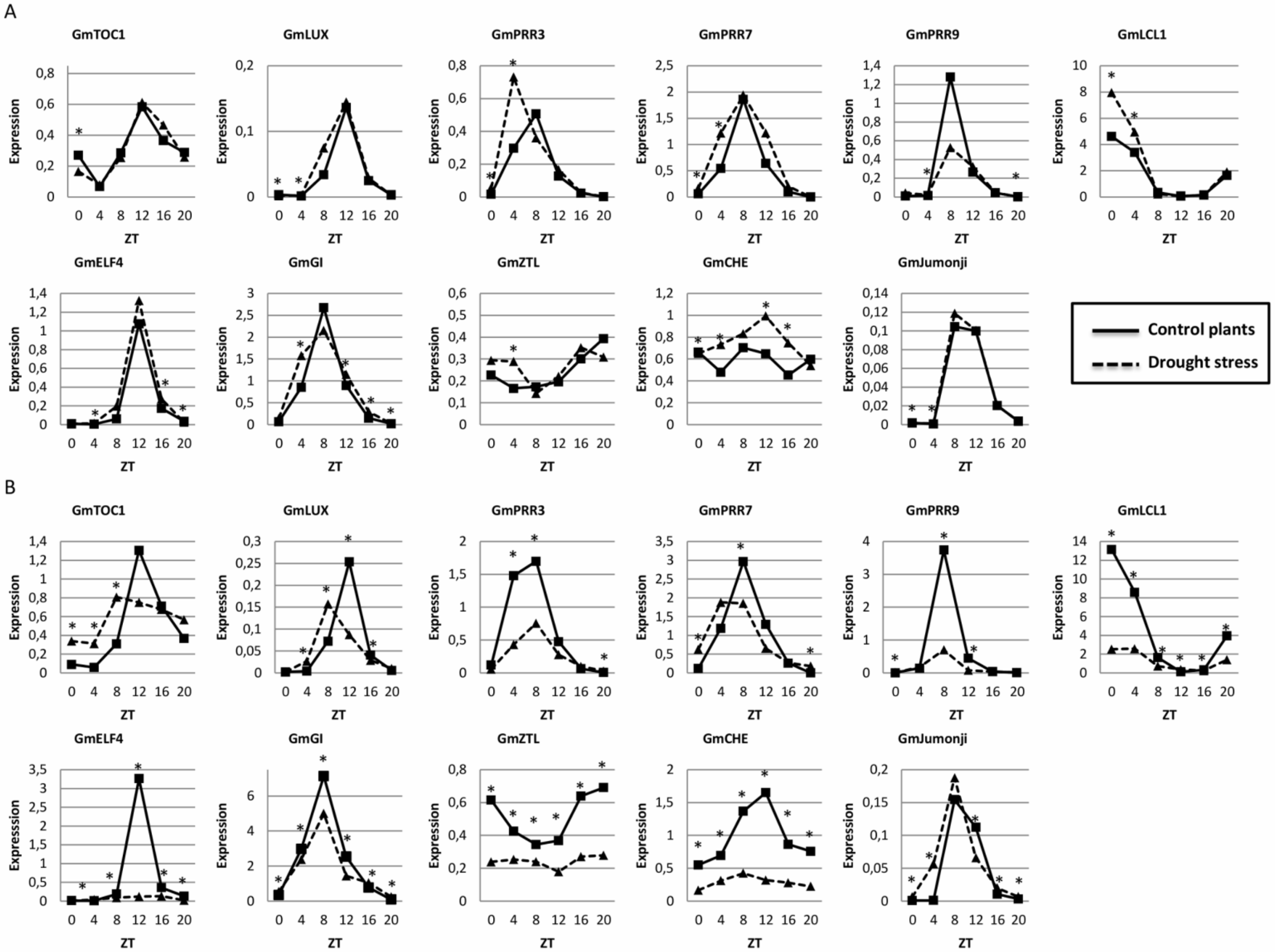


Fig. 1. Drought affects the expression of some circadian clock genes in soybean. Gene expression data regards qPCR analysis of soybean leaves during moderate (A) and severe (B) drought stress. Expression axis represents normalized expression (NE)=$2^{\wedge -(\text{Ct experimental} - \text{Ctn})}$. Collect time points are represented by ZT (Zeitgeiber Time) 0 to 20, starting from the time the lights came on (ZT0) and proceeding with 4 h intervals until ZT20. For easy viewing, asterisks represent significant differences between control and stressed plants in each time point (Duncan's test 5%, *time -treatment* interaction). The ANOVA and the complete Duncan's test results are presented in Table S3.

To compare the qPCR gene expression between control and stressed plants we performed ANOVA statistical analysis, followed by Duncan's test. In this analysis, we used the ΔCt values for treatment and control samples. This approach ensure the adjustment of additive effects of concentration, gene, and replicate variations since the ΔCt values are obtained subtracting Ct number of target gene from that of reference/normalizer genes [36]. Statistical results for ANOVA and Duncan's tests are presented in Table S3.

Under normal well-watered conditions, the potential soybean circadian clock genes showed expression patterns compatible with the circadian oscillation described in literature for soybean [23,24,26] and Arabidopsis [12,51–53]. On the other hand, the plants exposed to the moderate and severe drought stress conditions showed significant changes in the diurnal expression of several potential core circadian clock genes. The expression of *GmTOC1*-like under normal hydration conditions was compatible to previous circadian oscillation reports [23,26] with transcript levels for this gene peaked during the transition from day to night at the time corresponding to ZT12 (Fig. 1). The moderate stress condition did not affect the timing of *GmTOC1*-like expression (Fig. 1A), whereas lower amplitude of *GmTOC1*-like peak expression was observed 4 hours earlier under severe stress conditions (Fig. 1B). The evening-expressed *GmLUX*-like also displayed advanced-phase expression under severe hydration stress (Fig. 1B). Similar to *TOC1*, the *LUX* gene expressed in the evening were reported to be direct regulatory targets of CCA1 and LHY in *Arabidopsis* [19,54]. The advance in the peak expression of these evening genes was similar to that observed in *Arabidopsis* mutants lacking both the CCA1 and LHY functions [53]. The early phase expression observed in the *cca1 lhy* double mutants reflected the absence of the normal repression activities of these two transcription factors, suggesting that drought stress conditions modify the regulation of soybean *LHY/CCA1*-like gene expression.

To examine whether drought stress influenced the morning loop components of the soybean circadian clock, the expression of the *GmLCL1*-like gene was examined under conditions of drought stress. The expression of *GmLCL1*-like under conditions of normal hydration was comparable to that of its *Arabidopsis* ortholog (Fig. 1), with the peak expression of this gene being observed at dawn (ZT0). Strikingly, *GmLCL1*-like expression was substantially reduced in those plants experiencing severe dehydration stress (Fig. 1B). Therefore, the advancement of evening-phased gene expression under severe drought stress might reflect, at least in part, the loss of normal *GmLCL1*-like expression.

The lack of normal *GmLCL1*-like expression under drought conditions might also negatively affect the expression of *GmPRR9*-like and *GmPRR3*-like, as the transcript levels of these genes were markedly lower in the plants exposed to severe drought stress (Fig. 1B). The two *Arabidopsis* genes that participate in the morning loop, *PRR9* and *PRR7*, are directly upregulated through CCA1 and LHY [19]. Indeed, the genes we identified as *GmPRR9*-like and *GmPRR3*-like could potentially represent these components in the soybean clock system.

Although drought stress potentially influences Gm*LCL1*-like regulation, a potential contributor to the strong reduction in *LCL1* expression is the severe drought-induced loss of a *GmLCL1*-like activator. *GmELF4*-like is a putative activator of *GmLCL1*-like because *Arabidopsis ELF4* is required for the phytochrome-mediated light induction of *AtCCA1* and *AtLHY* [18]. Consistent with this idea, the *GmELF4*-like expression was low in the plants exposed to severe drought stress (Fig. 1). In *Arabidopsis*, ELF4 is repressed through the action of a multiprotein complex containing CCA1 and LHY and the transcription factors FAR1 and FHY3 [18,55,56]. FAR1 and FHY3 also activate ELF4 when CCA1 and LHY are absent [56]. Thus, severe drought stress could negatively affect the function of the soybean orthologs FAR1 and/or FHY3, resulting in limited *GmELF4*-like expression. Fig. S2 shows a proposed scheme for the effect of severe drought on the expression of *GmLCL*-like, *GmTOC1*-like, *GmLUX*-like, and *GmELF4*-like genes.

A general reduction in the amplitude of expression was observed after severe drought stress for most of the clock genes evaluated, including the *GmPRR3*-like, *GmPRR7*-like, *GmPRR9*-like, *GmGI*-like, *GmZTL*-like, and *GmCHE*-like genes (Fig. 1B). This observation is consistent with the effect of other abiotic stresses on the circadian rhythms in plants. Cold stress reduces the amplitude of cycles for clock components and reduces/disrupts the cycles of output genes in Arabidopsis, poplar, and chestnut [27,57,58].

Unlike the strong repression observed under the severe stress condition, the *GmLCL1*-like expression under moderate stress was slightly higher relative to that in the control plants at ZT0 and ZT4, with the most noticeable effect occurring at ZT0 and ZT4 (Fig. 1A). This increase in *LCL1* during early stress stages might be involved in the signaling and activation of drought response mechanisms. CCA1 directs the circadian regulation of the *CBF/DREB* genes in *Arabidopsis* [28], and CCA1 participates in the transcription factor complex that promotes the acute induction of CBF/DREB genes under cold stress [59]. A similar system might be induced in response to drought, as cross-talk between cold and dehydration response pathways is observed.

The *Jumonji* genes are upregulated by drought in soybean [34] and *Arachis hypogaea* [60]. The Jumonji transcription factors act as demethylases, controlling chromatin structure and, thus, gene expression [61]. The silencing of an *A. hypogaea Jumonji* ortholog in tobacco improved drought tolerance [60]. The *Arabidopsis JMJD5* (or *JMJ30*) gene shows circadian oscillation and acts as a regulator of period length [51]. It has been proposed that *JMJD5* and *AtTOC1* act in combination to control the circadian clock and that both genes are repressed through the activity of *AtCCA1* and *AtLHY* [62]. The soybean *GmJMJ*-like gene exhibited strong diurnal expression (Fig. 1). Severe drought treatment induced the upregulation of *GmJMJ*-like expression, particularly at ZT4, and the downregulation of *GmJMJ*-like expression at ZT12, relative to the control plants (Fig. 1B). Whether these changes in expression during drought stress might be important for the regulation of drought responses remains unknown. However, epigenetic mechanisms, such as DNA methylation and histone modification, play a crucial role in regulating gene expression during plant responses to environmental stress [63–66].

For crop species, such as soybean, corn or cotton, productivity is directly connected with flowering intensity/stability to maximize number of seeds produced. Clock genes expression is directly controlled by day length and temperature, interacting with other environmental factors (water status, soil fertility, presence of pathogens, etc) to establish flowering initiation, intensity and duration. Previous studies were able to increase soybean productivity by overexpressing the AtBBX32 transcription factor [67]. They demonstrated that the constitutive expression of AtBBX32 in soybean altered the abundance of transcript levels of the soybean clock genes *GmTOC1 and LHY-CCA1-like2* (*GmLCL2*), and by altering the abundance of circadian clock genes during the transition from dark to light, the timing of critical phases of reproductive development was altered [67]. Flower, pod, and seed number were increased, where the authors proposed that it was caused by changes in the timing of reproductive development in the transgenic soybean that lead to the increased duration of the pod and seed development period. Drought is well known to alter time of flowering in soybean where maintenance/abortion of flowers/seeds will depend on the intensity and duration of the stress. Our results are the first to show in soybean that water deficit alters the expression of circadian clock genes. Thus, in this context, an understanding of the means by which the circadian clock gene expression is altered by this environmental stress may give clues to how we can genetically manipulate the circadian clock to reduce yield losses during drought events.

2. The expression of drought responsive genes oscillates during the day

To evaluate the influence of the time of the day on the expression of soybean drought-related genes belonging to the ABA-dependent and ABA-independent drought response pathways and to determine the effect of drought on these genes temporal-dependent regulation, we analyzed the daily expression of 10 drought-responsive genes by qPCR and RNA-seq. The gene names and the respective orthologs in *Arabidopsis* and soybean are shown in Table 1.To compare the qPCR gene expression between control and stressed plants we also performed ANOVA statistical analysis, followed by Duncan's test, similarly to the analysis of the circadian clock genes, previously described. Statistical results for ANOVA and Duncan's tests are presented in Table S3.

Several studies have demonstrated the circadian oscillation of stress and hormone-responsive genes in *Arabidopsis* [10,11,35,68,69]. Consistent with these data, the drought- and ABA-responsive soybean genes showed diurnal oscillations in gene expression on qPCR data (Fig. 2, Table S3). The diurnal oscillation of these genes could also be confirmed in RNA-seq data, where ANOVA statistical analysis, followed by Tukey's test pointed differences in gene expression along the day (Fig. 3, Table S4).

In previous study, we demonstrated that some genes from the DREB subfamily displayed circadian oscillation under control and stressed conditions [70]. In the present study, we evaluate other DREB genes, to better understand the observed expression patterns and its relation with the soybean daily cycle. As expected, under control conditions, the expression of the *GmDREB*-like genes generally increased just before dawn and reached their peak expression between ZT0 and ZT4, in both, qPCR and RNA-seq analyses (Fig. 2 and 3). Generally, moderate stress induced the expression of *GmDREB1*-like, *GmDREB2*-like, and *GmDREB3*-like but had little effect on the timing of peak expression (Fig. 2).

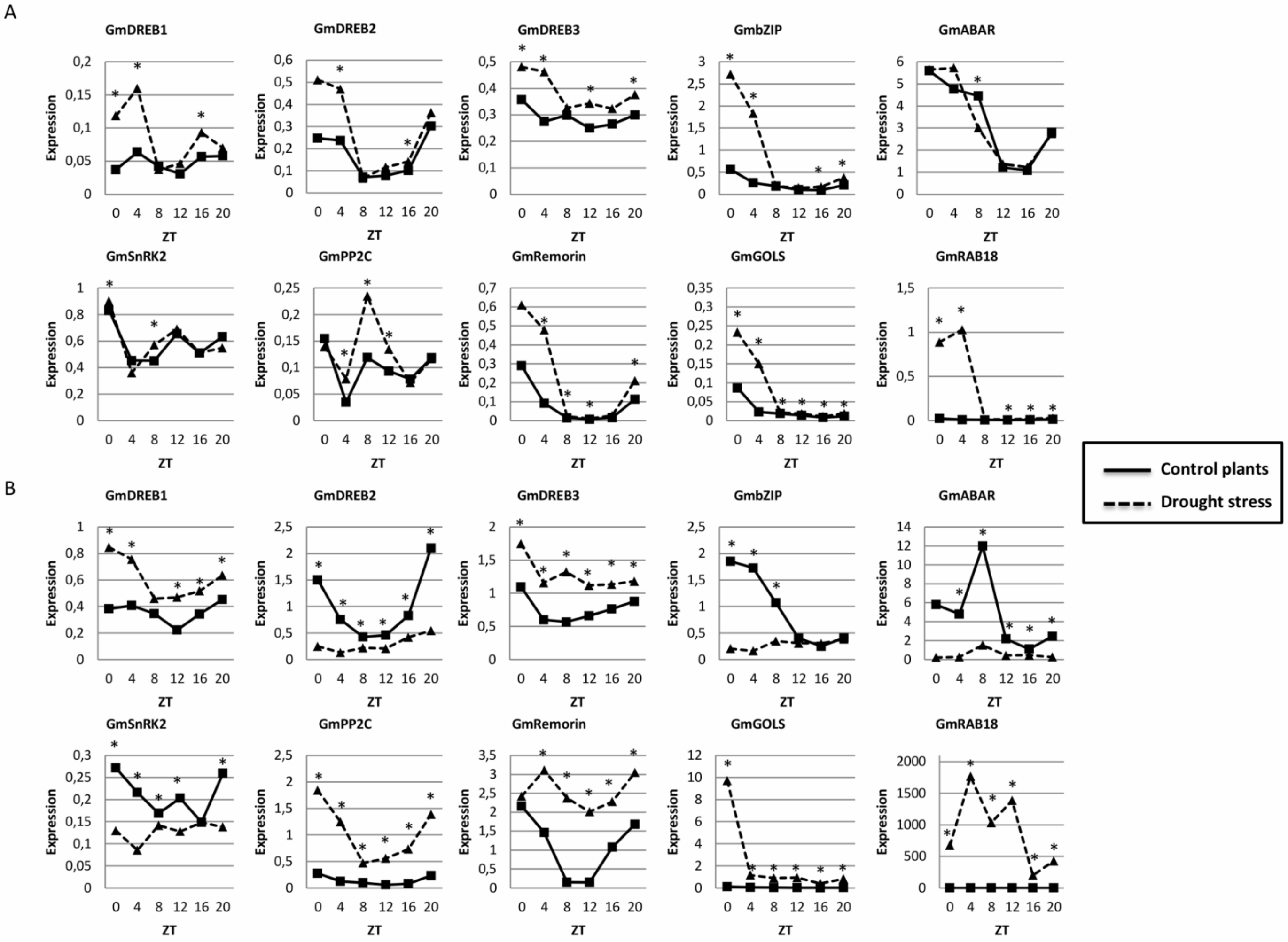


Fig. 2. Drought-responsive genes in soybean exhibit diurnal regulation, and the expression pattern is modified under drought stress. Gene expression data regards qPCR analysis of soybean leaves during moderate (A) and severe (B) drought. Expression axis represents normalized expression (NE)=2^$^{-(Ct}$ $_{experimental\ -\ Ctn)}$. Collect time points are represented by ZT (Zeitgeiber Time) 0 to 20, starting from the time the lights came on (ZT0) and proceeding with 4 h intervals until ZT20. For easy viewing, asterisks represent significant differences between control and stressed plants in each time point (Duncan' test 5%, *time -treatment* interaction). The ANOVA and the complete Duncan's tests results can be found in Table S3.

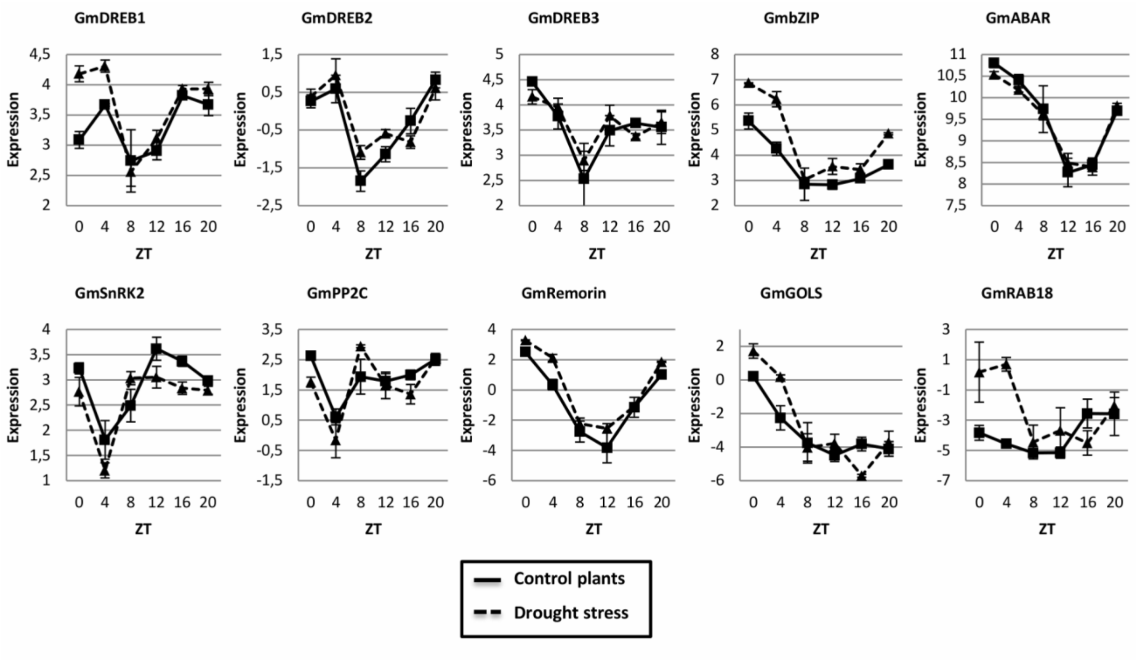


Fig. 3. The diurnal oscillation of drought-responsive genes in soybean leaves during moderate stress is confirmed by RNA-seq analysis. Collect time points are represented by ZT (Zeitgeiber Time) 0 to 20, starting from the time the lights came on (ZT0) and proceeding with 4 h intervals until ZT20. The error bars represent the standard error. The ANOVA and the Tukey HSD (95% family-wise confidence level) multiple comparison tests can be found in Table S4.

Interestingly, the morning-specific induction of these genes was associated with an increase in *GmLCL1*-like expression under the same treatment. The *Arabidopsis DREB* genes are positively regulated through the direct interaction of *CCA1* and *LHY* with the evening element motifs in the DREB gene promoters [28,59]. Additionally, the promoter regions of the *GmDREB*-like genes contain putative evening elements (Table S2); thus, *GmLCL1*-like might play an important role in inducing the expression of the soybean *DREB* genes.

The expression of *GmDREB2*-like suggested that this gene is only involved during the early responses to drought stress. Under moderate hydration stress, the expression of *GmDREB2*-like was induced (Fig. 2A), however, a significant reduction in *GmDREB2*-like expression was observed under severe conditions (Fig. 2B). In contrast, *GmDREB1*-like and *GmDREB3*-like gene expression induction was observed under this treatment (Fig. 2B).

The expression of *DREB* genes under cold stress is suppressed upon the expression of PRR transcription factors that also act in the core circadian oscillator

[59]. Because significant cross-talk occurs between cold and drought responses, these expression patterns indicate similar regulatory relationships in soybean. The peak expression of the *GmPRR*-like genes corresponds with the reduced expression of the *GmDREB1*-like gene (ZT8) (Fig. 1, 2 and 3). Additionally, the reduced expression of some *GmPRR*-like genes was observed under drought stress (Fig. 1), which is consistent with a mechanism for promoting DREB expression by reducing expression of these repressors.

Cold weather and drought affect plant growth and, ultimately, productivity, and many genes respond to both abiotic stresses at the transcriptional level. The DREB genes act as signaling intermediates for both cold and drought responses, and the promoter region of many *DREB*-induced genes contains a dehydration-responsive element, which is a *cis*-acting element that regulates both cold- and dehydration-responsive gene expression in *Arabidopsis* [71]. There is little definitive information regarding the cross-talk between these two different signal transduction pathways; however, the circadian behavior of the *DREB* genes in response to both cold [59] and drought stresses [70] suggests an association with the circadian clock. The circadian clock functions as a key moderator to coordinate metabolism under stress situations to fine-tune the synchronization of global transcription and physiological processes [11,29].

We also evaluated the expression of a *GmbZIP1* transcription factor that belongs to the AREB subfamily and is reported to be an abiotic stress- and ABA-responsive gene [72]. The authors suggest that the overexpression of *ZIP1* in *Arabidopsis* and wheat increases drought tolerance and improves the ABA-mediated control of stomatal aperture in plants. The results obtained in the present study show diurnal oscillation for this gene expression in response to moderate stress at dawn (ZT0) compared with the control treatment (Fig. 2A). In addition, the *GmbZIP1* expression was similar to that of the *GmDREB*-like genes, suggesting a putative mechanism involving transcription factors in plant stress defenses that are activated in the early morning before sunlight.

It has been proposed that the *AtABAR* gene functions as an ABA receptor, playing an important role in stomatal closure in response to drought in *Arabidopsis* [10]. Although it was possible to observe diurnal oscillation in the expression of

ABAR in the present study, we could not detect the upregulation of this gene in response to either drought stress treatment (Fig. 2). The function of ABAR as an ABA receptor has been examined in barley (*Hordeum vulgare* L.), and the authors showed that ABA has no effect on *ABAR/CHLH/GUN5* expression, and ABA binding to the barley protein could not be shown [73].

The *Arabidopsis* ABA signaling system is composed of ABA-bound PYR/RCAR proteins, phosphatases (PP2C) and kinases (SnRK2) and mediates the transmission of the hormone signal [2]. According to this mechanism, ABA binds to PYR/RCAR proteins, releasing SnRK2 from PP2C-induced repression; once activated, the kinases phosphorylate the transcription factors that activate ABA-responsive genes. We observed the upregulation of *GmPP2C*-like and downregulation of *GmSnRK2*-like at the transcriptional level, particularly when comparing the expression levels under severe stress to those under the control conditions (Fig. 2B). The expression of both gene transcripts showed a diurnal oscillation similar to that of *GmDREB*-like and *GmbZIP1*-like in which higher expression levels were observed at predawn hours. The observed regulation of *GmSnRK2*-like and *GmPP2C*-like is consistent with the observations of previous studies on *Vitis vinifera* [74] and *Oryza sativa* [75], respectively.

The *Arabidopsis* orthologs for the *GmRemorin*-like, *GmGols*-like, and *GmRAB18*-like genes used in our study showed diurnal oscillation and ABA upregulation [35]. This is consistent with the results of previous studies of the *Arabidopsis* Remorin ortholog showing the upregulation of gene expression in response to drought [8]. Our results show the upregulation of the *GmRemorin*-like gene under both moderate and severe drought stress (Fig. 2). Interestingly, this gene expression was much higher in response to severe stress and showed a significant diurnal oscillation, with lower expression values between ZT8 and ZT16 (Fig. 2B), whereas the expression differences were not significant under the control and stressed treatments at ZT0. Studies that characterize the function of Remorins are important to understand drought tolerance mechanisms, as Remorins have only been identified in plants [8]. The Remorins have a hydrophilic profile and attach to the plasma membrane. The evolution of the cell wall composition is likely associated with the emergence of different classes of proteins that maintain cell membrane/cell wall integrity and the acquisition of vascular tissue. Thus, it is reasonable to propose that these proteins played an important role during the plant colonization of land in that the chemical adaptation of the cell wall was vital to cope with the particularly rough selection pressure in a dry environment.

The *Gols* gene encodes a key enzyme for raffinose biosynthesis (galactinol synthase), which is an important osmoprotectant associated with defense mechanisms in response to abiotic stresses [6]. The *GmGols*-like expression observed in this study was similar to that of the other drought-induced genes discussed here, showing diurnal oscillation and expression peak just before dawn (Fig. 2 and 3). The *GmGols*-like promoter region also contained an element similar to TBS (Table S2), which is recognized by CHE factors, suggesting a possible interaction between the GmCHE factors and *GmGOLS*-like gene. Additionally, *Arabidopsis AtCHE* has been identified as a putative *AtCCA1* repressor [76], and consistent with this, our data show that the *GmLCL1*-like and *GmGOLS*-like genes have similar expression profiles, with peaks before daybreak (Fig. 1, 2 and 3).

AtRAB18 is a well-known drought- and ABA-responsive gene used in several studies as a marker for drought and ABA treatments [66,77,78]. *RAB18* encodes a protein involved in membrane vesicle transport, and its function has been associated with the recycling of molecules, the removal of existing molecules from cellular compartments, and their replacement with newly synthesized molecules during stress adaptation [7]. The *GmRAB18* gene showed diurnal oscillations under the stress conditions, with higher expression during the day between ZT0 and ZT4, in moderate stress (Fig. 2A and 3), and ZT4 and ZT12 in severe stress (Fig. 2B). The moderate drought stress applied appeared to specifically induce the peak expression of *GmRAB18*-like (Fig. 2A), whereas severe stress caused a significantly higher expression level that was approximately 2000 times greater than that in the control plants (Fig. 2B). The expression profile of the soybean gene suggests that the product of *GmRAB18*-like is important during the day because the moderate stress treatment specifically amplified the daytime expression.

It is important highlight that in the present study we evaluate the diurnal oscillation of drought-responsive genes using plants under simulated field conditions, which means light (14h light/10h dark) and temperature (28°C/20°C) cycles. It was not our intention tell apart the contribution of each factor (circadian clock or environment) to gene expression, but to study their interaction, since in real field conditions such separation does not occur. For example, we could not affirm that the early morning changes in gene expression of drought-responsive genes (*DREB2*-like, *bZIP*-like, *GOLS*-like and others) are exclusively due to light-induction, or due to the circadian clock alone. Previous reports suggest that the circadian clock and the environmental responses contribute to the diurnal molecular responses, suggesting that both components interact to regulate the biological processes in plants [79]. Additionally, as previously discussed, studies about *DREB* genes in Arabidopsis show the circadian oscillation of these genes (in constant light/temperature conditions), with gene expression peak at early morning [28], the same pattern observed in our analysis, making very unlikely the hypothesis that the pattern obtained in our study are only due light changes. Thus, based on our results and on literature evidences, we believe that the expression profiles obtained in our study are the result of the interaction between both circadian clock and environmental cues (light and temperature), which results in the diurnal oscillations.

3.ABA treatment influences the expression of drought response genes and some circadian clock genes

The phytohormone ABA is integral for the response to stressful environments and for the regulation of growth and development [80]. To understand the effects of ABA as a drought stress response component in soybean and to link the activity of this phytohormone with the diurnal expression patterns observed under drought stress, selected soybean genes were characterized for their responsiveness to exogenous ABA treatment. Plants grown under the control environmental conditions were sprayed with a solution of 100 mM ABA just after the start of the light period (ZT3), and gene expression was determined at 4 h intervals thereafter, beginning at ZT4. We selected representative ABA- and drought-responsive genes (*GmRAB18*-like, *GmGOLS*-like, *GmABAR*-like, and *GmbZIP*) as well as several circadian clock genes (*GmTOC1*-like, *GmLUX*-like, *GmLCL1*-like, *GmELF4*-like, *GmPRR3*-like, *GmPRR9*-like, and *GmJMJ*-like) for the analyses. All of these genes contain putative ABA-responsive elements in their promoter regions (Tables S1 and S2).

The candidate ABA- and drought-related genes *GmRAB18*-like, *GmDREB2*-like and *GmbZIP*-like responded to ABA application (Fig.4; Table S5). On the other hand, *GmABAR*-like, pointed as an ABA receptor for some authors, showed no significant change in its expression levels or patterns following the ABA treatment (Fig.S3). Thus, considering that *GmABAR*-like was nonresponsive to both moderate and severe drought stress (Fig. 2 and 3), it is likely that this soybean homolog is not involved in the ABA-mediated responses to drought.

Although we observed that ABA treatment induced ABA- and drought-related genes under our experimental conditions, we could not detect a strong effect on the circadian clock genes evaluated, with the exception of the evening genes *GmJumonji-like* and *GmTOC1-like* (Fig. 4; Fig.S3; Table S5). The *GmTOC1*-like responses to ABA were of particular interest, as exogenous ABA treatment induces the expression of *Arabidopsis TOC1* [10]. Furthermore, *Arabidopsis* TOC1 participates in a feedback loop with ABAR to modulate ABA responsiveness, a process which is involved in drought tolerance [10]. However *GmTOC1*-like expression showed a significant decrease in response to ABA treatment at ZT16 (Fig. 4A), requiring further investigation of this gene responses to ABA.

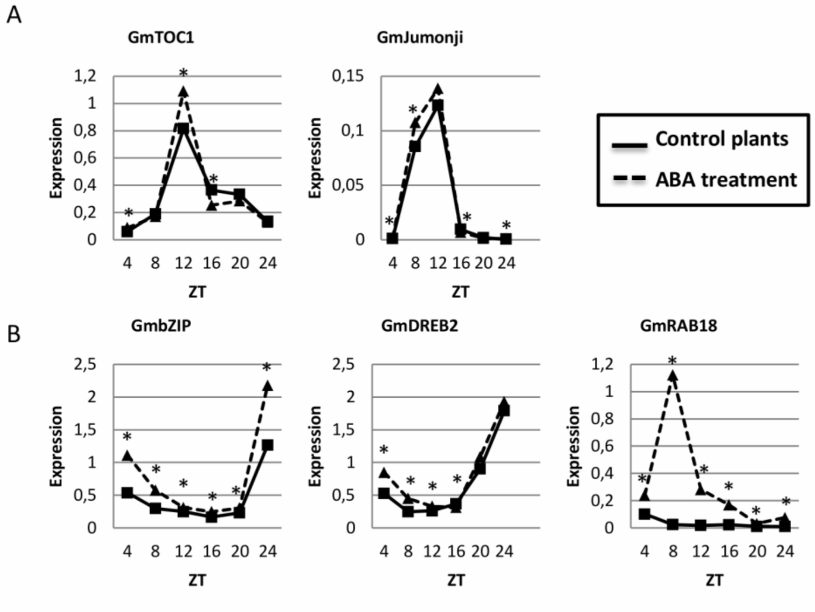


Fig.4. ABA treatment affects the regulation of drought-responsive and some circadian clock genes. Gene expression data regards qPCR analysis of circadian clock (A) and drought-responsive (B) genes. Expression axis represents normalized expression (NE)=$2^{-(Ct\ experimental\ -\ Ctn)}$. Collect time points are represented by ZT (Zeitgeiber Time) 4 to 24, starting 4h after the lights came on (ZT4) and proceeding with 4 h intervals until ZT24. For easy viewing, asterisks represent significant differences

between control and stressed plants in each time point (Duncan's test 5%, *time -treatment* interaction). The ANOVA and the complete Duncan's tests results can be found in Table S5.

4.Conserved cis-elements are located upstream of soybean circadian clock and drought-responsive genes

Diurnal oscillation was evident for both soybean clock and drought-responsive genes, suggesting the circadian regulation of these genes. Transcriptional regulation is expected to occur through conserved circadian and drought-responsive *cis*-elements present in the promoter regions of these genes. Therefore, the presumed promoter regions, corresponding to 100 bp located at the 5' end and the 500 bp located upstream of all 11 circadian clock and 10 drought response genes were analyzed for the presence of known *cis*-elements. Sequence motifs similar to circadian and light *cis*-elements were present in the promoters of several of the drought-responsive genes in our study (Table S2). These elements included potential evening elements (atgaaaAATAtcatc), GAP-box light response elements (taaaATGAagagtag), light-responsive element motifs (tcATCTataca), and TBS elements. These results confirm the gene expression data showing strong diurnal oscillation of the majority of the drought-responsive genes evaluated, and suggests the involvement of the circadian clock in these genes regulation.

In *Arabidopsis*, CCA1/LHY repress *AtTOC1* and *AtLUX* expression by directly binding to the evening element in the promoter regions of these genes [14,54]. The promoter regions of the *GmTOC1*-like and *GmLUX*-like genes reveal the presence of evening elements (Table S1). This result is consistent with the idea that LCL1 is involved in the direct regulation of these genes. This is further confirmed by the early phase expression of *GmTOC1*-like and *GmLUX*-like as well as the reduction of *GmLCL1*-like expression under severe stress conditions (Fig. 1B).

The promoter regions of the soybean circadian clock genes analyzed in this study (except the *GmPRR7*-like gene) contain the *cis*-elements involved in plant responses to abiotic stresses (Table S1). These elements include the dehydration-responsive elements (the binding site for DREB transcription factors), ethylene response elements (targets of AP2/EREBP transcription factors), salt/drought-responsive elements (targets of such stress-responsive elements as zinc-finger proteins), ABA-responsive elements, ER stress response elements, and heat shock elements (targets of heat shock transcription factors). The presence of these stress-responsive *cis*-elements in circadian clock gene promoters supports the idea of gene expression alterations in response to water deficit, and highlights the connection between the plant stress defenses induced to reduce cellular damage and the perception of the day/night environment, possibly through the circadian clock.

Conclusion

Our results show that the drought stress affects the gene expression of circadian clock components in soybeans. We also demonstrate the diurnal oscillation of soybean drought-responsive genes expression. This result explains discrepancies in the gene expression data available in the literature, as previously suggested [27], indicating that the daily expression fluctuations are the primary source of variation between independent experiments [27]. Together our results suggest a possible regulatory interaction between the drought responses and the circadian clock genes in soybean. Interestingly we observed that many of the drought-induced genes associated with the plant defense showed expression profiles with higher intensities before/during dawn under our experimental conditions. It is reasonable consider that may exist interaction between the expression of cell dehydration defense genes and the circadian clock genes to optimize plant metabolism to specific periods of the day, since such mechanism could play a key role in increasing survival and reproductive efficiency in arid environments. Characterizing the components of this mechanism will contribute to the development of genetic engineering strategies to improve drought tolerance in plants for desirable agronomic productivity.

Acknowledgements

We thank Juan Mendoza for technical assistance and the Plant Gene Expression Center, USDA-ARS, Albany, for the greenhouse and laboratory facilities.

References

1. Shinozaki K, Yamaguchi-Shinozaki K. Gene networks involved in drought stress response and tolerance. J Exp Bot. 2007;58: 221–7. doi:10.1093/jxb/erl164

2. Melcher K, Ng L-M, Zhou XE, Soon F-F, Xu Y, Suino-Powell KM, et al. A gate-latch-lock mechanism for hormone signalling by abscisic acid receptors. Nature. Nature Publishing Group; 2009;462: 602–8. doi:10.1038/nature08613

3. Wu F-Q, Xin Q, Cao Z, Liu Z-Q, Du S-Y, Mei C, et al. The magnesium-chelatase H subunit binds abscisic acid and functions in abscisic acid signaling: new evidence in Arabidopsis. Plant Physiol. 2009;150: 1940–54. doi:10.1104/pp.109.140731

4. Weltmeier F, Rahmani F, Ehlert A, Dietrich K, Schütze K, Wang X, et al. Expression patterns within the Arabidopsis C/S1 bZIP transcription factor network: availability of heterodimerization partners controls gene expression during stress response and development. Plant Mol Biol. 2009;69: 107–19. doi:10.1007/s11103-008-9410-9

5. Zhang G, Chen M, Li L, Xu Z, Chen X, Guo J, et al. Overexpression of the soybean GmERF3 gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. J Exp Bot. 2009;60: 3781–96. doi:10.1093/jxb/erp214

6. Santos TB, Budzinski IGF, Marur CJ, Petkowicz CLO, Pereira LFP, Vieira LGE. Expression of three galactinol synthase isoforms in Coffea arabica L. and accumulation of raffinose and stachyose in response to abiotic stresses. Plant Physiol Biochem. Elsevier Masson SAS; 2011;49: 441–8. doi:10.1016/j.plaphy.2011.01.023

7. Agarwal P, Reddy MK, Sopory SK, Agarwal PK. Plant Rabs: Characterization, Functional Diversity, and Role in Stress Tolerance. Plant Mol Biol Report. 2009;27: 417–430. doi:10.1007/s11105-009-0100-9

8. Raffaele S, Mongrand S, Gamas P, Niebel A, Ott T. Genome-wide annotation of remorins, a plant-specific protein family: evolutionary and functional perspectives. Plant Physiol. 2007;145: 593–600. doi:10.1104/pp.107.108639

9. Khanna-Chopra R, Selote DS. Acclimation to drought stress generates oxidative stress tolerance in drought-resistant than -susceptible wheat cultivar under field conditions. Environ Exp Bot. 2007;60: 276–283. doi:10.1016/j.envexpbot.2006.11.004

10. Legnaioli T, Cuevas J, Mas P. TOC1 functions as a molecular switch connecting the circadian clock with plant responses to drought. EMBO J. Nature Publishing Group; 2009;28: 3745–57. doi:10.1038/emboj.2009.297

11. Wilkins O, Bräutigam K, Campbell MM. Time of day shapes Arabidopsis drought transcriptomes. Plant J. 2010;63: 715–27. doi:10.1111/j.1365-313X.2010.04274.x

12. Pokhilko A, Fernández AP, Edwards KD, Southern MM, Halliday KJ, Millar AJ. The clock gene circuit in Arabidopsis includes a repressilator with additional feedback loops. Mol Syst Biol. 2012;8: 574. doi:10.1038/msb.2012.6

13. Harmer SL. Orchestrated Transcription of Key Pathways in Arabidopsis by the Circadian Clock. Science (80-). 2000;290: 2110–2113. doi:10.1126/science.290.5499.2110

14. Alabadí D, Oyama T, Yanovsky MJ, Harmon FG, Más P, Kay S a. Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock. Science. 2001;293: 880–3. doi:10.1126/science.1061320

15. Gendron JM, Pruneda-Paz JL, Doherty CJ, Gross AM, Kang SE, Kay S a. Arabidopsis circadian clock protein, TOC1, is a DNA-binding transcription factor. Proc Natl Acad Sci U S A. 2012;109: 3167–72. doi:10.1073/pnas.1200355109

16. Dixon LE, Knox K, Kozma-Bognar L, Southern MM, Pokhilko A, Millar AJ. Temporal repression of core circadian genes is mediated through EARLY FLOWERING 3 in Arabidopsis. Curr Biol. Elsevier Ltd; 2011;21: 120–5. doi:10.1016/j.cub.2010.12.013

17. Helfer A, Nusinow DA, Chow BY, Gehrke AR, Bulyk ML, Kay SA. LUX ARRHYTHMO encodes a nighttime repressor of circadian gene expression in the Arabidopsis core clock. Curr Biol. 2011;21: 126–33. doi:10.1016/j.cub.2010.12.021

18. Kikis EA, Khanna R, Quail PH. ELF4 is a phytochrome-regulated component of a negative-feedback loop involving the central oscillator components CCA1 and LHY. Plant J. 2005;44: 300–13. doi:10.1111/j.1365-313X.2005.02531.x

19. Farré EM, Harmer SL, Harmon FG, Yanovsky MJ, Kay SA. Overlapping and distinct roles of PRR7 and PRR9 in the Arabidopsis circadian clock. Curr Biol. 2005;15: 47–54. doi:10.1016/j.cub.2004.12.067

20. Nakamichi N, Kiba T, Henriques R, Mizuno T, Chua N-H, Sakakibara H. PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the Arabidopsis circadian clock. Plant Cell. 2010;22: 594–605. doi:10.1105/tpc.109.072892

21. Kim W-Y, Fujiwara S, Suh S-S, Kim J, Kim Y, Han L, et al. ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. Nature. 2007;449: 356–60. doi:10.1038/nature06132

22. Song YH, Ito S, Imaizumi T. Similarities in the circadian clock and photoperiodism in plants. Curr Opin Plant Biol. 2010;13: 594–603. doi:10.1016/j.pbi.2010.05.004

23. Liu H, Wang H, Gao P, Xü J, Xü T, Wang J, et al. Analysis of clock gene homologs using unifoliolates as target organs in soybean (Glycine max). J Plant Physiol. 2009;166: 278–89. doi:10.1016/j.jplph.2008.06.003

24. Sullivan S, Jenkins G, Nimmo H. Roots, cycles and leaves. Expression of the phosphoenolpyruvate carboxylase kinase gene family in soybean. Plant Physiol. 2004;135: 2078–2087. doi:10.1104/pp.104.042762.2000

25. Hudson KA. The Circadian Clock-controlled Transcriptome of Developing Soybean Seeds. Plant Genome J. 2010;3: 3. doi:10.3835/plantgenome2009.08.0025

26. Thakare D, Kumudini S, Dinkins RD. Expression of flowering-time genes in soybean E1 near-isogenic lines under short and long day conditions. Planta. 2010;231: 951–63. doi:10.1007/s00425-010-1100-6

27. Bieniawska Z, Espinoza C, Schlereth A, Sulpice R, Hincha DK, Hannah M a. Disruption of the Arabidopsis circadian clock is responsible for extensive variation in the cold-responsive transcriptome. Plant Physiol. 2008;147: 263–79. doi:10.1104/pp.108.118059

28. Fowler SG, Cook D, Thomashow MF. Low temperature induction of Arabidopsis CBF1, 2, and 3 is gated by the circadian clock. Plant Physiol. 2005;137: 961–8. doi:10.1104/pp.104.058354

29. Sanchez A, Shin J, Davis SJ. Abiotic stress and the plant circadian clock. Plant Signal Behav. 2011;6: 223–231. doi:10.4161/psb.6.2.14893

30. Oya T, Nepomuceno AL, Neumaier N, Renato J, Farias B, Tobita S, et al. Drought Tolerance Characteristics of Brazilian Soybean Cultivars: Evaluation and characterization of drought tolerance of various Brazilian soybean cultivars in the field. Plant Prod Sci. 2004;7: 129–137.

31. Fehr WR, Caviness CE, Burmood DT, Pennington JS. Stage of Development Descriptions for Soybeans, Glycine Max (L.) Merrill1. Crop Sci. 1971;11: 929. doi:10.2135/cropsci1971.0011183X001100060051x

32. Jian B, Liu B, Bi Y, Hou W, Wu C, Han T. Validation of internal control for gene expression study in soybean by quantitative real-time PCR. BMC Mol Biol. 2008;9: 59. doi:10.1186/1471-2199-9-59

33. Stolf-Moreira R, Lemos EG de M, Abdelnoor RV, Beneventi MA, Rolla AAP, Pereira S dos S, et al. Identification of reference genes for expression analysis by real-time quantitative PCR in drought-stressed soybean. Pesqui Agropecu Bras. Pesquisa agorpecuaria brasileira; 2011;46: 58–65.

34. Rodrigues FA, Marcolino-Gomes J, Carvalho JFC, do Nascimento LC, Neumaier N, Farias JRB, et al. Subtractive libraries for prospecting differentially expressed genes in the soybean under water deficit. Genet Mol Biol. 2012;35: 304–14. doi:10.1590/S1415-47572012000200011

35. Mizuno T, Yamashino T. Comparative transcriptome of diurnally oscillating genes and hormone-responsive genes in Arabidopsis thaliana: insight into circadian clock-controlled daily responses to common ambient stresses in plants. Plant Cell Physiol. 2008;49: 481–7. doi:10.1093/pcp/pcn008

36. Yuan JS, Reed A, Chen F, Stewart CN. Statistical analysis of real-time PCR data. BMC Bioinformatics. 2006;7: 85. doi:10.1186/1471-2105-7-85

37. Porter S, Olson NE, Smith T. Analyzing gene expression data from microarray and next-generation dna sequencing transcriptome profiling assays using GeneSifter analysis edition. [Internet]. Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.]. 2009. p. Unit 7.14 7.14.1–35. doi:10.1002/0471250953.bi0714s27

38. Altschul SF, Gish W, Pennsylvania T, Park U. Basic Local Alignment Search Tool. J Mol Biol. 1990;215: 403–410.

39. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4: 406–25. Available: http://www.ncbi.nlm.nih.gov/pubmed/18343690

40. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011;28: 2731–9. doi:10.1093/molbev/msr121

41. Felsenstein J. Confidence Limits on Phylogenies : An Approach Using the Bootstrap. Evolution (N Y). 1985;39: 783–791.

42. Tomii K, Sawada Y, Honda S. Convergent evolution in structural elements of proteins investigated using cross profile analysis. BMC Bioinformatics. BioMed Central Ltd; 2012;13: 11. doi:10.1186/1471-2105-13-11

43. Marsh HW. Pairwise deletion for missing data in structural equation models: Nonpositive definite matrices, parameter estimates, goodness of fit, and adjusted sample sizes. Struct Equ Model. 1998;5: 22–36.

44. Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff a, et al. MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics. 2005;21: 2933–42. doi:10.1093/bioinformatics/bti473

45. Quandt K, Frech K, Karas H, Wingender E, Werner T. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res. 1995;23: 4878–84. Available:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=307478&tool=pmcentrez&rendertype=abstract

46. Severin AJ, Cannon SB, Graham MM, Grant D, Shoemaker RC. Changes in twelve homoeologous genomic regions in soybean following three rounds of polyploidy. Plant Cell. 2011;23: 3129–36. doi:10.1105/tpc.111.089573

47. Jänkänpää HJ, Mishra Y, Schröder WP, Jansson S. Metabolic profiling reveals metabolic shifts in Arabidopsis plants grown under different light conditions. Plant Cell Environ. 2012;35: 1824–36. doi:10.1111/j.1365-3040.2012.02519.x

48. Thines B, Harmon FG. Ambient temperature response establishes ELF3 as a required component of the core Arabidopsis circadian clock. Proc Natl Acad Sci U S A. 2010;107: 3257–62. doi:10.1073/pnas.0911006107

49. Baerenfaller K, Massonnet C, Walsh S, Baginsky S, Bühlmann P, Hennig L, et al. Systems-based analysis of Arabidopsis leaf growth reveals adaptation to water deficit. Mol Syst Biol. 2012;8: 606. doi:10.1038/msb.2012.39

50. Jiang Y, Han YZ, Zhang XM. Expression profiles of a CONSTANS homologue GmCOL11 in Glycine max. Russ J Plant Physiol. 2011;58: 928–935. doi:10.1134/S1021443711040066

51. Lu SX, Knowles SM, Webb CJ, Celaya RB, Cha C, Siu JP, et al. The Jumonji C domain-containing protein JMJ30 regulates period length in the Arabidopsis circadian clock. Plant Physiol. 2011;155: 906–15. doi:10.1104/pp.110.167015

52. Huang W, Pérez-García P, Pokhilko A, Millar AJ, Antoshechkin I, Riechmann JL, et al. Mapping the core of the Arabidopsis circadian clock defines the network structure of the oscillator. Science. 2012;336: 75–9. doi:10.1126/science.1219075

53. Mizoguchi T, Wheatley K, Hanzawa Y, Wright L, Mizoguchi M, Song H-R, et al. LHY and CCA1 Are Partially Redundant Genes Required to Maintain Circadian Rhythms in Arabidopsis. Dev Cell. 2002;2: 629–641. doi:10.1016/S1534-5807(02)00170-3

54. Hazen SP, Schultz TF, Pruneda-Paz JL, Borevitz JO, Ecker JR, Kay S a. LUX ARRHYTHMO encodes a Myb domain protein essential for circadian rhythms. Proc Natl Acad Sci U S A. 2005;102: 10387–92. doi:10.1073/pnas.0503029102

55. McWatters HG, Kolmos E, Hall A, Doyle MR, Amasino RM, Gyula P, et al. ELF4 is required for oscillatory properties of the circadian clock. Plant Physiol. 2007;144: 391–401. doi:10.1104/pp.107.096206

56. Li G, Siddiqui H, Teng Y, Lin R, Wan X, Li J, et al. Coordinated transcriptional regulation underlying the circadian clock in Arabidopsis. Nat Cell Biol. Nature Publishing Group; 2011;13: 616–22. doi:10.1038/ncb2219

57. Ibáñez C, Kozarewa I, Johansson M, Ogren E, Rohde A, Eriksson ME. Circadian clock components regulate entry and affect exit of seasonal dormancy as well as winter hardiness in Populus trees. Plant Physiol. 2010;153: 1823–33. doi:10.1104/pp.110.158220

58. Ramos A, Pérez-Solís E, Ibáñez C, Casado R, Collada C, Gómez L, et al. Winter disruption of the circadian clock in chestnut. Proc Natl Acad Sci U S A. 2005;102: 7037–42. doi:10.1073/pnas.0408549102

59. Dong MA, Farré EM, Thomashow MF. Circadian clock-associated 1 and late elongated hypocotyl regulate expression of the C-repeat binding factor (CBF) pathway in Arabidopsis. Proc Natl Acad Sci U S A. 2011;108: 7241–6. doi:10.1073/pnas.1103741108

60. Govind G, Harshavardhan VT, ThammeGowda HV, Patricia JK, Kalaiarasi PJ, Dhanalakshmi R, et al. Identification and functional validation of a unique set of drought induced genes preferentially expressed in response to gradual water stress in peanut. Mol Genet Genomics. 2009;281: 591–605. doi:10.1007/s00438-009-0432-z

61. Chen Z, Zang J, Whetstine J, Hong X, Davrazou F, Kutateladze TG, et al. Structural insights into histone demethylation by JMJD2 family members. Cell. 2006;125: 691–702. doi:10.1016/j.cell.2006.04.024

62. Jones MA, Covington MF, DiTacchio L, Vollmers C, Panda S, Harmer SL. Jumonji domain protein JMJD5 functions in both the plant and human circadian systems. Proc Natl Acad Sci U S A. 2010;107: 21623–8. doi:10.1073/pnas.1014204108

63. Boyko A, Kathiria P, Zemp FJ, Yao Y, Pogribny I, Kovalchuk I. Transgenerational changes in the genome stability and methylation in pathogen-infected plants: (virus-induced plant genome instability). Nucleic Acids Res. 2007;35: 1714–25. doi:10.1093/nar/gkm029

64. Boyko A, Kovalchuk I. Epigenetic control of plant stress response. Environ Mol Mutagen. 2008;49: 61–72. doi:10.1002/em.20347

65. Choi C-S, Sano H. Abiotic-stress induces demethylation and transcriptional activation of a gene encoding a glycerophosphodiesterase-like protein in tobacco plants. Mol Genet Genomics. 2007;277: 589–600. doi:10.1007/s00438-007-0209-1

66. Wang M, Yu D, Guo X, Cui Y, Li X, Zhang J, et al. Casein kinase 1-Like 3 is required for abscisic acid regulation of seed germination , root growth , and gene expression in Arabidopsis. African J Biotechnol. 2011;10: 13219–13229. doi:10.5897/AJB11.2086

67. Preuss SB, Meister R, Xu Q, Urwin CP, Tripodi FA, Screen SE, et al. Expression of the Arabidopsis thaliana BBX32 gene in soybean increases grain yield. PLoS One. 2012;7: e30717. doi:10.1371/journal.pone.0030717

68. Mikkelsen MD, Thomashow MF. A role for circadian evening elements in cold-regulated gene expression in Arabidopsis. Plant J. 2009;60: 328–39. doi:10.1111/j.1365-313X.2009.03957.x

69. Robertson FC, Skeffington AW, Gardner MJ, Webb AA. Interactions between circadian and hormonal signalling in plants. Plant Mol Biol. 2009;69: 419–27. doi:10.1007/s11103-008-9407-4

70. Marcolino-Gomes J, Rodrigues FA, Oliveira MCN, Farias JRB, Neumaier N, Abdelnoor RV, et al. Expression Patterns of GmAP2/EREB-Like Transcription Factors Involved in Soybean Responses to Water Deficit. PLoS One. 2013;8: e62294. doi:10.1371/journal.pone.0062294

71. Maruyama K, Todaka D, Mizoi J, Yoshida T, Kidokoro S, Matsukura S, et al. Identification of cis-acting promoter elements in cold- and dehydration-induced transcriptional pathways in Arabidopsis, rice, and soybean. DNA Res. 2012;19: 37–49. doi:10.1093/dnares/dsr040

72. Gao S-Q, Chen M, Xu Z-S, Zhao C-P, Li L, Xu H, et al. The soybean GmbZIP1 transcription factor enhances multiple abiotic stress tolerances in transgenic plants. Plant Mol Biol. 2011;75: 537–53. doi:10.1007/s11103-011-9738-4

73. Müller AH, Hansson M. The barley magnesium chelatase 150-kd subunit is not an abscisic acid receptor. Plant Physiol. 2009;150: 157–66. doi:10.1104/pp.109.135277

74. Boneh U, Biton I, Schwartz A, Ben-Ari G. Characterization of the ABA signal transduction pathway in Vitis vinifera. Plant Sci. Elsevier Ireland Ltd; 2012;187: 89–96. doi:10.1016/j.plantsci.2012.01.015

75. Liu B, Fan J, Zhang Y, Mu P, Wang P, Su J, et al. OsPFA-DSP1, a rice protein tyrosine phosphatase, negatively regulates drought stress responses in transgenic tobacco and rice plants. Plant Cell Rep. 2012;31: 1021–32. doi:10.1007/s00299-011-1220-x

76. Pruneda-Paz JL, Breton G, Para A, Kay SA. A Functional Genomics Approach Reveals CHE as a Component of the Arabidopsis Circadian Clock. Nature. 2009;323: 1481–1486.

77. Rushton DL, Tripathi P, Rabara RC, Lin J, Ringler P, Boken AK, et al. WRKY transcription factors: key components in abscisic acid signalling. Plant Biotechnol J. 2012;10: 2–11. doi:10.1111/j.1467-7652.2011.00634.x

78. Ying S, Zhang D-F, Fu J, Shi Y-S, Song Y-C, Wang T-Y, et al. Cloning and characterization of a maize bZIP transcription factor, ZmbZIP72, confers drought and salt tolerance in transgenic Arabidopsis. Planta. 2012;235: 253–66. doi:10.1007/s00425-011-1496-7

79. Rugnone ML, Faigón Soverna A, Sanchez SE, Schlaen RG, Hernando CE, Seymour DK, et al. LNK genes integrate light and clock signaling networks at

the core of the Arabidopsis oscillator. Proc Natl Acad Sci U S A. 2013;110: 12120–5. doi:10.1073/pnas.1302170110

80. Zhu J-K. Salt and drought stress signal transduction in plants. Annu Rev Plant Biol. 2002;53: 247–73. doi:10.1146/annurev.arplant.53.091401.143329

Supporting Information Available at [http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0086402#s5]

Fig. S1: Phylogenetic tree for the PRR genes. The proteins encoded by *Arabidopsis* AtPRR3, AtPRR7, and AtPRR9, the soybean GmPRR homologs, and its paralogs were used to construct the tree using the ClustalW algorithm in the MEGA 5 program. The Neighbor-Joining method was used with the following parameters: Poisson correction, pairwise deletion, and bootstrapping (1000 replicates; random seed).

Fig. S2: Model of the impact of severe drought on circadian clock genes. Model of the impact of severe drought stress on *GmLCL1*, *GmTOC1*, *GmLUX* and *GmELF4-like* gene expression.

Fig. S3: The circadian clock genes that exhibited no response to ABA treatment. Gene expression data regards qPCR analysis. Expression axis represents normalized expression (NE)=$2^{\wedge -(\text{Ct experimental} - \text{Ctn})}$. Collect time points are represented by ZT (Zeitgeiber Time) 4 to 24, starting 4h after the lights came on (ZT4) and proceeding with 4 h intervals until ZT24.

Table S1: Putative *cis*-regulatory elements on circadian clock genes. Putative *cis*-regulatory elements located in the promoter regions of the soybean circadian clock genes.

Table S2: Putative *cis*-regulatory elements on drought-responsive genes. Putative *cis*-regulatory elements located in the promoter regions of soybean drought-responsive genes.

Table S3: ANOVA and Duncan's multiple range test for multiple comparisons among groups (5%) to evaluate the effects of water regimes (control, moderate and severe drought), time point (ZT0 to ZT20), and their interactions. Data regards qPCR analyses.

Table S4: ANOVA and Tukey's HSD multiple comparison test (95% family-wise confidence level) to evaluate the effects of water regimes (control and moderate drought), time point (ZT0 to ZT20), and their interactions. p-values are shown; bold numbers represent significant gene expression differences (p-values ≤0.05). Data regards RNA-seq analyses.

Table S5: ANOVA and Duncan's multiple range test for multiple comparisons among groups (5%) to evaluate the effects of ABA treatment, time point (ZT4 to ZT24), and their interactions. Data regards qPCR analyses.

5. ARTIGO 2: Soybean reference genes for expression analysis in drought stress responses and diurnal oscillations

Artigo a ser submetido no periódico BMC Molecular Biology

Juliana Marcolino-Gomes[a,b], Thiago Jonas Nakayama[a,c], Fabiana Aparecida Rodrigues[a], Renata Fuganti-Pagliarini[a], Rafaela Ribeiro Reis[a,b], Frank G. Harmon[d] and Alexandre Nepomuceno[a]

Institutions

[a]Embrapa Soybean, Brazilian Agricultural Research Corporation , Londrina, Paraná, Brazil
[b]Department of Biology, State University of Londrina, Londrina, Paraná, Brazil
[c]Department of Crop Science, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil
[d]Plant Gene Expression Center, ARS/USDA, Albany, California, USA and Department of Plant and Microbial Biology, University of California-Berkeley, Berkeley, California, USA

e-mail adresses
Juliana Marcolino-Gomes: marcolino.ju@gmail.com
Fabiana Aparecida Rodrigues: deganuti@gmail.com
Renata Fuganti-Pagliarini : renatafuganti@gmail.com
Rafaela Ribeiro Reis: rafaelarireis@gmail.com
Thiago Jonas Nakayama: thiago84naka@gmail.com
Frank G. Harmon: fharmon@berkeley.edu
Alexandre Nepomuceno (Corresponding Author):
alexandre.nepomuceno@embrapa.br
Phone: +55 43 3371 621
Fax: +55 43 3371 6100
Postal Addres: Embrapa Soybean- Rod. Carlos Joao Strass, Warta, Londrina, PR, BRAZIL
Zip code: 86001-970

ABSTRACT

Background

Drought is the abiotic stress with the greatest impact on soybean crop production worldwide. Recently, water deficit responses have been associated with circadian clock oscillations at the transcription level, revealing the existence of hitherto unknown processes and increasing the demand for studies on plant responses to drought stress and its oscillation during the day. To accurately quantify target gene expression via RT-qPCR, raw expression data must be normalized against stable reference genes. Here, we analyzed the expression stability of soybean candidate reference genes in response to both drought *vs.* normal conditions and diurnal oscillations.

Results

Data mining from a transcriptome-wide background using microarrays and RNA-seq databases allowed us to select candidate reference genes from soybean, which were specifically chosen for the normalization of gene expression in studies on drought responses and diurnal oscillations. Experimental validation and stability analysis showed that five of these newer reference genes indeed exhibited greater expression stability than the conventionally used housekeeping genes. We also demonstrated the effect of using reference candidate genes with different stability values to normalize expression data from a drought-inducible soybean gene.

Conclusions

We selected and experimentally validated an unpublished set of soybean genes suitable for the normalization of gene expression data in studies on drought responses at different times of day.

Keywords: water deficit, normalization, circadian clock, gene expression, gene prospecting, *Glycine max*

BACKGROUND

As sessile organisms, plants must endure environmental changes during the day and across seasons. These environmental oscillations strongly affect light, temperature, nutrient and water availability, acting as a powerful selective pressure that have shaped adaptive mechanisms in plants during their evolutionary history. As a result, these organisms have developed a complex molecular network that confers adaptive advantages by coordinating their metabolism with predictable daily and seasonal changes, known as the circadian clock [1].

The circadian clock is composed of a core of interconnected transcriptional–translational feedback loops, which are entrained by signals such as light and temperature to adjust metabolism to the environment. In plants, the clock controls a number of physiological and developmental processes. For example, the expression of chlorophyll biosynthesis genes is regulated by the circadian clock to peak at the end of the night, which is an important mechanism to ensure photosynthesis in subsequent light periods of the day, whereas the products of photosynthesis modulate the rhythm [2]. The circadian clock also allows plants to coordinate flowering with favorable seasons to increase their fitness [1], as well as it controls the rate of starch degradation [3] and nitrogen assimilation and utilization pathways [4].

In addition to normal day/night variations, plants are subject to other environmental variations via biotic and abiotic stresses. Among the abiotic stresses, drought stands out as the factor with the greatest impact on yield of important crops worldwide, including soybean. Different mechanisms are employed by plants to protect themselves against water deficits, including changes in stomatal conductance [5], osmotic adjustment [6], the accumulation of osmoprotectant molecules [7], and the activity of antioxidant proteins [8]. Because the circadian clock is known to improve organism fitness according to environmental conditions, a significant number of studies addressing the relationships between water deficit stress and the circadian clock have been conducted, providing consistent evidence of this interaction [9–13].

The metabolic and physiological adjustments performed in response to drought stresses usually involve the reconfiguration of the transcriptome [12, 13], and therefore the analysis of gene expression in response to water deficits during the day is an interesting strategy [10]. One of the most sensitive methods for the quantification of gene expression is the fluorescence-based quantitative real-time

PCR (RT-qPCR), which is increasingly being used. The advantages of this technique include its practical simplicity combined with the possibility of measuring small amounts of RNA in a wide range of samples, rapidly and with high specificity.

However, because most of the quantitative RNA data obtained are not absolute, but relative, accurate quantification of gene expression relies on the use of appropriate reference genes. These genes should be stably expressed, showing a transcript abundance that is strongly correlated with the total mRNA present in the samples to allow the normalization of gene expression data [14]. Normalization is a key step in RT-qPCR analysis, as it reduces/eliminates variations due to variations in RNA extraction, reverse transcription yields or amplification efficiency, allowing comparisons of mRNA concentrations across different samples [14]. Although several genes have been indicated as good references, it is known that even housekeeping genes may exhibit altered expression in response to experimental treatments, sampling times and the life cycle [15–18]. In this context, a reference gene must be experimentally validated for specific tissues, genotypes and experimental designs.

The soybean genes *TUA* (Glyma08g12140), *TUB* (Glyma03g27970), *ELF1-β* (Glyma13g04050), *β-actin* (Glyma15g05570) and *GAPDH* (Glyma06g01850) have been widely used as references in gene expression studies on drought responses [15, 19]. On the other hand, isopentenyl diphosphate (*IPP2*), actin and ubiquitin are the most commonly used reference genes in studies investigating circadian/diurnal oscillations [20–25]. Thus, no study conducted to date has evaluated the expression stability of reference genes for the study of both water deficit stress and circadian oscillations in soybean. Hence, in this study, after evaluating gene expression in response to drought during the day, we present a novel set of reference genes suitable for the normalization of relative expression data from combined studies on water deficit and diurnal oscillations.

RESULTS AND DISCUSSION

Screening of candidate reference genes

RT-qPCR is an important tool that allows the relative quantification of transcript abundance and can therefore be used to evaluate gene expression responses to environmental changes, such as diurnal oscillations and abiotic stresses, including drought. However, there are several pre-requisites of this technique, such as a high amplification efficiency, target amplification specificity and the use of reference genes [14]. Reference genes are employed for data normalization to enable comparisons of mRNA concentrations across different samples, playing a critical role in the accurate quantification of relative gene expression.

To date, most of the studies on reference genes have focused on validating a subset of commonly used reference genes for specific contexts [15, 16, 18]. Although these studies have their merits, they attempt to identify the best candidates from a small set of genes. A recent analysis demonstrated that reference genes are preferably selected by adopting a complete genome strategy, rather than from a handful of commonly used reference genes [26]. In this context, we searched reference genes showing high expression stability in 59 microarray libraries from soybean subjected to drought stress, heat and different photoperiods. Then, in a second strategy, we selected genes that exhibited minimal expression variance across 36 cDNA libraries synthesized from drought-stressed soybean plants sampled over a 24 h timecourse [10].

These data mining strategies allowed us to select a new set of candidate reference genes composed of seven soybean genes: Glyma13g24060, Glyma01g40510, Glyma13g17500, Glyma08g05790, Glyma11g38000, Glyma08g41240 and Glyma10g44020 (Table 1). The majority of the selected candidate genes are related to the plant's primary metabolism. For example, Glyma01g40510 encodes a cysteine desulfurase (CYST) similar to nitrogen fixation S (NIFS)-like 1 from *Arabidopsis*; Glyma08g05790 encodes a protein that participates in Golgi vesicles transport (Golgin-84) [27, 28]; Glyma11g38000 produces an RNA (cytosine-5)-methyltransferase (NCL1) involved in epigenetic modifications of tRNA [29, 30]; and Glyma13g17500 produces an FYVE domain protein, present in kinases and lipases in *Arabidopsis*, that recognizes phosphoinositide signals [31].

Table1. Information on reference and target genes

Gene name	Gene ID	Prospected from	Sense primer	Primer sequence	Tm ºC	Amplicon size (bp)	Efficiency (%)	Primer Concentration (nM)
NUDIX	Glyma13g24060	GeneVestigator (microarray)	F	TGAGTGTTAGAAGGGCTACTGG	58.5	108	88.02	200
			R	AACTTTGCCAACGGCATC	59.7			
CYST	Glyma01g40510	GeneVestigator (microarray)	F	TTCTTGGATCGGGGAGAG	58.7	126	95.49	100
			R	GCTAGAAATGGCGAAAGAGG	59.1			
FYVE	Glyma13g17500	GeneVestigator (microarray)	F	TTCTGTCTTCTGCAAGTGGTG	59.1	92	98.37	100
			R	GATCCCTCATCCATACATTTCAG	59.7			
Golgin-84	Glyma08g05790	GeneVestigator (microarray)	F	TTGGACAAGGAGAGACTCCAC	59.3	121	98.57	100
			R	TGCGAGGCTACGAAAACTTC	60.5			
NCL1	Glyma11g38000	GeneVestigator (microarray)	F	TCTTATCGGCATGGTTACGC	61.0	72	96.38	100
			R	ACATAACCAAACGCCAAAGC	60.0			
RNA-poly Mitovirus	Glyma08g41240	RNAseq libraries	F	TCTCATCACCGATCCACTTG	59.6	114	98.72	100
			R	GCAAACTCTACAGCACCAGTTG	60.0			
DNAJ	Glyma10g44020	RNAseq libraries	F	GATTGGGATGTTCTTCACCAG	59.4	136	90.50	200
			R	ATGACCAAGCCGATGGTTAG	60.0			
ELF1-β	Glyma13g04050	Literature [32]	F	GTTGAAAAGCCAGGGGACA	58.0	118	99.16	60
			R	TCTTACCCCTTGAGCGTGG	58.0			
β-actin	Glyma15g05570	Literature [33]	F	GAGCTATGAATTGCCTGATGG	58.0	118	97.92	60
			R	CGTITCATGAATTCCAGTAGC	55.0			
DREB5	Glyma12g33020	Literature [34]	F	TTGCCTACTACTACTCCTATATTCATTTCC	58.0	86	97.80	100
			R	CCTTGAAATACACGGAGCCTTAG	58.0			

The **Gene name** is proposed based on gene annotation in the Phytozome database; **Prospected from** refers to the source/database used to prospect the genes; the **Gene ID** corresponds to the gene's identification in the Phytozome database; **Sense primer, Primer sequence, Temperature of melting (Tm)** and **Amplicon size** refer to the features of the oligonucleotides.

The expression of these candidate reference genes in response to drought and to diurnal oscillations was compared with the expression of the commonly used soybean reference genes Glyma20g26690, Glyma12g02790, Glyma12g05510, Glyma08g12140, Glyma10g38460, Glyma15g05570, Glyma13g04050, Glyma13g04050, Glyma03g27970 and Glyma06g01850 [15, 19], as shown in analyses performed in microarray (Figure 1) and RNA-Seq (Figure 2) databases.

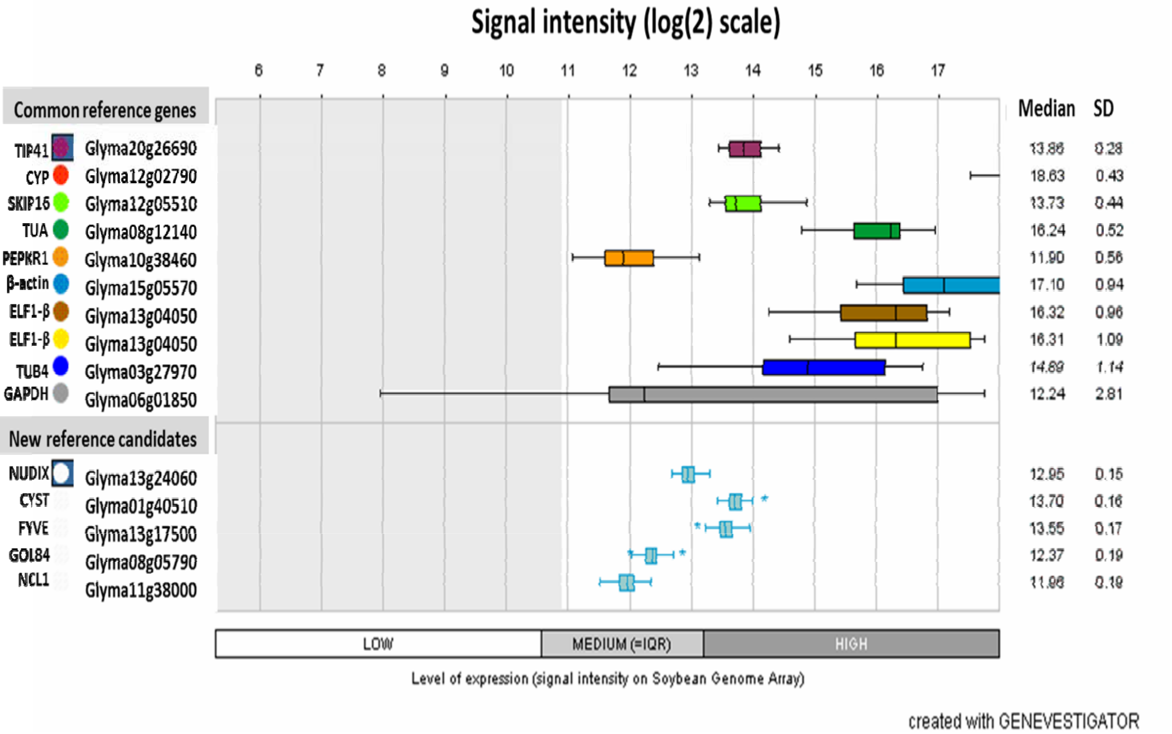
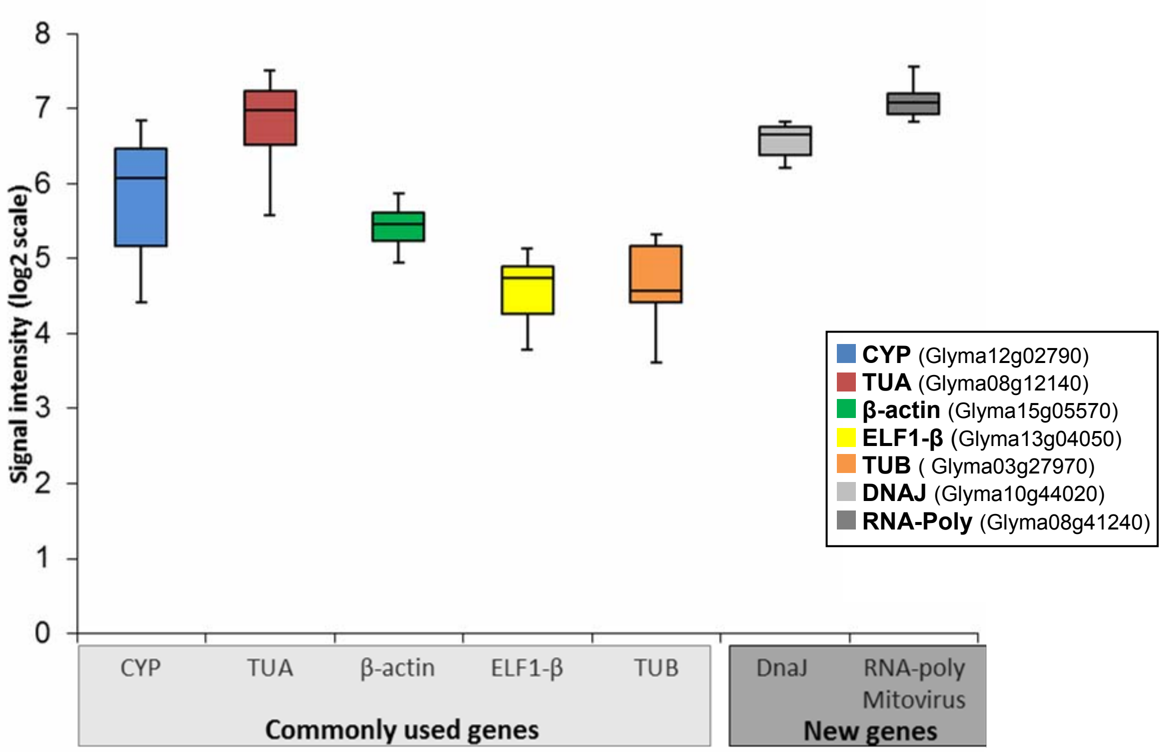


Figure 1. Expression of the commonly used reference genes and new candidates from microarray databases. The analyses were performed at the Genevestigator software, using data from soybean under drought stress, heat and different photoperiods [27]. The box plots represent the signal intensity (log$_2$ scale) variation. The interquartile range (IQR) values are shown for each gene across the dataset, represented by the middle boxes, which comprehends the middle 50% of scores for the group. The upper and lower whiskers represent scores outside the middle 50%. Outliers are plotted separately as asterisks on the



ch

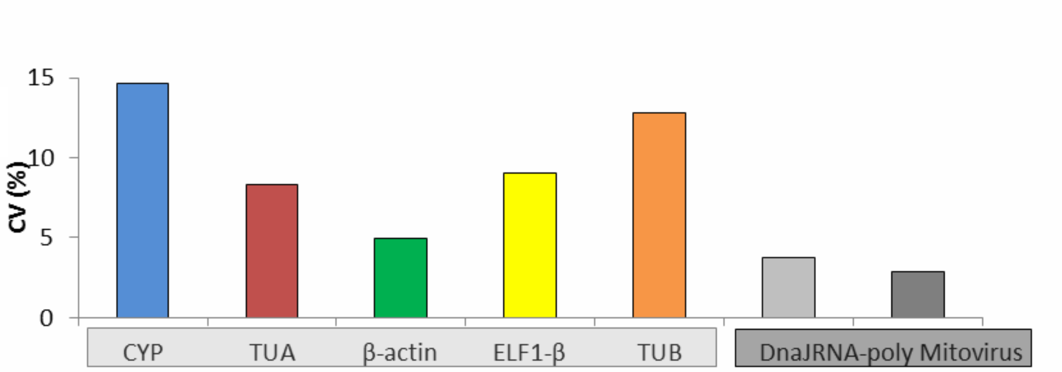


Figure 2. Expression variation of the commonly used reference genes and new candidates from RNA-Seq database. The gene expression analyses were performed using data from soybean under

The differential expression of some of the selected candidate genes has been reported during biotic and abiotic stress responses. The gene Glyma13g24060, for example, encodes a protein similar to a NUDIX hydrolase protein from *Arabidopsis*. The NUDIX hydrolase family is widespread, from eukaryotes to *Archaea*, and consists of pyrophosphohydrolases that act upon substrates with a general nucleoside diphosphate structure, including (deoxy)ribonucleoside diphosphates and triphosphates, nucleotide sugars, coenzymes and RNA caps [35, 36]. Members of the NUDIX family have been reported to be induced by salt, drought, heat, and cold in *Chrysanthemum lavandulifolium* [37]. Similarly, Glyma10g44020 encodes a protein from the DnaJ/Hsp40 cysteine-rich domain superfamily, which is described as being involved in diverse cellular processes (protein folding, translocation, and degradation) [38], including biotic and abiotic stress responses [39–42]. However, analysis of the soybean *NUDIX* (Figure 1) and *DNAJ* (Figure 2) genes in microarray and RNA-Seq databases showed high expression stability in response to drought and diurnal oscillations, suggesting that these soybean genes could be reliable candidate reference genes for drought studies.

Furthermore, we identified a gene (Glyma08g41240) that encodes an RNA-dependent RNA polymerase from a mitovirus (Figure 2, Table 1). A recent study on the soybean mitochondrial genome revealed the presence of a 0.5 kb insertion (at *rps10* intron) that is 57.4% identical to a mitovirus RNA polymerase gene, which might have been horizontally transferred during recent evolution. Although the effect of this insertion remains unknown, analysis of the insert's position suggests that it might affect the function of the mitochondrial *rps10* gene, which encodes the ribosomal protein S10 [43].

The preliminary analysis of this set of genes (Figures 1 and 2) revealed that in the majority of cases, the candidate reference genes presented less variation than the commonly used reference genes selected from the literature [15, 19]. Additionally, *in silico* pre-validation of the *NUDIX*, *CYST*, *FYVE*, *Golgin-84*, and *NCL1* genes across 3,458 microarrays using GeneVestigator platform revealed that these genes are unresponsive to a wide variety of conditions, including abiotic stresses, such as heat, salinity, and cold, and show little variation between developmental stage and genotypes, being responsive only to infections by *Phytophthora sojae*, *Phakopsora pachyrhizi* and *Bradyrhizobium japonicum* (data not shown).

Stability analysis

To experimentally validate this new set of candidate reference genes, we examined their individual properties and compared the stability of their expression with the commonly used reference genes *ELF1-β* and *β-actin* (Table 1) using RT-qPCR assays. In previous studies, *ELF1-β* and *β-actin* showed high expression stability across different levels of drought [15, 32, 34]. However, no study conducted to date has investigated the expression stability of these genes during the diurnal cycle. Validation was carried out on soybean leaves from plants subjected to a moderate water deficit (30% of gravimetric humidity (GH)), sampled across a 24 h timecourse, with 4 h intervals, from 8:00 a.m. to 4:00 a.m., corresponding to Zeitgeber Time (ZT) 0 to ZT20. The expression level of each reference gene was evaluated separately in drought-treated and control plants at all sampling times (ZT0, ZT4, ZT8, ZT12, ZT16 and ZT20).

To assess the stability values for the reference genes, we performed analysis using the NormFinder [44] and Genorm softwares [45]. The NormFinder software employs a variance estimation approach to rank genes according to combined inter- and intragroup expression variation across a given set of experimental conditions, which were drought stress and diurnal oscillations in this case. Although NormFinder [44] is known to perform in a more robust manner and with less sensitivity toward the co-regulation of candidate genes compared with other software, we performed an additional stability analysis using Genorm software [45] to compare the performance of the candidate genes using different stability analysis tools. In general, the results of Genorm analysis were similar to those from NormFinder, with slight differences regarding ranking positions (Figure 3A and B). According these stability analysis, many of the newer reference genes indeed exhibited greater expression stability than the conventionally used reference genes (*ELF1-β* and *β -actin*) (Figure 3A and B). The *FYVE*, *NUDIX* and *Golgin-84* genes were the most stable, suggesting that they are the most suitable for normalizing expression data from combined studies addressing drought treatment and diurnal oscillation (Figure 3A and B).

Although most of the candidate reference genes performed well in response to the applied experimental conditions, the *NCL1* and *DNAJ* genes showed lower stability (Figure 3A and B).

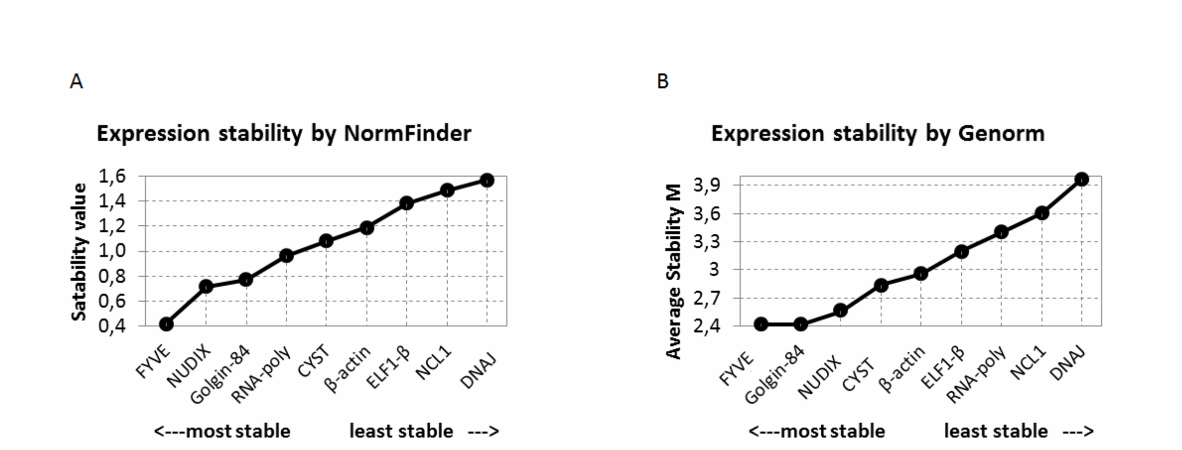


Figure 3. Stability analysis of candidate reference genes. Gene expression stability was measured in leaf tissues of soybean subjected to drought stress and control conditions at different times of day (ZT0, ZT4, ZT8, ZT12, ZT16 and ZT20). The analysis was performed with NormFinder and Genorm softwares. Genes were ranked according to their stability values and *M values* from NormFinder and Genorm, respectively. The genes were plotted on the x-axis from the most to the least stable.

The analysis of intragroup variation performed on NormFinder for *NCL1* revealed that under control conditions this was one of the least stable genes (Figure 4A), whereas in stressed plants it was among the most stable (Figure 4B), indicating that *NCL1* is not a suitable reference gene for studies on gene expression oscillation during the day in plants under control conditions. Furthermore, this result illustrates that gene expression stability during the day may vary according to the plant's water status (under normal hydration versus water deficit conditions).

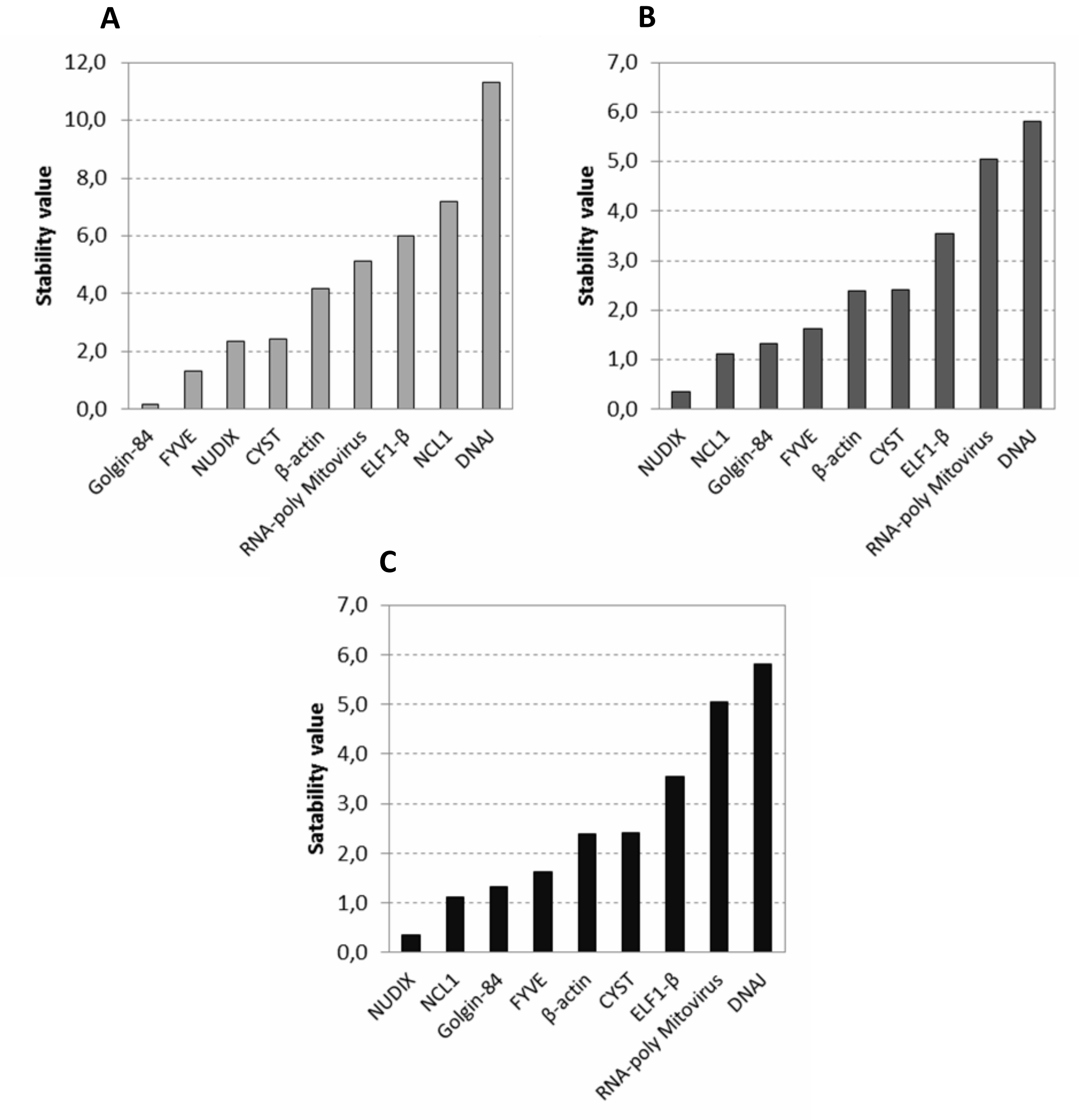


Figure 4. Intra- and intergroup variation of the gene expression. The intragroup variation within **(A)** control (non-stressed plants) and **(B)** stress (plants under drought stress), and **(C)** the intergroup variation are presented.

In contrast, the expression of the *DNAJ* gene was unstable in both control and stressed conditions (Figure 4ABC). Genes from the DNAJ family have been reported to be drought responsive in many species, and its lack of stability may therefore be explained in part by its possible involvement in drought responses in soybean, as studies have reported that drought-responsive genes oscillate in response to diurnal oscillations [10, 12, 13, 34, 46]. Additionally, the intragroup analysis shows that *Golgin-84* and *FYVE* (Figure 4A) are the most reliable reference genes for gene expression normalization when studying diurnal oscillations in plants under normal water conditions, whereas *NUDIX* and *NCL1* (Figure 4B) are the most strongly indicated for use in studies under drought conditions. Furthermore, the intergroup variation analysis allowed us to identify the genes *NUDIX* and *DNAJ* as the most and least stable genes, respectively, for data normalization in studies comparing gene expression under control and drought stress, without considering the time of day (Figure 4C).

Normalization of a target gene

A previous study showed that the normalization of soybean genes under circadian regulation using unstable reference genes may lead to erroneous data interpretation [19]. To demonstrate the effect of data normalization using reference genes with different stability values in response to drought stress during the day, we evaluate the relative expression of the drought-responsive gene *GmDREB5*-like [34] using a combination of 2 reference genes with high (*FYVE and GOL84*), intermediate (*ELF1-β* and *β -actin*) and low (*DNAJ and NCL1*) expression stabilities (Figure 5). As expected, this analyses demonstrated that the expression of *GmDREB5*-like in response to drought oscillates during the day, as described for other genes of the DREB subfamily [10, 47–52] (Figure 5).

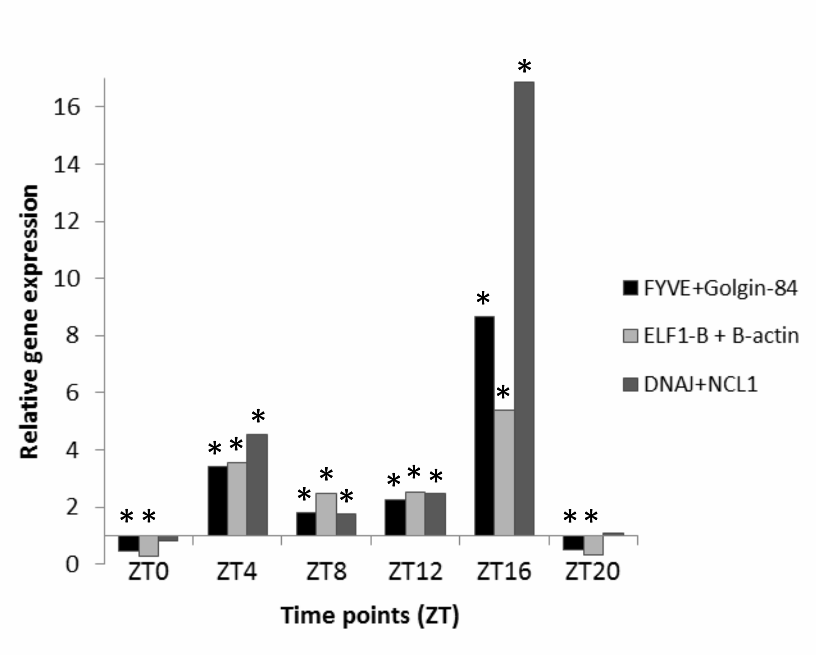


Figure 5. Normalization of the target gene *GmDREB5*-like. Gene expression was measured in leaf tissues of soybean subjected to drought stress at different times of day (ZT0, ZT4, ZT8, ZT12, ZT16 and ZT20). The raw data were normalized using a combination of 2 candidate reference genes with high (FYVE and GOL84), intermediate (ELF1-β and β -actin) and low (DNAJ and NCL1) expression stabilities. The relative expression of *GmDREB5*-like under drought stress was determined after comparison with the control sample (non-stressed plants). Asterisks indicate statistically significant changes in gene expression between drought stressed and non-stressed plants. The analyses were performed using REST 2009 software.

Our results demonstrated that choosing reference genes with diverse stability of expression can lead to differences on gene expression data interpretation when evaluating combined studies on water deficit and diurnal oscillations. As shown in Figure 5, at ZT16 the normalization of *GmDREB5*-like gene using the least stable genes (*DNAJ and NCL1*) resulted in higher expression levels than observed for normalization using genes with high (FYVE and GOL84) and intermediate expression stability (ELF1-β and β –actin). Additionally, slight changes in gene expression, such as the down-regulation of the target gene at ZT0 and ZT20, were detected only by normalization using genes with high (*FYVE* and *GOL84*) and intermediate expression stability (*ELF1-β* and *β –actin*) (Figure 5A). These results emphasize the importance of selecting reference genes with stable expression for accurate gene expression analysis on drought responses and diurnal oscillations.

CONCLUSION

Here, by analyzing experiments involving both drought and diurnal oscillations, we demonstrated the importance of selecting reference genes under the specific studied conditions. From a transcriptome-wide background, we selected a new set of candidate reference genes for the normalization of data obtained in studies on drought and diurnal oscillations.

The experimental validation of this new set of candidate reference genes revealed that *FYVE*, *NUDIX* and *Golgin-84* were the most stably expressed genes in soybean plants under control and drought conditions along the day, and are therefore considered the best reference genes for the studied conditions. Our results highlight that the selection of reference genes is crucial for the proper quantification of relative expression data obtained under specific experimental conditions.

MATERIAL AND METHODS

Selection of reference genes

To evaluate the stability of genes expressed in response to drought during the day, we used the RefGenes tool from the Genevestigator platform [26], available at [https://www.genevestigator.com/gv/plant.jsp]. This tool allowed us to perform *in silico* identification of genes showing high expression stability in 59 microarray libraries from soybean subjected to drought, heat and distinct light periods. To select the candidate reference genes presenting range of expression levels detected by RT-qPCR, we uploaded commonly used RT-qPCR reference genes from soybean (Glyma20g26690, Glyma12g02790, Glyma12g05510, Glyma08g12140, Glyma10g38460, Glyma15g05570, Glyma13g04050, Glyma03g27970 and Glyma06g01850 [19]. The candidate reference genes obtained in this analysis were pre-validated by checking their expression across all microarrays available on the Genevestigator platform (3458 arrays) (data not shown). The expression profiles of the six most stable new reference genes and the commonly used references are shown in Figure 1.

Under a second strategy, we evaluated the expression stability of genes from 36 cDNA libraries synthesized from drought-stressed soybean plants sampled over a 24 h timecourse [10]. In this analysis, we selected genes that exhibited minimal expression variance across the libraries (presenting the lowest standard deviation), with a range of expression similar to that of commonly used soybean reference genes [19]. The expression profiles of the two most stable new reference genes and the commonly used references are shown in Figure 2. In summary, a total of seven new candidate reference genes were selected for validation via RT-qPCR (Table 1).

Primer design

Primers for the new candidate reference genes were designed based on soybean GeneModel sequences [http://www.phytozome.net/search.php?method=Org_Gmax] using the program Primer3 Plus [53], available at [http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi]. The primer sequences were determined for the 3' end of each gene whenever possible, and the amplicons spanned up to 150 base pairs (bp). The primer sequences were subjected to BLAST searches against the soybean genome [http://www.phytozome.net/search.php?method=Org_Gmax] to verify the specificity of each primer, as recommended by the Minimum Information for Publication of Quantitative Real-Time PCR Experiments guideline (MIQE) [14]. The primers for the commonly used reference genes *ELF1-β* and *β-actin* were selected from [32] and [33], respectively. Information on the primers may be visualized in Table 1.

Plant material and treatment application

Plant material was obtained from experiments performed as described by Marcolino-Gomes and colleagues [10]. Briefly, seeds from the soybean BR16 genotype, which exhibits drought-sensitive characteristics [54], were cultivated in peat pots (Jiffy) with Supersoil® (Scotts Miracle-Gro Company, Marysville, Ohio, USA) under optimal growth conditions in controlled growth chambers until reaching the V_2 developmental stage [55], when water was withheld to induce a moderate water deficit. Control plants were maintained near field capacity for the unstressed treatment. The soil moisture was calculated by the gravimetric humidity (GH), which corresponds to the percentage of water in the soil in relation to the dry weight of the soil. The volume of irrigation was adjusted to 70% (GH) (near field capacity) for the unstressed treatment, 30% GH for the moderate stress treatment. The pots were weighed twice a day, and water was added to maintain the treatments at the desired GH values. The middle leaflet of the first trifoliate leaf was collected from plants in each treatment group at 4 h intervals from the time the lights came on (8:00 a.m. = Zeitgeiber Time (ZT) 0), during a 24 h timecourse (ZT0, ZT4, ZT8, ZT12, ZT16 and ZT20), and were immediately frozen in liquid nitrogen and stored at −80°C until further use. All of the experiments were conducted with three biological replicates, with each replicate consisting of two plants, whose tissues were collected together and pooled.

From RNA extraction to cDNA synthesis

Each replicate tissue set was ground to a fine powder in liquid nitrogen, and total RNA was isolated using the TRIzol reagent (Invitrogen) according to the manufacturer’s instructions. The obtained RNA concentration and purity were measured using a spectrophotometer (NanoDrop, ND-1000), and contaminating DNA in the total RNA was removed using the Turbo DNA-free kit, according to the manufacturer’s instructions (Life Technologies, Grand Island, NY, USA). After DNAse treatment, the integrity of the molecules was analyzed on 1% agarose gels stained with ethidium bromide, and high-quality total RNA was used to synthesize cDNA strands (Superscript III First Strand Synthesis, Invitrogen/Life Technologies, Grand Island, NY, USA). The quality of the cDNA and contamination with genomic DNA were examined using a standard PCR assay with primers that spanned an intronic region of the *β-actin* soybean gene. High-quality cDNA was used to analyze the transcripts in each treatment.

RT-qPCR analyses

Standard curves were produced from serial dilutions of a cDNA pool to estimate the efficiency of the PCR amplification with each pair of primers. RT-qPCR amplifications were performed in a 7300 RT-qPCR Thermocycler (Applied

Biosystems/Life Technologies, Grand Island, NY, USA) with the following cycling parameters: 50°C for 2 min, 95°C for 10 min and 45 cycles at 95°C for 2 min, 60°C for 30 seconds and 72°C for 30 seconds. Each amplification reaction contained 2 μL of cDNA from serial dilutions, 60-200 nM each forward and reverse primer (Table 1), 500 nM ROX (passive reference), 6.5 μL of Platinum® SYBR® Green qPCR SuperMix (Invitrogen/ Grand Island, NY, USA), and ultrapure water to a final volume of 12.5 μL. Data were collected during the extension phase, and dissociation curves were generated by heating each reaction from 60 to 95°C and taking readings at one-degree intervals to verify the specificity of the primers. A control sample, obtained via performing RT-qPCR with no template, was also assayed to confirm that the samples were not contaminated. The primer concentrations were adjusted to achieve efficiency rates higher than 85%, as detailed in Table 1.

After carrying out the efficiency analysis, the expression levels of the candidate reference genes were analyzed separately at 6 time points (ZT0, ZT4, ZT8, ZT12, ZT16 and ZT20) in plants under control *vs.* drought stress condition in order to assess their expression stability along the day. The expression of the target gene (*GmDREB5*-like; Glyma12g33020) [34] was also measured under the experimental conditions described above. The reactions were performed in triplicate with cycling parameters similar to those described above for the amplification efficiency analysis.

Stability analysis

To validate and compare the suitability of the candidate reference genes for use in normalization, we evaluated their expression stability in response to drought along the day under the experimental conditions described above. For this purpose, Cycle threshold values (Ct) were transformed into non-normalized relative quantities (Q; linear scale). Here, Q = EΔCt, where E is the amplification efficiency, and ΔCt is the lowest Ct from the data set minus the sample Ct. The non-normalized relative quantities were analyzed using NormFinder [44] and GeNorm [45] software to assess the expression stability of the reference genes.

Normalization of target gene expression

The relative expression level of the drought-responsive gene *GmDREB5*-like [34] was measured in leaf samples from BR16 plants subjected to moderate drought, sampled over a 24 h timecourse, under the experimental conditions described above. For each time point (ZT0, ZT4, ZT8, ZT12, ZT16 and ZT20), three biological replicates, with three technical replicates each, were analyzed. Target expression was normalized using the endogenous genes *FYVE* (Glyma13g17500) and *DNAJ* (Glyma10g44020) separately (Table1). Plants grown under normal water conditions (control plants) were used to calibrate relative expression. The gene expression analysis was performed using the Rest2009 software package [56], which allows the input of different amplification efficiencies for the reference and target genes and provides the statistical significance of expression levels through randomization (10,000 interactions) and bootstrapping of the data. Hypothesis testing was conducted to determine whether the differences between the control and treatment conditions were significant [56].

COMPETING INTERESTS

The author(s) declare that they have no competing interests. Neither the interpretation of data nor presentation of information was influenced by personal or financial relationships with other people or organizations.

AUTHORS' CONTRIBUTIONS

Conceived and designed the experiments: JMG, FGH and ALN.

Performed the experiments: JMG, TJN, RRR.

Analyzed the data: JMG, TJN, FGH, ALN.

Wrote the paper: JMG, FGH, ALN.

Helped to draft/correct the manuscript: JMG, FAR, RFP, FGH and ALN.

All authors read and approved the final manuscript

ACKNOWLEDGEMENTS

This study was supported by the National Council for Scientific and Technological Development - CNPq - and the Brazilian Federal Agency for the Support and Evaluation of Graduate Education -CAPES. We also thank the National Center for Soybean Research - Embrapa Soybean - for the use of the greenhouse and laboratory facilities.

REFERENCES

1. Yerushalmi S, Green RM: **Evidence for the adaptive significance of circadian rhythms.** *Ecol Lett* 2009, **12**:970–81.

2. Haydon MJ, Mielczarek O, Robertson FC, Hubbard KE, Webb A a R: **Photosynthetic entrainment of the Arabidopsis thaliana circadian clock.** *Nature* 2013, **502**:689–92.

3. Graf A, Smith AM: **Starch and the clock: the dark side of plant productivity.** *Trends Plant Sci* 2011, **16**:169–75.

4. Gutiérrez R, Stokes T: **Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene CCA1**. *Proc Natl Acad Sci* 2008, **105**:4939–4944.

5. Mahdieh M, Mostajeran A, Horie T, Katsuhara M: **Drought stress alters water relations and expression of PIP-type aquaporin genes in Nicotiana tabacum plants.** *Plant Cell Physiol* 2008, **49**:801–13.

6. Babita M, Maheswari M, Rao LM, Shanker AK, Rao DG: **Osmotic adjustment, drought tolerance and yield in castor (Ricinus communis L.) hybrids**. *Environ Exp Bot* 2010, **69**:243–249.

7. Molinari HBC, Marur CJ, Daros E, de Campos MKF, de Carvalho JFRP, Filho JCB, Pereira LFP, Vieira LGE: **Evaluation of the stress-inducible production of proline in transgenic sugarcane (Saccharum spp.): osmotic adjustment, chlorophyll fluorescence and oxidative stress**. *Physiol Plant* 2007, **130**:218–229.

8. Carvalho MHC: **Drought stress and reactive oxygen species**. *Plant Signal Behav* 2008, **3**:156–165.

9. Legnaioli T, Cuevas J, Mas P: **TOC1 functions as a molecular switch connecting the circadian clock with plant responses to drought.** *EMBO J* 2009, **28**:3745–57.

10. Marcolino-Gomes J, Rodrigues FA, Fuganti-Pagliarini R, Bendix C, Nakayama TJ, Celaya B, Molinari HBC, de Oliveira MCN, Harmon FG, Nepomuceno A: **Diurnal**

oscillations of soybean circadian clock and drought responsive genes. *PLoS One* 2014, **9**:e86402.

11. Sanchez A, Shin J, Davis SJ: **Abiotic stress and the plant circadian clock**. *Plant Signal Behav* 2011, **6**:223–231.

12. Wilkins O, Bräutigam K, Campbell MM: **Time of day shapes Arabidopsis drought transcriptomes**. *Plant J* 2010, **63**:715–727.

13. Wilkins O, Waldron L, Nahal H, Provart NJ, Campbell MM: **Genotype and time of day shape the Populus drought response.** *Plant J* 2009, **60**:703–15.

14. Bustin S a, Benes V, Garson J a, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT: **The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments.** *Clin Chem* 2009, **55**:611–22.

15. Ma S, Niu H, Liu C, Zhang J, Hou C, Wang D: **Expression stabilities of candidate reference genes for RT-qPCR under different stress conditions in soybean.** *PLoS One* 2013, **8**:e75271.

16. Miranda VDJ, Coelho RR, Viana AAB, de Oliveira Neto OB, Carneiro RMDG, Rocha TL, de Sa MFG, Fragoso RR: **Validation of reference genes aiming accurate normalization of qPCR data in soybean upon nematode parasitism and insect attack.** *BMC Res Notes* 2013, **6**:196.

17. Nakayama TJ, Rodrigues FA, Neumaier N, Marcelino-Guimarães FC, Farias JRB, de Oliveira MCN, Borém A, de Oliveira ACB, Emygdio BM, Nepomuceno AL: **Reference genes for quantitative real-time polymerase chain reaction studies in soybean plants under hypoxic conditions.** *Genet Mol Res* 2014, **13**:860–71.

18. Le DT, Aldrich DL, Valliyodan B, Watanabe Y, Ha C Van, Nishiyama R, Guttikonda SK, Quach TN, Gutierrez-Gonzalez JJ, Tran L-SP, Nguyen HT: **Evaluation of candidate reference genes for normalization of quantitative RT-PCR in soybean tissues under various abiotic stress conditions.** *PLoS One* 2012, **7**:e46487.

19. Hu R, Fan C, Li H, Zhang Q, Fu Y-F: **Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR.** *BMC Mol Biol* 2009, **10**:93.

20. Song YH, Lee I, Lee SY, Imaizumi T, Hong JC: **CONSTANS and ASYMMETRIC LEAVES 1 complex is involved in the induction of FLOWERING LOCUS T in photoperiodic flowering in Arabidopsis.** *Plant J* 2012, **69**:332–42.

21. Huang W, Pérez-García P, Pokhilko A, Millar AJ, Antoshechkin I, Riechmann JL, Mas P: **Mapping the core of the Arabidopsis circadian clock defines the network structure of the oscillator.** *Science (80-)* 2012, **336**:75–9.

22. Jiang Y, Han YZ, Zhang XM: **Expression profiles of a CONSTANS homologue GmCOL11 in Glycine max**. *Russ J Plant Physiol* 2011, **58**:928–935.

23. Galeou A, Prombona A: **Light at night resynchronizes the evening-phased rhythms of TOC1 and ELF4 in Phaseolus vulgaris**. *Plant Sci* 2012, **184**:141–147.

24. Hudson KA: **The Circadian Clock-controlled Transcriptome of Developing Soybean Seeds**. *Plant Genome J* 2010, **3**:3.

25. Li G, Siddiqui H, Teng Y, Lin R, Wan X, Li J, Lau O-S, Ouyang X, Dai M, Wan J, Devlin PF, Deng XW, Wang H: **Coordinated transcriptional regulation underlying the circadian clock in Arabidopsis.** *Nat Cell Biol* 2011, **13**:616–22.

26. Hruz T, Wyss M, Docquier M, Pfaffl MW, Masanetz S, Borghi L, Verbrugghe P, Kalaydjieva L, Bleuler S, Laule O, Descombes P, Gruissem W, Zimmermann P: **RefGenes: identification of reliable and condition specific reference genes for RT-qPCR data normalization.** *BMC Genomics* 2011, **12**:156.

27. Munro S: **The golgin coiled-coil proteins of the Golgi apparatus.** *Cold Spring Harb Perspect Biol* 2011, **3**.

28. Osterrieder a: **Tales of tethers and tentacles: golgins in plants.** *J Microsc* 2012, **247**:68–77.

29. Pavlopoulou A, Kossida S: **Phylogenetic analysis of the eukaryotic RNA (cytosine-5)-methyltransferases.** *Genomics* 2009, **93**:350–7.

30. Squires JE, Preiss T: **5-Methylcytosine as a modification in RNA**. In *Epigenetics: A Reference Manual*. 1st edition. Edited by Craig J, Wong NC. Norfolk: Horizon Scientific Press; 2011:449.

31. Jensen RB, Cour TLA, Albrethsen J, Nielsen M, Skriver K: **identification of PtdIns3 P -binding residues by comparison of classic and variant FYVE domains**. 2001, **173**:165–173.

32. Jian B, Liu B, Bi Y, Hou W, Wu C, Han T: **Validation of internal control for gene expression study in soybean by quantitative real-time PCR.** *BMC Mol Biol* 2008, **9**:59.

33. Stolf-Moreira R, Lemos EG de M, Abdelnoor RV, Beneventi MA, Rolla AAP, Pereira S dos S, Oliveira MCN, Nepomuceno AL, Marcelino-Guimarães FC: **Identification of reference genes for expression analysis by real-time quantitative PCR in drought-stressed soybean**. *Pesqui Agropecu Bras* 2011, **46**:58–65.

34. Marcolino-Gomes J, Rodrigues FA, Oliveira MCN, Farias JRB, Neumaier N, Abdelnoor RV, Marcelino-Guimarães FC, Nepomuceno AL: **Expression Patterns of GmAP2/EREB-Like Transcription Factors Involved in Soybean Responses to Water Deficit.** *PLoS One* 2013, **8**:e62294.

35. Ogawa T, Yoshimura K, Miyake H, Ishikawa K, Ito D, Tanabe N, Shigeoka S: **Molecular characterization of organelle-type Nudix hydrolases in Arabidopsis.** *Plant Physiol* 2008, **148**:1412–24.

36. Goyer A, Hasnain G, Frelin O, Ralat M a, Gregory JF, Hanson AD: **A cross-kingdom Nudix enzyme that pre-empts damage in thiamin metabolism.** *Biochem J* 2013, **454**:533–42.

37. Huang H, Cao H, Niu Y, Dai S: **Expression Analysis of Nudix Hydrolase Genes in Chrysanthemum lavandulifolium**. *Plant Mol Biol Report* 2012, **30**:973–982.

38. Rajan VB V, D'Silva P: **Arabidopsis thaliana J-class heat shock proteins: cellular stress sensors.** *Funct Integr Genomics* 2009, **9**:433–46.

39. Wang G, Cai G, Kong F, Deng Y, Ma N, Meng Q: **Overexpression of tomato chloroplast-targeted DnaJ protein enhances tolerance to drought stress and resistance to Pseudomonas solanacearum in transgenic tobacco.** *Plant Physiol Biochem* 2014, **82**:95–104.

40. So H-A, Chung E, Lee J-H: **Molecular characterization of soybean GmDjp1 encoding a type III J-protein induced by abiotic stress**. *Genes Genomics* 2013, **35**:247–256.

41. Liu JZ, Whitham S a.: **Overexpression of a soybean nuclear localized type-III DnaJ domain-containing HSP40 reveals its roles in cell death and disease resistance**. *Plant J* 2013, **74**:110–121.

42. Bekh-Ochir D, Shimada S, Yamagami A, Kanda S, Ogawa K, Nakazawa M, Matsui M, Sakuta M, Osada H, Asami T, Nakano T: **A novel mitochondrial DnaJ/Hsp40 family protein BIL2 promotes plant growth and resistance against environmental stress in brassinosteroid signaling**. *Planta* 2013, **237**:1509–1525.

43. Chang S, Wang Y, Lu J, Gai J, Li J, Chu P, Guan R, Zhao T: **The mitochondrial genome of soybean reveals complex genome structures and gene evolution at intercellular and phylogenetic levels.** *PLoS One* 2013, **8**:e56502.

44. Andersen CL, Jensen JL, Ørntoft TF: **Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets.** *Cancer Res* 2004, **64**:5245–50.

45. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: **Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.** *Genome Biol* 2002, **3**:RESEARCH0034.

46. Mizuno T, Yamashino T: **Comparative transcriptome of diurnally oscillating genes and hormone-responsive genes in Arabidopsis thaliana: insight into**

circadian clock-controlled daily responses to common ambient stresses in plants. *Plant Cell Physiol* 2008, **49**:481–7.

47. Kidokoro S, Maruyama K, Nakashima K, Imura Y, Narusaka Y, Shinwari ZK, Osakabe Y, Fujita Y, Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K: **The phytochrome-interacting factor PIF7 negatively regulates DREB1 expression under circadian control in Arabidopsis.** *Plant Physiol* 2009, **151**:2046–57.

48. Nakamichi N, Kusano M, Fukushima A, Kita M, Ito S, Yamashino T, Saito K, Sakakibara H, Mizuno T: **Transcript profiling of an Arabidopsis PSEUDO RESPONSE REGULATOR arrhythmic triple mutant reveals a role for the circadian clock in cold stress response.** *Plant Cell Physiol* 2009, **50**:447–62.

49. Fowler SG, Cook D, Thomashow MF: **Low temperature induction of Arabidopsis CBF1, 2, and 3 is gated by the circadian clock.** *Plant Physiol* 2005, **137**:961–8.

50. Dong MA, Farré EM, Thomashow MF: **Circadian clock-associated 1 and late elongated hypocotyl regulate expression of the C-repeat binding factor (CBF) pathway in Arabidopsis.** *Proc Natl Acad Sci U S A* 2011, **108**:7241–6.

51. Mikkelsen MD, Thomashow MF: **A role for circadian evening elements in cold-regulated gene expression in Arabidopsis**. *Plant J* 2009, **60**:328–339.

52. Bieniawska Z, Espinoza C, Schlereth A, Sulpice R, Hincha DK, Hannah M a: **Disruption of the Arabidopsis circadian clock is responsible for extensive variation in the cold-responsive transcriptome.** *Plant Physiol* 2008, **147**:263–279.

53. Rozen S, Skaletsky H: **Primer3 on the WWW for General Users and for Biologist Programmers**. In *Bioinformatics Methods and Protocols Methods in Molecular Biology*. Edited by Misener S, Krawetz SA. Humana Press; 1999:365–386.

54. Oya T, Nepomuceno AL, Neumaier N, Renato J, Farias B, Tobita S, Ito O: **Drought Tolerance Characteristics of Brazilian Soybean Cultivars: Evaluation and characterization of drought tolerance of various Brazilian soybean cultivars in the field.** *Plant Prod Sci* 2004, **7**(October 2003):129–137.

55. Fehr WR, Caviness CE, Burmood DT, Pennington JS: **Stage of Development Descriptions for Soybeans, Glycine Max (L.) Merrill1**. *Crop Sci* 1971, **11**:929.

56. Pfaffl MW, Horgan GW, Dempfle L: **Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR.** *Nucleic Acids Res* 2002, **30**:e36.

6. CONSIDERAÇÕES FINAIS

Os genes do ciclo circadiano desempenham o controle rítmico de diversos processos biológicos, otimizando-os para cada período do dia e permitindo que as plantas antecipem e sincronizem sua fisiologia às mudanças ambientais. Este relógio interno pode ainda participar das respostas a estresses ambientais, como o déficit hídrico. Todavia, até o presente momento desconhecia-se a participação dos genes do ciclo circadiano na resposta ao déficit hídrico em soja. O presente estudo identificou 11 genes do ciclo circadiano em soja e mostrou como o déficit hídrico afeta a expressão destes ao longo dia. Os perfis transcricionais obtidos nesse estudo demonstraram também que genes importantes na resposta ao déficit hídrico em soja apresentam oscilação ao longo do dia, indicando sua regulação pelo ciclo circadiano.

O presente estudo gerou também um novo conjunto de genes para a normalização de dados de expressão de RT-qPCR, que serão essenciais para a acurácia dos dados de expressão gênica relativa. Esses normalizadores poderão ser amplamente empregados em futuros estudos combinando-se o déficit hídrico e as oscilações diurnas controladas pelo ciclo circadiano. Os resultados gerados contribuíram para melhor compreensão de como o ciclo circadiano influencia as respostas das plantas às mudanças no ambiente, como o déficit hídrico. Esses dados serão de grande importância para o desenvolvimento de futuras estratégias visando o melhoramento de soja, seja por engenharia genética ou por melhoramento convencional, que auxiliem na solução do problema da seca.

APÊNDICES

APÊNDICE A

Artigo 1: Figura suplementar S1

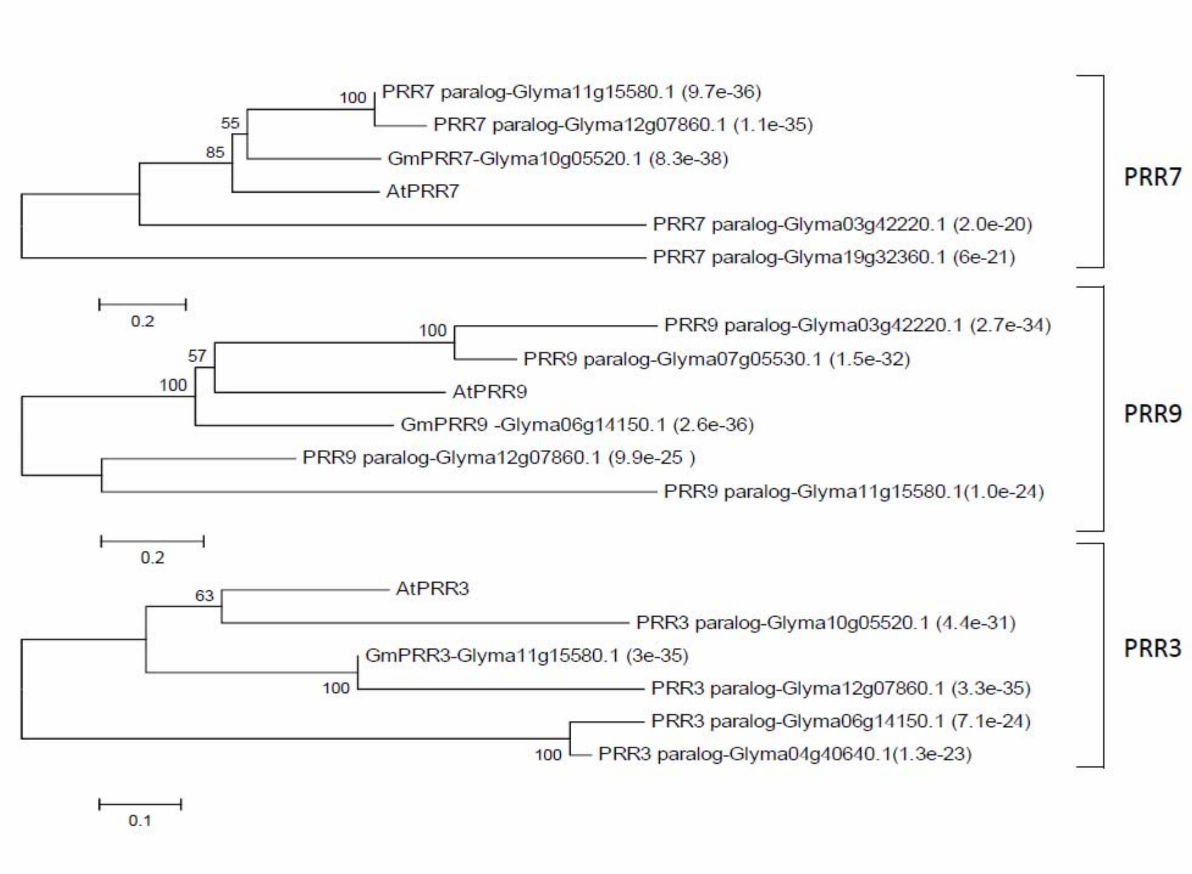


Fig. S1. Phylogenetic tree for the PRR genes. The proteins encoded by *Arabidopsis* AtPRR3, AtPRR7, and AtPRR9, the soybean GmPRR homologs, and its paralogs were used to construct the tree using the ClustalW algorithm in the MEGA 5 program. The Neighbor-Joining method was used with the following parameters: Poisson correction, pairwise deletion, and bootstrapping (1000 replicates; random seed).

APÊNDICE B

Artigo 1: Figura suplementar S2

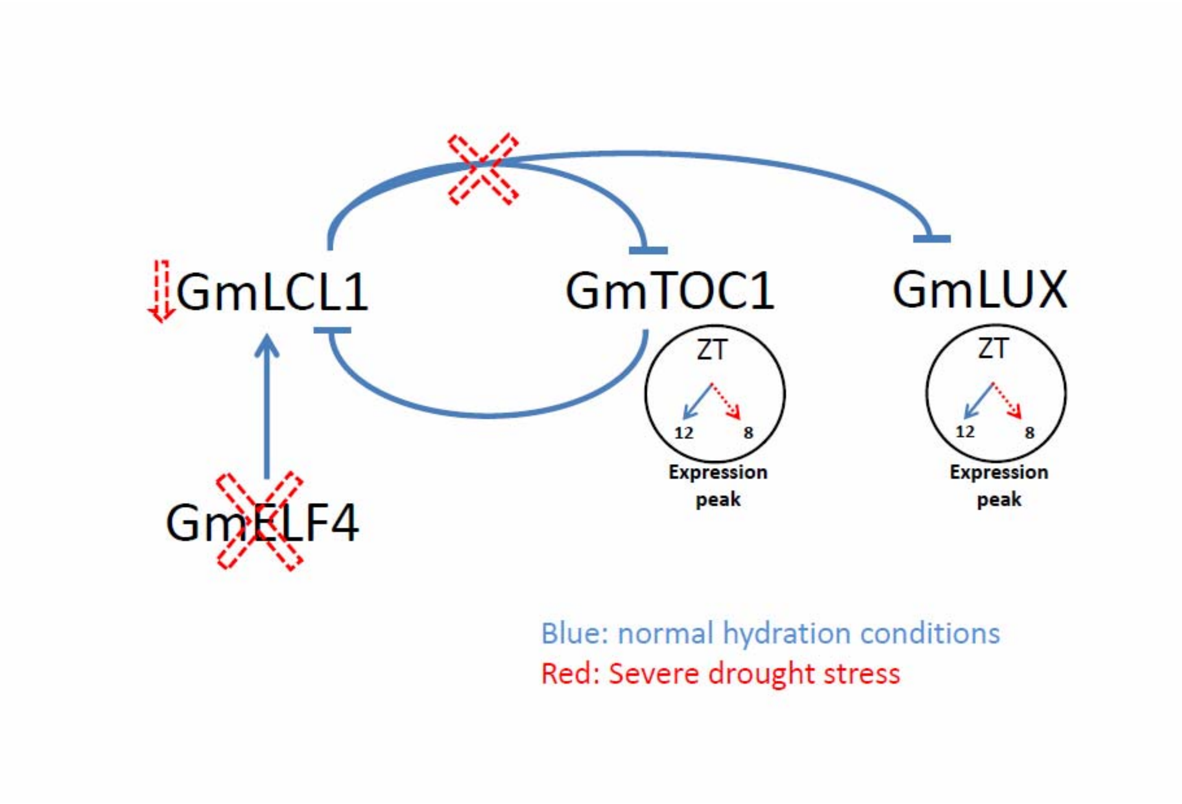


Fig. S2. Model of the impact of severe drought on circadian clock genes. Model of the impact of severe drought stress on *GmLCL1*, *GmTOC1*, *GmLUX* and *GmELF4-like* gene expression.

APÊNDICE C

Artigo 1: Figura suplementar S3

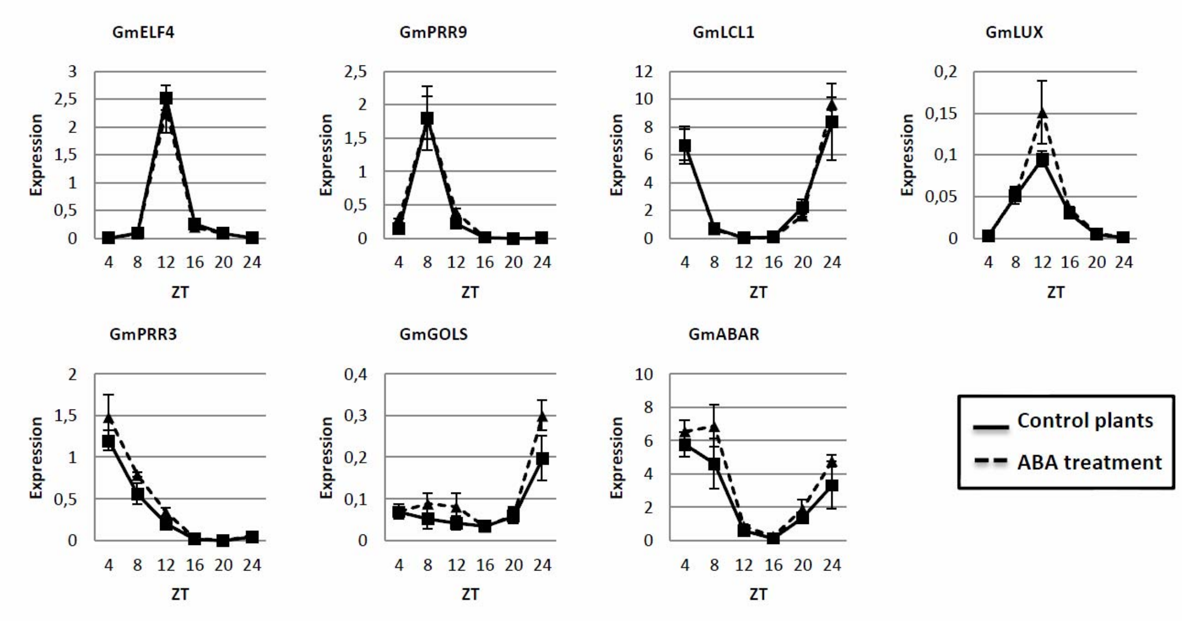


Fig. S3. The circadian clock genes that exhibited no response to ABA treatment. Gene expression data regards qPCR analysis. Expression axis represents normalized expression (NE)=$2^{\wedge-(\text{Ct experimental} - \text{Ctn})}$. Collect time points are represented by ZT (Zeitgeiber Time) 4 to 24, starting 4h after the lights came on (ZT4) and proceeding with 4 h intervals until ZT24.

APÊNDICE D

Artigo 1: Tabela suplementar S1

Table S1. Putative *cis*-regulatory elements on circadian clock genes. Putative *cis*-regulatory elements located in the promoter regions of the soybean circadian clock genes.

Gene	Matrix Family	Detailed Family Information	Matrix	Detailed Matrix Information	Sequence [a]	Start position[b]	End position[c]	Anchor position[d]	Strand[e]	Matrix sim.[f]	Core sim.
GmCHE	P$CCAF	Circadian control factors	P$EE.01	Evening element	atgaaaAATAtcatc	21	35	28	+	0.866	1
					taataaAATAtcgat	119	133	126	+	0.864	1
					aattaaCATAtctcg	169	183	176	+	0.863	0.75
	P$CE1F	Coupling element 1 binding factors	P$SBOX.01	Sugar and ABA responsive element conserved in several rbcS promoters	tattCACCtccat	247	259	253	-	1	1
	P$DREB	Dehydration responsive element binding factors	P$HVDRF1.01	H. vulgare dehydration-response factor 1	caccACCGccctcgc	581	595	588	+	0.946	1
	P$EREF	Ethylen respone element	P$ANT.01	ANT (Arabidopsis protein AINTEGUMENTA),	ggccccagTCCCcatgttc	414	432	423	+	0.826	1

		factors		member of the plant-specific family of AP2/EREBP-transcription factors							
GmCHE (continued)	P$GAPB	GAP-Box (light response elements)	P$GAP.01	Cis-element in the GAPDH promoters conferring light inducibility	tataATGAaaaatat	17	31	24	+	0.934	1
					gagaATGAacataac	69	83	76	+	0.88	1
					taaaATGAagagtag	262	276	269	+	0.959	1
					agaaATGAaaaaatg	298	312	305	+	0.91	1
					aagcATGAaaagcat	393	407	400	-	0.937	1
					atgaATGAagaatcc	439	453	446	-	0.894	1
	P$LREM	Light responsive element motif, not modulated by different light qualities	P$RAP22.01	RAP2.2, involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	tcATCTataca	31	41	36	+	0.878	1
					ttATCTaatgc	386	396	391	+	0.921	1
	P$MIIG	MYB IIG-type binding sites	P$ACTYP.01	AC-type motifs, MYB46/MYB83-responsive elements	ttaaaGTTGgtggtg	202	216	209	-	1	1
					gagtgGTTGgtggtg	527	541	534	-	1	1
					gtgttGTTGttggag	539	553	546	-	0.945	1
					ggttgGTTGctggtg	567	581	574	-	0.927	1
			P$P_ACT.01	Maize activator P of flavonoid biosynthetic genes	tggtGGTTggttgct	571	585	578	-	0.975	0.967
	P$MYBL	MYB-like proteins	P$WER.01	Myb-domain transcription factor werewolf	gggatattGTTAtaatg	7	23	15	+	0.876	1

GmCHE (continued)			P$GAMYB.01	GA-regulated myb gene from barley	gaagagtaGTTAtatgg	268	284	276	+	0.914	1
			P$CARE.01	CAACTC regulatory elements, GA-inducible	atggtagAGTTggcaca	336	352	344	+	0.834	1
					ccgtgagAGTTggacat	492	508	500	-	0.897	1
			P$NTMYBAS1.01	Anther-specific myb gene from tobacco	cggtggtgGTTGgttgc	572	588	580	-	0.961	0.949
	P$MYBS	MYB proteins with single DNA binding repeat	P$ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	aaatatcgaTATTttat	121	137	129	-	0.794	0.778
			P$PHR1.01	Phosphate starvation response 1	tcaaATATgcttaaagt	210	226	218	-	0.848	1
			P$MYBST1.01	MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat	ttcattATCCaccttcc	319	335	327	-	0.963	1
			P$ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	atctattttTATCtaat	378	394	386	+	0.851	1
	P$SALT	Salt/drought responsive elements	P$ALFIN1.02	Zinc-finger protein in alfalfa roots, regulates salt tolerance	cgagggcGGTGgtgg	580	594	587	-	0.962	1
	P$TCPF	DNA-binding proteins with the plant specific TCP-	P$PCF2.01	TCP class I transcription factor	ctggggCCCAcag	409	421	415	-	1	1
			P$PCF5.01	TCP class II transcription factor	tgtgGGCCcagt	410	422	416	+	0.977	0.897

		domain									
GmELF4	P$ABRE	ABA response elements	P$ABF1.03	ABA (abscisic acid) inducible transcriptional activator	gttgtcgtCGTGgcggg	459	475	467	-	0.84	1
GmELF4 (continued)	P$CCAF	Circadian control factors	P$CCA1.01	Circadian clock associated 1	taaaaaaaAATCata	75	89	82	+	0.865	1
			P$EE.01	Evening element	gatgtgAATAtcttg	217	231	224	+	0.898	1
	P$DREB	Dehydration responsive element binding factors	P$HVDRF1.01	H. vulgare dehydration-response factor 1	aaccACCGccgctcc	414	428	421	+	0.946	1
					cagcGCCGccaccgc	438	452	445	+	0.895	0.826
					cgccACCGccaccac	444	458	451	+	0.971	1
			P$CRT_DRE.01	C-repeat/dehydration response element	gatccCCGAcgtcct	520	534	527	-	0.927	1
					cagcaCCGActgcac	573	587	580	-	0.932	1
	P$EREF	Ethylen respone element factors	P$WRI1.01	WRINKLED 1	cACGAcgacaacaacggtg	464	482	473	+	0.893	0.962
					aACGGtgagcacgacgacc	476	494	485	+	0.913	0.848
					cTCGGgatccccgacgtcc	521	539	530	-	0.874	0.886
					gTCGGggatcccgaggcgt	525	543	534	+	0.899	0.886
	P$ERSE	ER stress-response elements	P$ERSE_I.01	ERSE I (ER stress-response element I)-like motif	ccgccaccaccccgcCACG	449	467	458	+	0.795	1
	P$GCCF	GCC box family	P$ERE_JERE.01	Ethylene-responsive elements (ERE) and jasmonate- and elicitor-responsive elements (JERE)	cagcgcCGCCacc	438	450	444	+	0.865	1
	P$MIIG	MYB IIG-type binding sites	P$NTLIM1.01	LIM domain protein binding to a PAL-box like sequence	ccaagttttgTTGGg	319	333	326	-	0.884	1

GmELF4 (continued)			P$PALBOXP.01	Putative cis-acting element in various PAL and 4CL gene promoters	cgGTGGttggagggg	407	421	414	-	0.85	0.936
			P$ATMYB15.01	R2R3-type myb-like transcription factor (IIG-type binding site)	cggcGGTGgttggag	410	424	417	-	0.792	0.775
	P$MYBL	MYB-like proteins	P$CARE.01	CAACTC regulatory elements, GA-inducible	ttatttgAGTTgaataa	96	112	104	-	0.837	1
			P$WER.01	Myb-domain transcription factor werewolf	tgctcaccGTTGttgtc	470	486	478	-	0.889	0.944
	P$MYBS	MYB proteins with single DNA binding repeat	P$MYBST1.01	MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat	ttttttATCCatgggca	189	205	197	-	0.959	1
			P$HVMCB1.01	Hordeum vulgare Myb-related CAB-promoter-binding protein 1	gaactcATCCtcctccg	368	384	376	+	0.941	1
	P$SALT	Salt/drought responsive elements	P$ALFIN1.01	Zinc-finger protein in alfalfa roots, regulates salt tolerance	tcgtcGTGGcggggt	457	471	464	-	0.93	1
GmGI	P$CCAF	Circadian control factors	P$CCA1.01	Circadian clock associated 1	caaacaaaAATAtca	117	131	124	+	0.859	0.757
			P$EE.01	Evening element	aacaaaAATAtcatt	119	133	126	+	0.858	1
	P$GBOX	Plant G-box/C-box bZIP	P$TGA1.01	Arabidopsis leucine zipper protein TGA1	cttccaTGACgaaacatcctc	20	40	30	+	0.911	1

GmGI (continued)		proteins	P$BZIP910.01	bZIP transcription factor from Antirrhinum majus	ggcgggTGATgtgggacttcc	178	198	188	-	0.777	0.75
	P$HEAT	Heat shock factors	P$HSE.01	Heat shock element	tagaataggcaAGAAct	303	319	311	-	0.853	1
					gcgagcctactAGAAta	313	329	321	-	0.842	1
	P$LREM	Light responsive element motif, not modulated by different light qualities	P$RAP22.01	RAP2.2, involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	aaATCTaataa	454	464	459	+	0.936	1
	P$MYBL	MYB-like proteins	P$MYBPH3.02	Myb-like protein of Petunia hybrida	aggaatTAGTtttgttt	60	76	68	+	0.786	1
			P$AS1_AS2_I.01	AS1/AS2 repressor complex binding motif I	atatatCTGTtaggcaa	220	236	228	+	1	1
			P$NTMYBAS1.01	Anther-specific myb gene from tobacco	ggggggggtGTTAggaag	353	369	361	+	0.984	1
			P$AS1_AS2_I.01	AS1/AS2 repressor complex binding motif I	atatatCTGTtagacaa	407	423	415	+	1	1
			P$CARE.01	CAACTC regulatory elements, GA-inducible	atcttggAGTTgaaggc	538	554	546	-	0.861	1
			P$GAMYB.01	GA-regulated myb gene from barley	gatgggctGTTGccaat	552	568	560	+	0.91	0.885
	P$MYBS	MYB proteins with single	P$PHR1.01	Phosphate starvation response 1	tagaATAGgcaagaact	303	319	311	-	0.848	0.75

		DNA binding repeat									
GmGI (continued)	P$MYCL	Myc-like basic helix-loop-helix binding factors	P$ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	cctccACATatgcctgcca	336	354	345	-	0.953	1
	P$SALT	Salt/drought responsive elements	P$ALFIN1.02	Zinc-finger protein in alfalfa roots, regulates salt tolerance	gagggggGGTGttag	351	365	358	+	0.952	1
	P$TCPF	DNA-binding proteins with the plant specific TCP-domain	P$ATTCP20.01	TCP class I transcription factor (Arabidopsis)	aacaGCCCatctt	550	562	556	-	0.941	1
GmJMJ	P$ABRE	ABA response elements	P$ABRE.01	ABA response elements	ctaataaACGTggcttc	141	157	149	+	0.84	1
					ataaaataCGTGgcttc	200	216	208	+	0.836	1
					aggtctcACGTcgcaat	250	266	258	+	0.846	1
					aattgcgACGTgagacc	251	267	259	-	0.826	1
	P$CCAF	Circadian control factors	P$EE.01	Evening element	caaaaaAATAtcata	341	355	348	+	0.871	1
					gacaaaAATAtctga	367	381	374	+	0.991	1
					gagaaaAATAtctga	404	418	411	+	0.999	1
					ttctaaAGTAtctta	451	465	458	-	0.861	0.75
					tttagaAATAtctcc	459	473	466	+	0.961	1
	P$GBOX	Plant G-box/C-box bZIP proteins	P$HBP1B.01	Wheat bZIP transcription factor HBP1B (histone gene binding	attagcaaACGTcacagaaat	125	145	135	-	0.919	1

GmJMJ (continued)				protein 1b)						
			P$TGA1.01	Arabidopsis leucine zipper protein TGA1	tttctgTGACgtttgctaata	126	146	136	+	0.979	1
			P$HBP1A.01	HBP-1a, suggested to be involved in the cell cycle-dependent expression	gtgaagcCACGtttattagca	139	159	149	-	0.919	1
			P$GBF1.01	bZIP protein G-Box binding factor 1	gctaataaACGTggcttcact	140	160	150	+	0.943	1
			P$HBP1A.01	HBP-1a, suggested to be involved in the cell cycle-dependent expression	atgaagcCACGtattttatta	198	218	208	-	0.914	1
			P$GBF1.01	bZIP protein G-Box binding factor 1	aataaaatACGTggcttcata	199	219	209	+	0.957	1
			P$HBP1B.01	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	taattgcgACGTgagacctgt	248	268	258	-	0.831	1
	P$HEAT	Heat shock factors	P$HSE.01	Heat shock element	tggatatattaAGAAta	33	49	41	-	0.837	1
			P$HSFA1A.01	Arabidopsis thaliana class A heat shock factor 1a	aagcaaCTTCaaaaaac	268	284	276	+	0.792	1
	P$MIIG	MYB IIG-type binding sites	P$ATMYB15.01	R2R3-type myb-like transcription factor (IIG-type binding site)	taacGGTAggatgga	46	60	53	-	0.794	1

GmJMJ (continued)	P$MYBL	MYB-like proteins	P$WER.01	Myb-domain transcription factor werewolf	atcctaccGTTAtgaaa	49	65	57	+	0.9	1
			P$MYBPH3.02	Myb-like protein of Petunia hybrida	cggagtTCGTcaacatt	584	600	592	-	0.786	0.779
	P$MYBS	MYB proteins with single DNA binding repeat	P$OSMYBS.01	Rice MYB proteins with single DNA binding domains, binding to the amylase element (TATCCA)	taataTATCcatcctac	39	55	47	+	0.896	1
			P$TAMYB80.01	MYB protein from wheat	cgggATATtcagatatt	373	389	381	-	0.934	1
			P$PHR1.01	Phosphate starvation response 1	ctgaATATcccgtgtgc	378	394	386	+	0.984	1
GmCLC1	P$ABRE	ABA response elements	P$ABF1.03	ABA (abscisic acid) inducible transcriptional activator	ccggccagCGTGgccac	22	38	30	-	0.843	1
					tccggtgaCGTGgcgag	108	124	116	-	0.969	1
					tcgcgACAAgtggcgca	322	338	330	-	0.853	0.75
	P$ERSE	ER stress-response elements	P$ERSE_I.01	ERSE I (ER stress-response element I)-like motif	ccaaaagtgagtggcCACG	12	30	21	+	0.91	1
					ccagtaacattgcgcCACT	312	330	321	+	0.843	0.75
	P$GARP	Myb-related DNA binding proteins (Golden2, ARR, Psr)	P$ARR10.01	Type-B response regulator (ARR10), member of the GARP-family of plant myb-related DNA binding motifs	AGATtcgag	193	201	197	+	0.99	1
	P$GBOX	Plant G-box/C-box bZIP proteins	P$TGA1.01	Arabidopsis leucine zipper protein TGA1	ctccggTGACgtggcgaggtc	105	125	115	-	0.988	1
			P$HBP1A.01	HBP-1a, suggested	acctcgcCACGtcaccggagg	106	126	116	+	1	1

GmCLC1 (continued)				to be involved in the cell cycle-dependent expression							
			P$OSBZ8.01	Oryza sativa bZIP protein 8	atcgcgacAAGTggcgcaatg	319	339	329	-	0.854	0.75
	P$HEAT	Heat shock factors	P$HSE.01	Heat shock element	gtgaatcaatcAGAAga	526	542	534	-	0.815	1
					cagaaaaatcgAGAAgt	541	557	549	-	0.934	1
	P$LREM	Light responsive element motif, not modulated by different light qualities	P$RAP22.01	RAP2.2, involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	cgATCTaatat	581	591	586	+	0.916	1
	P$MYBL	MYB-like proteins	P$NTMYBAS1.01	Anther-specific myb gene from tobacco	tactggctGTTGgtgct	301	317	309	-	0.968	0.949
	P$MYBS	MYB proteins with single DNA binding repeat	P$MYBST1.01	MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat	ggcgatATCCggccagc	30	46	38	-	0.918	1
			P$ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	tcttttttttTATCtctc	201	217	209	-	0.808	1
	P$MYCL	Myc-like basic helix-loop-helix binding factors	P$PIL5.01	Phytochrome interacting factor3-like 5	cctcgcCACGtcaccggag	107	125	116	+	0.928	1
			P$OSIRO2.01	Rice iron-related transcription factor	tcgcgaCAAGtggcgcaat	320	338	329	-	0.856	0.75

				2							
GmCLC1 (continued)	P$TCPF	DNA-binding proteins with the plant specific TCP-domain	P$ATTCP20.01	TCP class I transcription factor (Arabidopsis)	ctagGCCCaaatg	51	63	57	+	0.978	1
					gtggggCCCTccg	121	133	127	-	0.895	0.784
					gagggcCCCAcca	123	135	129	+	0.997	1
GmLUX	P$ABRE	ABA response elements	P$ABF1.01	ABA (abscisic acid) inducible transcriptional activator	gaaacACACgtgcctcc	77	93	85	-	0.843	1
					gtcggCCACgtaggcac	210	226	218	+	0.805	1
					ggtgcctaCGTGgccga	211	227	219	-	0.939	1
	P$CCAF	Circadian control factors	P$EE.01	Evening element	cctctaAATAtctct	175	189	182	-	0.961	1
					taaaaaAATAtctgg	191	205	198	+	0.994	1
					aagtgcAATAtctta	248	262	255	-	0.899	1
					tcccgaAATAtcttc	380	394	387	+	0.95	1
	P$DREB	Dehydration responsive element binding factors	P$CRT_DRE.01	C-repeat/dehydration response element	cgtggCCGActtagc	205	219	212	-	0.928	1
	P$GBOX	Plant G-box/C-box bZIP proteins	P$CPRF.01	Common plant regulatory factor (CPRF) from parsley	cgaaacacACGTgcctccgcc	74	94	84	-	0.971	1
					gcggaggcACGTgtgtttcgt	75	95	85	+	0.97	1
			P$ABZ1.01	Anaerobic basic leucine zipper	tggtgcctACGTggccgactt	208	228	218	-	0.969	1
			P$HBP1A.01	HBP-1a, suggested to be involved in the cell cycle-dependent expression	agtcggcCACGtaggcaccac	209	229	219	+	0.914	1
	P$MYBL	MYB-like	P$MYBPH3.01	Myb-like protein of	ataaaacaATTAagtat	25	41	33	+	0.813	0.75

GmLUX (continued)		proteins		Petunia hybrida							
			P$NTMYBAS1.01	Anther-specific myb gene from tobacco	agatgggtGTTGgggtg	227	243	235	-	0.97	0.949
			P$MYBPH3.01	Myb-like protein of Petunia hybrida	ctgaaacgGTTAtggta	510	526	518	-	0.964	1
	P$MYBS	MYB proteins with single DNA binding repeat	P$ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	tcttttcttTATTattt	150	166	158	-	0.808	0.778
					aaatatctcTATCtttc	168	184	176	-	0.986	1
					ttttttATCCtctaaat	181	197	189	-	0.958	1
					tcttctttcTATCtatc	433	449	441	+	0.838	1
					ctttctatcTATCcaaa	437	453	445	+	0.806	1
			P$MYBST1.01	MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat	ctatctATCCaaatcca	441	457	449	+	0.928	1
	P$MYCL	Myc-like basic helix-loop-helix binding factors	P$OSBHLH66.01	Rice bHLH protein	gaaacaCACGtgcctccgc	75	93	84	-	0.954	1
			P$PIL5.01	Phytochrome interacting factor3-like 5	cggaggCACGtgtgtttcg	76	94	85	+	0.947	1
			P$PIF3.01	Phytochrome B-regulated transcription factor PIF3 (nuclear-localized bHLH-protein)	gtcggcCACGtaggcacca	210	228	219	+	0.836	1
			P$AMS.01	ABORTED MICROSPORES	aacaccCATCtggaataag	233	251	242	+	0.972	1
GmPRR3	P$ABRE	ABA response elements	P$ABF1.02	ABA (abscisic acid) inducible transcriptional	gtgagCCACgtggtggg	271	287	279	+	0.895	1
					tcccaCCACgtggctca	272	288	280	-	0.97	1
					tttgcagaCGTGgacaa	350	366	358	-	0.822	1

GmPRR3 (continued)				activator	ccacatcACGTgggccc	386	402	394	+	0.834	1
					gcgttCCACgtggagca	404	420	412	+	0.881	1
					gtgctCCACgtggaacg	405	421	413	-	0.887	1
	P$CCAF	Circadian control factors	P$EE.01	Evening element	cacaatAATAtcttc	462	476	469	+	0.941	1
	P$ERSE	ER stress-response elements	P$ERSE_I.01	ERSE I (ER stress-response element I)-like motif	gcatttgattttgtcCACG	340	358	349	+	0.814	1
	P$GAPB	GAP-Box (light response elements)	P$GAP.01	Cis-element in the GAPDH promoters conferring light inducibility	tctaATGAaaaaaac	521	535	528	-	0.96	1
	P$GARP	Myb-related DNA binding proteins (Golden2, ARR, Psr)	P$ARR10.01	Type-B response regulator (ARR10), member of the GARP-family of plant myb-related DNA binding motifs	AGATctgga	289	297	293	-	0.973	1
					AGATctgga	292	300	296	+	0.973	1
	P$GBOX	Plant G-box/C-box bZIP proteins	P$TAF1.01	Tobacco bZip transcription activator (TAF-1)	atcccaccACGTggctcactc	269	289	279	-	0.991	1
			P$TAF1.01	Tobacco bZip transcription activator (TAF-1)	agtgagccACGTggtgggatc	270	290	280	+	0.991	1
			P$ABZ1.01	Anaerobic basic leucine zipper	atttgcagACGTggacaaaat	347	367	357	-	0.98	1
			P$HBP1A.01	HBP-1a, suggested to be involved in the cell cycle-dependent expression	ttttgtcCACGtctgcaaatc	348	368	358	+	0.918	1

GmPRR3 (continued)			P$OSBZ8.01	Oryza sativa bZIP protein 8	cccacgtgATGTggcgcgttc	379	399	389	-	0.855	0.75
			P$ROM.01	Regulator of MAT (ROM1, ROM2)	aacgcgCCACatcacgtgggc	380	400	390	+	0.86	1
			P$CPRF.01	Common plant regulatory factor (CPRF) from parsley	ctgggcccACGTgatgtggcg	384	404	394	-	0.978	1
			P$GBF1.01	bZIP protein G-Box binding factor 1	gccacatcACGTgggcccagc	385	405	395	+	0.983	1
			P$CPRF.01	Common plant regulatory factor (CPRF) from parsley	ggtgctccACGTggaacgctg	402	422	412	-	0.998	1
			P$CPRF.01	Common plant regulatory factor (CPRF) from parsley	agcgttccACGTggagcaccg	403	423	413	+	0.998	1
			P$GBF1.01	bZIP protein G-Box binding factor 1	agatgctgACGTgtataaaac	440	460	450	-	0.999	1
			P$HBP1B.01	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	ttttatacACGTcagcatctc	441	461	451	+	0.922	1
			P$BZIP910.02	bZIP transcription factor from Antirrhinum majus	ttgtggagatgcTGACgtgta	446	466	456	-	0.884	1
	P$HEAT	Heat shock factors	P$HSE.01	Heat shock element	aagaaaagaaaAGAAaa	563	579	571	+	0.841	1
	P$IBOX	Plant I-Box sites	P$IBOX.01	I-Box in rbcS genes and other light regulated genes	tatatGATTaggtttca	72	88	80	-	0.812	0.75
			P$GATA.01		gaatgGATAattggata	91	107	99	+	0.932	1
			P$IBOX.01		caaaaGAAAagaaaaga	560	576	568	+	0.816	0.75
			P$IBOX.01		gaaaaGAAAagaaaaac	565	581	573	+	0.823	0.75
	P$MYBL	MYB-like	P$CARE.01	CAACTC regulatory	cttagtgAGTTagtaga	19	35	27	+	0.839	1

GmPRR3 (continued)		proteins		elements, GA-inducible							
			P$MYBPH3.02	Myb-like protein of Petunia hybrida	gtgagtTAGTagacacg	23	39	31	+	0.798	1
			P$ATMYB77.01	R2R3-type myb-like transcription factor (I-type binding site)	ccttgaCAGTtgacaca	164	180	172	-	0.93	1
			P$GAMYB.01	GA-regulated myb gene from barley	tttcttttGTTAttccg	552	568	560	-	0.914	1
	P$MYBS	MYB proteins with single DNA binding repeat	P$MYBST1.01	MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat	ccaattATCCattcagt	88	104	96	-	0.937	1
			P$HVMCB1.01	Hordeum vulgare Myb-related CAB-promoter-binding protein 1	atatatATCCaattatc	96	112	104	-	0.952	1
			P$ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	ttatatataTATCcaat	100	116	108	-	0.847	1
					ttttatataTATAtcca	102	118	110	-	0.794	0.778
	P$MYCL	Myc-like basic helix-loop-helix binding factors	P$OSIRO2.01	Rice iron-related transcription factor 2	tcccacCACGtggctcact	270	288	279	-	0.973	1
					gtgagcCACGtggtgggat	271	289	280	+	0.977	1
			P$PIL5.01	Phytochrome interacting factor3-like 5	tgggccCACGtgatgtggc	385	403	394	-	0.989	1
			P$OSIRO2.01	Rice iron-related transcription factor 2	ccacatCACGtgggcccag	386	404	395	+	0.976	1
					gtgctcCACGtggaacgct	403	421	412	-	0.971	1
					gcgttcCACGtggagcacc	404	422	413	+	0.971	1
			P$PIL5.01	Phytochrome interacting factor3-	tttataCACGtcagcatct	442	460	451	+	0.912	1

				like 5							
GmPRR3 (continued)	P$TCPF	DNA-binding proteins with the plant specific TCP-domain	P$PCF2.01	TCP class I transcription factor	gctgggCCCAcgt	393	405	399	-	0.994	1
			P$PCF5.01	TCP class II transcription factor	cgtgGGCCcagcg	394	406	400	+	0.974	0.897
GmPRR7	P$CCAF	Circadian control factors	P$CCA1.01	Circadian clock associated 1	gtaagtaaAATCtcg	26	40	33	+	0.935	1
	P$GBOX	Plant G-box/C-box bZIP proteins	P$HBP1B.01	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	aatccaaaACGTcaaagcagt	1	21	11	-	0.899	1
			P$TGA1.01	Arabidopsis leucine zipper protein TGA1	ctgcttTGACgttttggattt	2	22	12	+	0.958	1
	P$IBOX	Plant I-Box sites	P$IBOX.01	I-Box in rbcS genes and other light regulated genes	aagaaGTTAaggttgaa	526	542	534	+	0.813	0.75
	P$LREM	Light responsive element motif, not modulated by different light qualities	P$RAP22.01	RAP2.2, involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	atATCTagtga	272	282	277	+	0.853	1
	P$MIIG	MYB IIG-type binding sites	P$PALBOXP.01	Putative cis-acting element in various PAL and 4CL gene	ccGTGGttagttttg	119	133	126	+	0.81	0.936

GmPRR7 (continued)				promoters							
			P$ACTYP.01	AC-type motifs, MYB46/MYB83-responsive elements	tgagtGTTGatggcg	406	420	413	+	0.927	1
	P$MYBL	MYB-like proteins	P$MYBPH3.02	Myb-like protein of Petunia hybrida	cgtggtTAGTtttgaat	120	136	128	+	0.816	1
			P$WER.01	Myb-domain transcription factor werewolf	gatatgcaGTTAaaaga	260	276	268	-	0.918	1
			P$CARE.01	CAACTC regulatory elements, GA-inducible	tctagtgATTTgcagtt	275	291	283	+	0.839	0.75
	P$MYBS	MYB proteins with single DNA binding repeat	P$OSMYBS.01	Rice MYB proteins with single DNA binding domains, binding to the amylase element (TATCCA)	ccataTATCcacagcac	192	208	200	-	0.846	1
			P$PHR1.01	Phosphate starvation response 1	tggtATATgctgctgga	571	587	579	-	0.975	1
			P$TAMYB80.01	MYB protein from wheat	cagcATATaccacagcc	576	592	584	+	0.961	1
GmPRR9	P$ABRE	ABA response elements	P$ABF1.02	ABA (abscisic acid) inducible transcriptional activator	cagagCCACgtgtctcg	177	193	185	+	0.871	1
					ccgagACACgtggctct	178	194	186	-	0.958	1
					attgctgaCGTGgatca	485	501	493	+	0.835	1
					ttggaagaGGTGgccca	504	520	512	-	0.829	0.75
	P$CCAF	Circadian control factors	P$EE.01	Evening element	ctcaaaAATAtcttt	230	244	237	+	0.987	1
					taaattAATAtcttt	258	272	265	+	0.92	1
					acataaAATAtccac	310	324	317	+	0.864	1

GmPRR9 (continued)					ccaaaaAATAtccga	337	351	344	-	0.87	1
					tgaaaaAATAtccac	403	417	410	+	0.866	1
					tgggaaAATAtctag	525	539	532	-	0.993	1
					gccgaaAATAtcctg	543	557	550	-	0.856	1
	P$CGCG	Calmodulin binding / CGCG box binding proteins	P$ATSR1.01	Arabidopsis thaliana signal-responsive gene1, Ca2+/ calmodulin binding protein homolog to NtER1 (tobacco early ethylene-responsive gene)	gggCGCGtgaggagctc	100	116	108	+	0.876	1
	P$EREF	Ethylen respone element factors	P$WRI1.01	WRINKLED 1	cTCGGtgtcgtccatgtga	190	208	199	+	0.862	0.886
	P$ERSE	ER stress-response elements	P$ERSE_I.01	ERSE I (ER stress-response element I)-like motif	cccctaggacagagcCACG	168	186	177	+	0.866	1
	P$GBOX	Plant G-box/C-box bZIP proteins	P$BZIP911.01	bZIP transcription factor from Antirrhinum majus	aatgcgTGAAgtggcctgcat	26	46	36	-	0.826	0.75
			P$EMBP1.01	bZIP transcription factor implicated in ABA induced gene expression	accgagacACGTggctctgtc	175	195	185	-	1	1
			P$HBP1A.01	HBP-1a, suggested to be involved in the cell cycle-dependent expression	acagagcCACGtgtctcggtg	176	196	186	+	0.983	1

GmPRR9 (continued)			P$BZIP910.02	bZIP transcription factor from Antirrhinum majus	agatctcattgcTGACgtgga	478	498	488	+	0.853	1
			P$UPRE.01	UPRE (unfolded protein response element) like motif	catgatCCACgtcagcaatga	483	503	493	-	0.964	1
			P$GBF1.01	bZIP protein G-Box binding factor 1	cattgctgACGTggatcatgt	484	504	494	+	0.988	1
			P$EMBP1.01	bZIP transcription factor implicated in ABA induced gene expression	atctagacACTTggaagaggt	510	530	520	-	0.847	0.75
	P$LREM	Light responsive element motif, not modulated by different light qualities	P$RAP22.01	RAP2.2, involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	ctATCTacgag	122	132	127	-	0.879	1
					agATCTaaagc	454	464	459	-	0.971	1
					atATCTagaca	522	532	527	-	0.902	1
	P$MYBS	MYB proteins with single DNA binding repeat	P$PHR1.01	Phosphate starvation response 1	atgaACATgcaggccac	20	36	28	+	0.848	0.75
			P$OSMYBS.01	Rice MYB proteins with single DNA binding domains, binding to the amylase element (TATCCA)	gcgagTATCcagcacga	75	91	83	+	0.823	1
					taaaaTATCcacatgaa	313	329	321	+	0.84	1
					aaaaaTATCcaccacca	406	422	414	+	0.823	1
	P$MYCL	Myc-like basic helix-loop-helix binding	P$OSIRO2.01	Rice iron-related transcription factor 2	ccgagaCACGtggctctgt	176	194	185	-	0.99	1
			P$PIL5.01	Phytochrome	cagagcCACGtgtctcggt	177	195	186	+	0.965	1

GmPRR9 (continued)		factors		interacting factor3-like 5	atgatcCACGtcagcaatg	484	502	493	-	0.921	1
	P$SALT	Salt/drought responsive elements	P$ALFIN1.01	Zinc-finger protein in alfalfa roots, regulates salt tolerance	gtttgGTGGtggata	411	425	418	-	0.953	1
	P$TCPF	DNA-binding proteins with the plant specific TCP-domain	P$PCF2.01	TCP class I transcription factor	cacgcgCCCActg	96	108	102	-	0.892	1
			P$PCF5.01	TCP class II transcription factor	agtgGGCCccgcc	279	291	285	-	0.976	0.897
			P$PCF2.01	TCP class I transcription factor	gcggggCCCActc	280	292	286	+	0.999	1
					aggtggCCCAcat	501	513	507	-	0.89	1
GmTOC1	P$ABRE	ABA response elements	P$ABF1.03	ABA (abscisic acid) inducible transcriptional activator	tcaggtgaCGTGgccgg	98	114	106	+	0.988	1
			P$ABRE.01		cctttccACGTctccct	187	203	195	+	0.901	1
			P$ABF1.03		gagggagaCGTGgaaag	188	204	196	-	0.827	1
			P$ABF1.01		agtgcACACggggcgtg	215	231	223	-	0.828	1
	P$CCAF	Circadian control factors	P$EE.01	Evening element	atgaatAATAtctat	11	25	18	-	0.949	1
					gtctaaAATAtccca	35	49	42	-	0.859	1
	P$DREB	Dehydration responsive element binding factors	P$CRT_DRE.01	C-repeat/dehydration response element	cttcgCCGAcctctg	361	375	368	+	0.961	1
	P$ERSE	ER stress-response elements	P$ERSE_I.01	ERSE I (ER stress-response element I)-like motif	ccctctttcctacgcCACG	200	218	209	+	0.793	1
	P$GBOX	Plant G-box/C-box bZIP proteins	P$HBP1A.01	HBP-1a, suggested to be involved in the cell cycle-dependent	agccggcCACGtcacctgaag	96	116	106	-	1	1

GmTOC1 (continued)				expression							
			P$BZIP911.01	bZIP transcription factor from Antirrhinum majus	ttcaggTGACgtggccggctc	97	117	107	+	1	1
			P$ABZ1.01	Anaerobic basic leucine zipper	agagggagACGTggaaaggaa	185	205	195	-	0.977	1
			P$HBP1A.01	HBP-1a, suggested to be involved in the cell cycle-dependent expression	tcctttcCACGtctccctctt	186	206	196	+	0.947	1
			P$OSBZ8.01	Oryza sativa bZIP protein 8	cacacgggGCGTggcgtagga	207	227	217	-	0.848	0.75
					cagtgcacACGGggcgtggcg	212	232	222	-	0.844	0.75
			P$HBP1B.01	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	ccccgcgcACGTtagctcacc	244	264	254	+	0.831	1
			P$BZIP910.02	bZIP transcription factor from Antirrhinum majus	aagggggtgagcTAACgtgcg	249	269	259	-	0.846	0.75
			P$HBP1B.01	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	gacggcgcACGTtagcagata	418	438	428	-	0.837	1
	P$HEAT	Heat shock factors	P$HSFA1A.01	Arabidopsis thaliana class A heat shock factor 1a	cagaacTTTCataaaaa	526	542	534	-	0.861	1
	P$LREM	Light	P$RAP22.01	RAP2.2, involved in	atATCTattaa	8	18	13	-	0.868	1

GmTOC1 (continued)		responsive element motif, not modulated by different light qualities		carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes							
	P$MIIG	MYB IIG-type binding sites	P$ATMYB84.01	R2R3-type myb-like transcription factor (IIG-type binding site)	agggagtgGGTGgtc	140	154	147	-	0.838	1
			P$MYBC1.01	Maize C1 myb-domain protein	gagggGTAGttggtg	563	577	570	+	0.95	1
			P$PALBOXL.01	Cis-acting element conserved in various PAL and 4CL promoters	gggtagttGGTGagt	566	580	573	+	0.837	1
	P$MYBL	MYB-like proteins	P$CARE.01	CAACTC regulatory elements, GA-inducible	tctggatAGTTgggtag	125	141	133	-	0.862	1
			P$NTMYBAS1.01	Anther-specific myb gene from tobacco	gaggggtaGTTGgtgag	563	579	571	+	0.965	0.949
			P$MYBPH3.02	Myb-like protein of Petunia hybrida	ggtagtTGGTgagtgag	567	583	575	+	0.784	0.779
	P$MYBS	MYB proteins with single DNA binding repeat	P$MYBST1.01	MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat	ccaactATCCagaccac	129	145	137	+	0.902	1
	P$MYCL	Myc-like basic helix-loop-helix	P$PIL5.01	Phytochrome interacting factor3-like 5	gccggcCACGtcacctgaa	97	115	106	-	0.928	1
					cctttcCACGtctccctct	187	205	196	+	0.892	1

GmTOC1 (continued)		binding factors									
	P$SALT	Salt/drought responsive elements	P$ALFIN1.02	Zinc-finger protein in alfalfa roots, regulates salt tolerance	ggaagggGGTGagct	257	271	264	-	0.961	1
					gaaaagcGGTGccta	288	302	295	-	0.959	1
GmZTL	P$CCAF	Circadian control factors	P$CCA1.01	Circadian clock associated 1	aacaaaaaAATCtac	72	86	79	+	0.968	1
			P$EE.01	Evening element	actacaAATAtattt	229	243	236	+	0.871	1
					ttttaaAATAtattt	234	248	241	-	0.902	1
					tacaaaAATAtattt	318	332	325	+	0.896	1
					tgggaaAATAtattt	323	337	330	-	0.902	1
	P$GAPB	GAP-Box (light response elements)	P$GAP.01	Cis-element in the GAPDH promoters conferring light inducibility	ccacATGTaaaagaa	459	473	466	+	0.89	0.769
					ggaaATGAatatgat	577	591	584	-	0.915	1
	P$LREM	Light responsive element motif, not modulated by different light qualities	P$RAP22.01	RAP2.2, involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	aaATCTactta	79	89	84	+	0.874	1
					tcATCTacttt	484	494	489	+	0.86	1
	P$MIIG	MYB IIG-type binding sites	P$PALBOXP.01	Putative cis-acting element in various PAL and 4CL gene promoters	aaGGGGtgagttgag	511	525	518	-	0.85	1
	P$MYBL	MYB-like proteins	P$MYBPH3.02	Myb-like protein of Petunia hybrida	ctctgtTAGTtaaaaaa	6	22	14	-	0.897	1
			P$AS1_AS2_I.01	AS1/AS2 repressor complex binding	ttaactCTGTtagttaa	10	26	18	-	1	1

GmZTL (continued)				motif I							
			P$MYBPH3.02	Myb-like protein of Petunia hybrida	taaattTAGTtttttta	104	120	112	-	0.792	1
			P$MYBPH3.02		tgtattTGGTtaagaga	305	321	313	-	0.784	0.779
			P$MYBPH3.02		ttgaggTTGTtccagaa	499	515	507	-	0.768	0.817
			P$CARE.01	CAACTC regulatory elements, GA-inducible	aggggtgAGTTgaggtt	508	524	516	-	0.867	1
	P$MYBS	MYB proteins with single DNA binding repeat	P$PHR1.01	Phosphate starvation response 1	tagaATATggagatact	215	231	223	+	0.848	1
			P$PHR1.01		atgaATATgatgatgag	571	587	579	-	0.848	1
	P$SALT	Salt/drought responsive elements	P$ALFIN1.02	Zinc-finger protein in alfalfa roots, regulates salt tolerance	agaaaggGGTGagtt	514	528	521	-	0.953	1

[a] The element sequence is shown in “sequence”, whereby the capital letters represent the element core.

[bc] The “start” and “end” positions correspond to the position of the cis-element on the promoter sequence.

[d] The “anchor” position corresponds to the position of the element core.

[e] The “strand” corresponds to the DNA strand position.

[f] The “matrix similarity score” reflects the similarity between our sequence and the matrix sequence in which a value of 1 corresponds to sequences with the most conserved nucleotides at each position of the matrix. A matrix similarity >0.80 was used as the threshold.

APÊNDICE E

Artigo 1: Tabela suplementar S2

Table S2. Putative *cis*-regulatory elements on drought-responsive genes. Putative *cis*-regulatory elements located in the promoter regions of soybean drought-responsive genes.

Gene	Matrix Family	Detailed Family Information	Matrix	Detailed Matrix Information	Sequence [a]	Start position[b]	End position[c]	Anchor position[d]	Strand[e]	Matrix sim.[f]	Core sim.
GmABAR											
	P$ABRE	ABA response elements	P$ABF1.03	ABA (abscisic acid) inducible transcriptional activator	gtggatgaCATGtccta	217	233	225	-	0.823	0.75
					taagaccaCGTGaccaa	327	343	335	-	0.903	1
			P$ABRE.01	ABA response elements	tttggtcACGTggtctt	326	342	334	+	0.839	1
	P$CCAF	Circadian control factors	P$CCA1.01	Circadian clock associated 1	aaacaaaaAATCtca	35	49	42	+	0.994	1
			P$E**E**.01	Evening element	gggatgAAATAtctca	74	88	81	-	0.904	1
	P$GAPB	GAP-Box (light response elements)	P$GAP.01	Cis-element in the GAPDH promoters conferring light inducibility	actaATTAaaaaaaa	288	302	295	-	0.898	0.808
	P$LREM	Light responsive	P$RAP22.01	RAP2.2, involved in	aaATCTaaaag	5	15	10	+	0.99	1
					gaATCTatatc	20	30	25	+	0.88	1

		element motif, not modulated by different light qualities		carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes						8	
					atATCTaaaaa	26	36	31	+	1	1
GmABAR (continued)	P$MIIG	MYB IIG-type binding sites	P$PALBOXP.01	Putative cis-acting element in various PAL and 4CL gene promoters	gaGCGGttgttgaat	431	445	438	-	0.81	0.83
	P$MYBL	MYB-like proteins	P$AS1_AS2_I.01	AS1/AS2 repressor complex binding motif I	cactatCTGTtagtact	195	211	203	-	1	1
			P$MYBPH3.02	Myb-like protein of Petunia hybrida	gtcaatTAGTtaataaa	301	317	309	+	0.861	1
			P$WER.01	Myb-domain transcription factor werewolf	tatttacgGTTAgtggt	350	366	358	-	0.973	1
	P$MYBS	MYB proteins with single DNA binding repeat	P$ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	aggaatctaTATCtaaa	18	34	26	+	0.879	1
					atatctttcTATCtgtt	171	187	179	-	0.82	1
					tatttgctaTATCtttc	179	195	187	-	0.835	1
			P$OSMYBS.01	Rice MYB proteins with single DNA binding domains, binding to the	tgagaTATTcatccctc	74	90	82	+	0.826	0.75

GmABAR (continued)				amylase element (TATCCA)							
	P$MYCL	Myc-like basic helix-loop-helix binding factors	P$PIL5.01	Phytochrome interacting factor3-like 5	taagacCACGtgaccaaat	325	343	334	-	0.995	1
			P$OSIRO2.01	Rice iron-related transcription factor 2	tttggtCACGtggtcttaa	326	344	335	+	0.976	1
	P$DPBF	Dc3 promoter binding factors	P$DPBF.01	bZIP factors DPBF-1 and 2 (Dc3 promoter binding factor-1 and 2)	aCCACgtgacc	329	339	334	-	0.906	0.867
	P$GBOX	Plant G-box/C-box bZIP proteins	P$BZIP911.02	bZIP transcription factor from Antirrhinum majus	ggtggaTGACatgtcctagtt	214	234	224	-	0.869	1
			P$CPRF.01	Common plant regulatory factor (CPRF) from parsley	ttaagaccACGTgaccaaatg	324	344	334	-	0.983	1
			P$TAF1.01	Tobacco bZip transcription activator (TAF-1)	atttggtcACGTggtcttaat	325	345	335	+	0.988	1
GmAREB	P$ABRE	ABA response elements	P$ABRE.01	ABA response elements	ttgtcttACGTagcact	433	449	441	-	0.825	1
	P$CCAF	Circadian control factors	P$CCA1.01	Circadian clock associated 1	aaactacaAATCtat	278	292	285	-	0.865	1

GmAREB (continued)	P$GBOX	Plant G-box/C-box bZIP proteins	P$HBP1B.01	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	tgagtgctACGTaagacaaag	431	451	441	+	0.836	1
	P$HEAT	Heat shock factors	P$HSFA1A.01	Arabidopsis thaliana class A heat shock factor 1a	agtaagCTTCtaaaagt	460	476	468	+	0.767	1
					gatgaaCTTCaggaact	500	516	508	+	0.772	1
					accaaaGTTCctgaagt	505	521	513	-	0.852	0.857
			P$HSE.01	Heat shock element	acgatacttttAGAAgc	465	481	473	-	0.864	1
	P$LREM	Light responsive element motif, not modulated by different light qualities	P$RAP22.01	RAP2.2, involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	aaATCTattat	275	285	280	-	0.858	1
					agATCTaacca	375	385	380	-	0.901	1
					tgATCTactta	416	426	421	+	0.86	1
	P$MYBL	MYB-like proteins	P$NTMYBAS1.01	Anther-specific myb gene from tobacco	agaaggttGTTAacttc	53	69	61	-	0.977	1
			P$WER.01	Myb-domain transcription factor werewolf	tttatatgGTTAgatct	369	385	377	+	0.897	1
					aacattcaGTTAgaagg	64	80	72	-	0.938	1
			P$MYBPH3.02	Myb-like protein of Petunia hybrida	tcagatTAGTtaaaatt	152	168	160	-	0.77	1
					attaatTAGTtaaaaac	289	305	297	-	0.862	1

GmAREB (continued)			P$GAMYB.01	GA-regulated myb gene from barley	gtggggtgGTTAtctcc	525	541	533	-	0.978	1
	P$MYBS	MYB proteins with single DNA binding repeat	P$ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	ttatctcttTATCgaat	170	186	178	-	0.873	1
	P$SALT	Salt/drought responsive elements	P$ALFIN1.02	Zinc-finger protein in alfalfa roots, regulates salt tolerance	caagtggGGTGgtta	530	544	537	-	0.973	1
GmDREB1	P$GAPB	GAP-Box (light response elements)	P$GAP.01	Cis-element in the GAPDH promoters conferring light inducibility	caaaATTAaaaatat	266	280	273	+	0.889	0.808
					tgaaATGAaaatagt	583	597	590	-	0.88	1
	P$LREM	Light responsive element motif, not modulated by different light qualities	P$RAP22.01	RAP2.2, involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	gaATCTatatt	6	16	11	+	0.888	1
					cgATCTaaatt	359	369	364	+	0.961	1
	P$MYBL	MYB-like proteins	P$MYBPH3.02	Myb-like protein of Petunia hybrida	tttaatTAGTtgaaaaa	101	117	109	+	0.796	1
					ttaattTAGTttttaaa	158	174	166	-	0.786	1
					ttaaaacaTTTAtagaa	212	228	220	-	0.812	0.75

GmDREB1 (continued)					aaaatacgGTCAatcaa	285	301	293	+	0.802	0.75
			P$AS1_AS2_I.01	AS1/AS2 repressor complex binding motif I	tcgcagCTGTtatatag	565	581	573	+	1	1
	P$MYBS	MYB proteins with single DNA binding repeat	P$PHR1.01	Phosphate starvation response 1	ttgtATATtcataaatt	252	268	260	-	0.944	1
					atgaATATacaaaatta	257	273	265	+	0.979	1
	P$MYCL	Myc-like basic helix-loop-helix binding factors	P$MYCRS.01	Myc recognition sequences	ttaattcATGTgtttatta	226	244	235	-	0.942	0.864
			P$ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	aggaaACACttgtttaccc	503	521	512	-	0.985	0.955
					ggtaaACAAgtgtttcctg	504	522	513	+	0.973	0.955
	P$NACF	Plant specific NAC [NAM (no apical meristem), ATAF172, CUC2 (cup-shaped cotyledons 2)] transcription factors	P$ANAC019.01	Arabidopsis NAC domain containing protein 19	gaaaTACGaaaagaa	400	414	407	-	0.987	1
				Transcription factor NAC2	gactCAGGaaacact	512	526	519	-	0.862	0.803
	P$SPF1	Sweet potato	P$SP8BF.01	DNA-binding protein of sweet	aaTACTatattt	195	207	201	+	0.924	1

GmDREB1 (continued)		DNA-binding factor with two WRKY-domains		potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT) sequences of sporamin and beta-amylase genes	acTACTatgctat	319	331	325	-	0.917	1
					agTACTataaaaa	329	341	335	+	0.919	1
					acTACTatgactt	431	443	437	-	0.939	1
					agTACTatataac	573	585	579	-	0.919	1
					agTACTattttca	580	592	586	+	0.986	1
	P$WBXF	W Box family	P$WRKY11.01	Calmodulin binding WRKY transcription factor 11	attgaTTGAccgtattt	286	302	294	-	0.973	1
					gaccaTTGActactatg	435	451	443	-	0.962	1
			P$ERE.01	Elicitor response element	ggaaaaTGACcattgac	442	458	450	-	0.906	1
GmDREB2.1	P$ABRE	ABA response elements	P$ABRE.01	ABA response elements	gactgacACGTgtaagg	279	295	287	+	0.879	1
					accttacACGTgtcagt	280	296	288	-	0.95	1
	P$CCAF	Circadian control factors	P$EE.01	Evening element	ttaataAATAtatta	241	255	248	+	0.88	1
	P$CE1F	Coupling element 1 binding factors	P$SBOX.01	Sugar and ABA responsive element conserved in several rbcS promoters	cttcCACCtcaac	437	449	443	+	0.877	1
	P$EREF	Ethylen respone element factors	P$WRI1.01	WRINKLED 1	aGCGAagctcgcaaagtga	492	510	501	+	0.908	0.987

GmDREB2.1 (continued)	P$ERSE	ER stress-response elements	P$ERSE_I.01	ERSE I (ER stress-response element I)-like motif	ccagtacatgactgaCACG	270	288	279	+	0.885	1
					ccaaacagcaccttaCACG	287	305	296	-	0.847	1
					cctataggaaaatgcCACC	344	362	353	-	0.811	0.75
	P$MYBS	MYB proteins with single DNA binding repeat	P$TAMYB80.01	MYB protein from wheat	ggtaATATacctcattc	325	341	333	-	0.836	1
					aggtATATtacctaggt	330	346	338	+	0.837	1
			P$PHR1.01	Phosphate starvation response 1	aggaAAATgccacctag	341	357	349	-	0.848	0.75
			P$HVMCB1.01	Hordeum vulgare Myb-related CAB-promoter-binding protein 1	atggctATCCtatagga	354	370	362	-	0.94	1
	P$MYCL	Myc-like basic helix-loop-helix binding factors	P$ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	ttaaaACAAttgcttaaat	130	148	139	-	0.979	0.955
					accttACACgtgtcagtca	278	296	287	-	0.956	0.955
					ggtggACATttgccagtca	305	323	314	+	0.959	1
					gagccACACttggtttata	457	475	466	-	0.967	0.955
			P$PIL5.01	Phytochrome interacting factor3-like 5	gactgaCACGtgtaaggtg	279	297	288	+	0.957	1
	P$GBOX	Plant G-box/C-box bZIP									
			P$CPRF.01	Common plant regulatory factor (CPRF)	caccttacACGTgtcagtcat	277	297	287	-	0.986	1

GmDREB2.1 (continued)				from parsley							
			P$GBF1.01	bZIP protein G-Box binding factor 1	tgactgacACGTgtaaggtgc	278	298	288	+	0.976	1
		proteins	P$ROM.01	Regulator of MAT (ROM1, ROM2)	caacttCCACctcaacatgag	434	454	444	+	0.863	1
GmDREB2.2	P$ABRE	ABA response elements	P$ABRE.01	ABA response elements	gattgacACGTgtaagg	289	305	297	+	0.912	1
					accttacACGTgtcaat	290	306	298	-	0.95	1
			P$ABF1.03	ABA (abscisic acid) inducible transcriptional activator	gaagttgaCGCGgcaag	400	416	408	+	0.857	0.75
	P$CCAF	Circadian control factors	P$EE.01	Evening element	agtttaAATAtgtat	111	125	118	-	0.865	1
					ataataAATAtatat	233	247	240	+	0.875	1
	P$ERSE	ER stress-response elements	P$ERSE_I.01	ERSE I (ER stress-response element I)-like motif	ccagcacatgattgaCACG	280	298	289	+	0.812	1
					cctataggaaaaggcCACC	353	371	362	-	0.808	0.75
	P$GBOX	Plant G-box/C-box bZIP proteins	P$CPRF.01	Common plant regulatory factor (CPRF) from parsley	taccttacACGTgtcaatcat	287	307	297	-	0.986	1
			P$GBF1.01	bZIP protein G-Box binding factor 1	tgattgacACGTgtaaggtaa	288	308	298	+	0.976	1
			P$BZIP910.01	bZIP transcription	tgaagtTGACgcggcaagtga	399	419	409	+	0.857	1

GmDREB2.2 (continued)				factor from Antirrhinum majus	tgttgaTGAGgtggaagttgc	442	462	452	-	0.812	0.75
	P$MIIG	MYB IIG-type binding sites	P$PALBOXP.01	Putative cis-acting element in various PAL and 4CL gene promoters	tgGGGGtggggtata	584	598	591	-	0.821	1
	P$MYBL	MYB-like proteins	P$GAMYB.01	GA-regulated myb gene from barley	ttaatgttGTTAtttca	165	181	173	+	0.925	1
					ccgtgatgGTTAtccta	368	384	376	-	0.931	1
			P$MYBPH3.01	Myb-like protein of Petunia hybrida	cccaaacaGTTAcctta	301	317	309	-	0.93	1
	P$MYBS	MYB proteins with single DNA binding repeat	P$ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	atatatataTATTtatt	235	251	243	-	0.807	0.778
					atatatataTATAttta	237	253	245	-	0.794	0.778
					aaatatataTATAtata	238	254	246	+	0.807	0.778
					atatatataTATAtatt	239	255	247	-	0.807	0.778
					atatatataTATAtata	240	256	248	+	0.807	0.778
					atatatataTATAtata	241	257	249	-	0.807	0.778
					atatatataTATAtata	242	258	250	+	0.807	0.778
					atatatataTATAtata	243	259	251	-	0.807	0.778
					atatatataTATAtata	244	260	252	+	0.807	0.778

GmDREB2.2 (continued)					atatatataTATAtata	245	261	253	-	0.807	0.778
					atatatataTATAtata	246	262	254	+	0.807	0.778
					atatatataTATAtata	247	263	255	-	0.807	0.778
					atatatataTATAtatt	248	264	256	+	0.807	0.778
					taatatataTATAtata	249	265	257	-	0.807	0.778
					atatatataTATAttaa	250	266	258	+	0.794	0.778
			P$HVMCB1.01	Hordeum vulgare Myb-related CAB-promoter-binding protein 1	atggttATCCtatagga	363	379	371	-	0.946	1
	P$MYCL	Myc-like basic helix-loop-helix binding factors	P$MYCRS.01	Myc recognition sequences	accttacACGTgtcaatca	288	306	297	-	0.978	1
			P$PIL5.01	Phytochrome interacting factor3-like 5	gattgaCACGtgtaaggta	289	307	298	+	0.957	1
			P$ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	ggtggACATttggcagtca	316	334	325	+	0.953	1
GmDREB3.1	P$ABRE	ABA response elements	P$ABF1.03	ABA (abscisic acid) inducible transcriptional activator	agtgggcaCGTGtcttc	273	289	281	+	0.884	1
			P$ABF1.01		agaagACACgtgcccac	274	290	282	-	0.905	1
					gaaggACACgcgccgct	302	318	310	-	0.837	1

GmDREB3.1 (continued)	P$CE3S	Coupling element 3 sequence	P$CE3.01	Coupling element 3 (CE3), non-ACGT ABRE	agaagaCACGtgcccactc	272	290	281	-	0.794	0.75
					agtgggCACGtgtcttctg	273	291	282	+	0.784	0.75
					cagcggCGCGtgtccttct	301	319	310	+	0.911	1
	P$DREB	Dehydration responsive element binding factors	P$CRT_DRE.01	C-repeat/dehydration response element	aaacaCCGAcgggag	260	274	267	+	0.948	1
	P$ERSE	ER stress-response elements	P$ERSE_I.01	ERSE I (ER stress-response element I)-like motif	ccaatgagagaaggaCACG	309	327	318	-	0.912	1
	P$GAPB	GAP-Box (light response elements)	P$GAP.01	Cis-element in the GAPDH promoters conferring light inducibility	aaaaATGAagacaaa	87	101	94	+	0.967	1
					gaaaATGAagagaag	554	568	561	-	0.951	1
	P$GBOX	Plant G-box/C-box bZIP proteins	P$CPRF.01	Common plant regulatory factor (CPRF) from parsley	cagaagacACGTgcccactcc	271	291	281	-	0.971	1
					gagtgggcACGTgtcttctgc	272	292	282	+	0.975	1
			P$TGA1.01	Arabidopsis leucine zipper protein TGA1	ttgtcaTGACgcaagacagct	349	369	359	-	0.922	1
			P$HBP1B.01	Wheat bZIP transcription factor HBP1B (histone gene binding protein	gctgtcttGCGTcatgacaag	350	370	360	+	0.833	0.795

GmDREB3.1 (continued)				1b)							
	P$HEAT	Heat shock factors	P$HSE.01	Heat shock element	acaagaaatcaAGAAtc	366	382	374	+	0.821	1
					gagaagagaagAGAAgt	538	554	546	-	0.85	1
					gagaagagaagAGAAga	543	559	551	-	0.85	1
					atgaagagaagAGAAga	548	564	556	-	0.839	1
					tagaagattagAGAAaa	585	601	593	-	0.91	1
	P$MIIG	MYB IIG-type binding sites	P$PALBOXL.01	Cis-acting element conserved in various PAL and 4CL promoters	gtgtgtttGGTGaga	523	537	530	-	0.894	1
	P$MYBL	MYB-like proteins	P$MYBPH3.01	Myb-like protein of Petunia hybrida	gaaaaaacGTTAgcttg	171	187	179	+	0.824	1
					acgggtTAGTttacgct	383	399	391	-	0.816	1
					ttagcgcgGTTAattgg	434	450	442	-	0.808	1
	P$MYBS	MYB proteins with single DNA binding repeat	P$HVMCB1.01	Hordeum vulgare Myb-related CAB-promoter-binding protein 1	aaacttATCCactaacc	99	115	107	+	0.958	1
	P$MYCL	Myc-like basic helix-loop-helix binding	P$MYCRS.01	Myc recognition sequences	agaagacACGTgcccactc	272	290	281	-	0.985	1
			P$PIL5.01	Phytochrome interacting factor3-like 5	agtgggCACGtgtcttctg	273	291	282	+	0.952	1

		factors									
GmDREB3.2	P$ABRE	ABA response elements	P$ABF1.03	ABA (abscisic acid) inducible transcriptional activator	agtgggcaCGTGtcttc	270	286	278	+	0.884	1
			P$ABF1.01		agaagACACgtgcccac	271	287	279	-	0.905	1
GmDREB3.2 (continued)	P$CCAF	Circadian control factors	P$EE.01	Evening element	tttaatAATAtatat	26	40	33	+	0.858	1
	P$CE3S	Coupling element 3 sequence	P$CE3.01	Coupling element 3 (CE3), non-ACGT ABRE	agaagaCACGtgcccactc	269	287	278	-	0.794	0.75
					agtgggCACGtgtcttctg	270	288	279	+	0.784	0.75
					agcagcGGCGtgtcctcct	300	318	309	+	0.82	0.75
	P$ERSE	ER stress-response elements	P$ERSE_I.01	ERSE I (ER stress-response element I)-like motif	ccaatgggaggaggaCACG	308	326	317	-	0.894	1
	P$GBOX	Plant G-box/C-box bZIP proteins	P$CPRF.01	Common plant regulatory factor (CPRF) from parsley	cagaagacACGTgcccactcc	268	288	278	-	0.971	1
					gagtgggcACGTgtcttctgc	269	289	279	+	0.975	1
			P$TGA1.01	Arabidopsis leucine zipper protein TGA1	tcgtgaTGACgcaagacagct	348	368	358	-	0.923	1
	P$HEAT	Heat shock factors	P$HSE.01	Heat shock element	gagaagtgagaAGAAga	581	597	589	-	0.813	1
	P$MYBL	MYB-like proteins	P$MYBPH3.01	Myb-like protein of Petunia hybrida	aaaaaaagGTTAgcttg	168	184	176	+	0.846	1
					aggggtTAGTtttcgct	374	390	382	-	0.819	1
			P$AS1_AS2_I.	AS1/AS2	ttagcgCTGTtaatcgg	432	448	440	-	1	1

			01	repressor complex binding motif I							
GmDREB3.2 (continued)	P$MYCL	Myc-like basic helix-loop-helix binding factors	P$MYCRS.01	Myc recognition sequences	agaagacACGTgcccactc	269	287	278	-	0.985	1
			P$PIL5.01	Phytochrome interacting factor3-like 5	agtgggCACGtgtcttctg	270	288	279	+	0.952	1
GmGOLS	P$ABRE	ABA response elements	P$ABF1.01	ABA (abscisic acid) inducible transcriptional activator	tgggtCCAAgtggcgag	90	106	98	+	0.869	0.75
					cgtgactgCGTGgcatg	138	154	146	+	0.849	1
					tttggACACgtggcaaa	281	297	289	+	1	1
					cgtgcctaCGTGacatt	558	574	566	+	0.898	1
			P$ABRE.01	ABA response elements	gtttgccACGTgtccaa	282	298	290	-	0.964	1
	P$CCAF	Circadian control factors	P$EE.01	Evening element	catggcAATAtctaa	166	180	173	-	0.899	1
	P$CE3S	Coupling element 3 sequence	P$CE3.01	Coupling element 3 (CE3), non-ACGT ABRE	ggtcacCGAGtgtcagcat	120	138	129	-	0.778	0.75
					cgtgacTGCGtggcatgtg	138	156	147	+	0.785	0.75
					gtttgcCACGtgtccaaat	280	298	289	-	0.784	0.75
					tttggaCACGtggcaaaca	281	299	290	+	0.79	0.75
	P$EREF	Ethylen respone element factors	P$ANT.01	ANT (Arabidopsis protein AINTEGUMENTA	actacggtTCCCgcggtca	212	230	221	+	0.827	1

GmGOLS (continued)				), member of the plant-specific family of AP2/EREBP-transcription factors							
	P$GBOX	Plant G-box/C-box bZIP proteins	P$HBP1A.01	HBP-1a, suggested to be involved in the cell cycle-dependent expression	tgtttgcCACGtgtccaaatt	279	299	289	-	0.983	1
			P$EMBP1.01	bZIP transcription factor implicated in ABA induced gene expression	atttggacACGTggcaaacat	280	300	290	+	1	1
			P$ROM.01	Regulator of MAT (ROM1, ROM2)	ccatccCCACctcatttatac	360	380	370	+	0.868	1
			P$OSBZ8.01	Oryza sativa bZIP protein 8	gtgggtccAAGTggcgagacc	89	109	99	+	0.849	0.75
					gcgtgcctACGTgacattcct	557	577	567	+	0.899	1
	P$HEAT	Heat shock factors	P$HSFA1A.01	Arabidopsis thaliana class A heat shock factor 1a	ttgaacTTTCgcacaat	38	54	46	-	0.759	1
				Arabidopsis thaliana class A heat shock factor 1a	gcgaaaGTTCaagagtt	43	59	51	+	0.84	0.857

GmGOLS (continued)			P$HSFA1A.01	Arabidopsis thaliana class A heat shock factor 1a	gggaagGTTCtggaaaa	234	250	242	-	0.909	0.857
			P$HSE.01	Heat shock element	tagaagagactAGAAta	186	202	194	-	0.914	1
					gagagtaatgcAGAAgt	420	436	428	-	0.842	1
	P$LREM	Light responsive element motif, not modulated by different light qualities	P$RAP22.01	RAP2.2, involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	atATCTaagtc	163	173	168	-	0.936	1
	P$MIIG	MYB IIG-type binding sites		Cis-acting element conserved in various PAL and 4CL promoters	aaagtgttGGTGagt	455	469	462	-	0.837	1
			P$PALBOXL.01	Putative cis-acting element in various PAL and 4CL gene promoters	gtGGGGatggtgtga	355	369	362	-	0.824	1
			P$ATMYB84.01	R2R3-type myb-like transcription factor (IIG-type binding site)	agggggtgGGTGtgt	381	395	388	-	0.924	1

GmGOLS (continued)			P$ACTYP.01	AC-type motifs, MYB46/MYB83-responsive elements	gttttGTTGgtggag	402	416	409	-	1	1
	P$MYBL	MYB-like proteins	P$CARE.01	CAACTC regulatory elements, GA-inducible	gttcaagAGTTgtttga	49	65	57	+	0.845	1
			P$MYBPH3.01	Myb-like protein of Petunia hybrida	cggaaacgGTGActacg	581	597	589	+	0.821	0.75
	P$MYCL	Myc-like basic helix-loop-helix binding factors	P$OSIRO2.01	Rice iron-related transcription factor 2	tctcgcCACTtggacccac	89	107	98	-	0.851	0.75
			P$MYCRS.01	Myc recognition sequences	gtttgccACGTgtccaaat	280	298	289	-	0.985	1
					tttggacACGTggcaaaca	281	299	290	+	0.973	1
			P$ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	tggttACATttgatcatgc	323	341	332	-	0.988	1
	P$SALT	Salt/drought responsive elements	P$ALFIN1.02	Zinc-finger protein in alfalfa roots, regulates salt tolerance	taaatgaGGTGggga	363	377	370	-	0.981	1
					gggagggGGTGggtg	384	398	391	-	0.967	1
	P$TCPF	DNA-binding proteins with the plant specific TCP-	P$PCF2.01	TCP class I transcription factor	cttggaCCCActt	87	99	93	-	0.992	1

		domain									
GmPP2C	P$CCAF	Circadian control factors	P$EE.01	Evening element	atattaAATAtttat	115	129	122	+	0.865	1
					aaaataAATAtttaa	118	132	125	-	0.872	1
					tttcaaAATAtttga	148	162	155	+	0.895	1
					tattcaAATAttttg	151	165	158	-	0.861	1
			P$CCA1.01	Circadian clock associated 1	agaaacaaAATCaat	186	200	193	+	0.851	1
					tataacacAATCttg	213	227	220	+	0.852	1
	P$IBOX	Plant I-Box sites	P$IBOX.01	I-Box in rbcS genes and other light regulated genes	aatatAATAagattaat	89	105	97	-	0.823	0.75
					aaaaaGAAAaggagaaa	130	146	138	-	0.813	0.75
					caaaaGAAAaggctcaa	254	270	262	+	0.831	0.75
	P$WBXF	W Box family	P$ERE.01	Elicitor response element	caatctTGACcctttat	220	236	228	+	0.899	1
					ccaatcTGACctagaac	435	451	443	+	0.921	1
			P$WRKY.01	WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box	cttctTTGAccacctta	317	333	325	-	0.977	1
					gaattTTGActgaagac	537	553	545	+	0.943	1
			P$WRKY11.01	Calmodulin binding WRKY	acttgTTGActttatcg	42	58	50	-	0.964	1

				transcription factor 11	taaaaTTGActgtgtga	338	354	346	+	0.94	1
GmPP2C (continued)	P$GBOX	Plant G-box/C-box bZIP proteins	P$BZIP911.02	bZIP transcription factor from Antirrhinum majus	gaggggTGAGgtgtttataaa	470	490	480	-	0.76	0.75
GmRAB18	P$ABRE	ABA response elements	P$ABRE.01	ABA response elements	tattttgACGTgagtat	27	43	35	-	0.828	1
					agtccacACGTgtcaag	307	323	315	+	0.926	1
					gcttgacACGTgtggac	308	324	316	-	0.891	1
					aaccaacACGTgtccat	365	381	373	+	0.885	1
					acgcagcACGTgtcaga	409	425	417	+	0.87	1
			P$ABF1.01	ABA (abscisic acid) inducible transcriptional activator	catggACACgtgttggt	366	382	374	-	0.865	1
					tccatgcaCGTGgcagc	377	393	385	+	0.869	1
					gtctgACACgtgctgcg	410	426	418	-	0.829	1
	P$CCAF	Circadian control factors	P$EE.01	Evening element	cgtcaaAATAttttt	34	48	41	+	0.892	1
					atcaaaAATAttttg	37	51	44	-	0.888	1
			P$CCA1.01	Circadian clock associated 1	agcataaaAATCtgg	443	457	450	-	0.93	1
	P$CE1F	Coupling element 1 binding factors	P$SBOX.01	Sugar and ABA responsive element conserved in	aattCACAtccat	252	264	258	+	0.917	0.783

GmRAB18 (continued)				several rbcS promoters							
	P$CE3S	Coupling element 3 sequence	P$CE3.01	Coupling element 3 (CE3), non-ACGT ABRE	agtccaCACGtgtcaagcc	307	325	316	+	0.776	0.75
					aaccaaCACGtgtccatgc	365	383	374	+	0.771	0.75
					acgtgcTGCGtggcgttga	401	419	410	-	0.795	0.75
	P$DREB	Dehydration responsive element binding factors	P$CRT_DRE.01	C-repeat/dehydration response element	tgttaCCGActctgg	230	244	237	-	0.903	1
	P$GBOX	Plant G-box/C-box bZIP proteins	P$CPRF.01	Common plant regulatory factor (CPRF) from parsley	gagtccacACGTgtcaagcca	306	326	316	+	0.978	1
					tagctgccACGTgcatggaca	375	395	385	-	0.974	1
					gtccatgcACGTggcagctac	376	396	386	+	0.993	1
					agtctgacACGTgctgcgtgg	407	427	417	-	0.964	1
					cacgcagcACGTgtcagactt	408	428	418	+	0.973	1
			P$GBF1.01	bZIP protein G-Box binding factor 1	ggcttgacACGTgtggactct	305	325	315	-	0.976	1
					gcatggacACGTgttggtttc	363	383	373	-	0.976	1
					aaaccaacACGTgtccatgca	364	384	374	+	0.977	1
			P$HBP1B.01	Wheat bZIP transcription	ttatactcACGTcaaaatatt	25	45	35	+	0.906	1

GmRAB18 (continued)				factor HBP1B (histone gene binding protein 1b)						
			P$OSBZ8.01	Oryza sativa bZIP protein 8	cacgtgctGCGTggcgttgag	400	420	410	-	0.848	0.75
			P$TGA1.01	Arabidopsis leucine zipper protein TGA1	atatttTGACgtgagtataac	24	44	34	-	0.964	1
	P$MIIG	MYB IIG-type binding sites	P$ACTYP.01	AC-type motifs, MYB46/MYB83-responsive elements	tagttGTTGttggtt	549	563	556	-	0.945	1
	P$MYBL	MYB-like proteins	P$GAMYB.01	GA-regulated myb gene from barley	taagcggtGTTAccgac	235	251	243	-	0.912	1
			P$CARE.01	CAACTC regulatory elements, GA-inducible	ttgccatAGTTgttgtt	553	569	561	-	0.841	1
	P$MYBS	MYB proteins with single DNA binding repeat	P$ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	ttgaatcttTATCaaaa	45	61	53	-	0.811	1
			P$MYBST1.01	MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat	acacttATCCaaatttt	284	300	292	+	0.955	1
					gactctATCCaaaattt	294	310	302	-	0.958	1
			P$OSMYBS.01	Rice MYB proteins with single DNA binding	acacgTGTCcatgcacg	370	386	378	+	0.859	0.75

GmRAB18 (continued)				domains, binding to the amylase element (TATCCA)							
	P$MYCL	Myc-like basic helix-loop-helix binding factors	P$ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	caaatACAAatgatcaaaa	125	143	134	+	0.973	0.955
					catggACACgtgttggttt	364	382	373	-	0.967	0.955
					gtctgACACgtgctgcgtg	408	426	417	-	0.961	0.955
			P$MYCRS.01	Myc recognition sequences	gcttgacACGTgtggactc	306	324	315	-	0.98	1
					agtccacACGTgtcaagcc	307	325	316	+	0.978	1
					aaccaacACGTgtccatgc	365	383	374	+	0.973	1
			P$OSIRO2.01	Rice iron-related transcription factor 2	tccatgCACGtggcagcta	377	395	386	+	0.95	1
			P$PIL5.01	Phytochrome interacting factor3-like 5	tatactCACGtcaaaatat	26	44	35	+	0.914	1
					agctgcCACGtgcatggac	376	394	385	-	0.957	1
					acgcagCACGtgtcagact	409	427	418	+	0.95	1
	P$WBXF	W Box family	P$WRKY.01	WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box	atattTTGAcgtgagta	28	44	36	-	0.926	1
					acaatTTGAcaaagata	194	210	202	-	0.921	1
			P$WRKY11.01	Calmodulin	actccTTGAccagattt	434	450	442	+	0.95	1

GmRAB18 (continued)				binding WRKY transcription factor 11						7	
	P$WNAC	Wheat NAC-domain transcription factors	P$TANAC69.01	Wheat NAC-domain DNA binding factor	gacgtgagtatAACGtttacaaa	15	37	26	-	0.709	0.812
					aacgttatactCACGtcaaaata	21	43	32	+	0.777	0.896
GmRemorin	P$CCAF	Circadian control factors	P$EE.01	Evening element	agaaatAATAtataa	281	295	288	-	0.863	1
					tcagaaAATAttttg	295	309	302	+	0.894	1
					aagcaaAATAttttc	298	312	305	-	0.892	1
					gttaaaAATAttttt	408	422	415	+	0.896	1
					tttaaaAATAttttt	411	425	418	-	0.896	1
	P$ERSE	ER stress-response elements	P$ERSE_I.01	ERSE I (ER stress-response element I)-like motif	ccaagaatagaaataCACG	198	216	207	+	0.839	1
	P$HEAT	Heat shock factors	P$HSE.01	Heat shock element	aggagacattaAGAAga	159	175	167	+	0.834	1
			P$HSFA1A.01	Arabidopsis thaliana class A heat shock factor 1a	ttgaccTTTCgtgtatt	209	225	217	-	0.767	1
	P$LREM	Light responsive element motif, not modulate	P$RAP22.01	RAP2.2, involved in carotenoid and tocopherol biosynthesis and	gcATCTaagcc	375	385	380	-	0.897	1

GmRemorin (continued)		d by different light qualities		in the expression of photosynthesis-related genes							
	P$MIIG	MYB IIG-type binding sites	P$PALBOXP.01	Putative cis-acting element in various PAL and 4CL gene promoters	gaGGGGtgagtgtag	551	565	558	-	0.827	1
			P$P_ACT.01	Maize activator P of flavonoid biosynthetic genes	ttttGGTTggttttt	643	657	650	-	0.974	0.967
	P$MYBL	MYB-like proteins	P$GAMYB.01	GA-regulated myb gene from barley	acttgttgGTTAgtctg	14	30	22	+	0.916	1
			P$AS1_AS2_I.01	AS1/AS2 repressor complex binding motif I	gttagtCTGTtattata	22	38	30	+	1	1
			P$CARE.01	CAACTC regulatory elements, GA-inducible	ttaaatgAGTTgctata	125	141	133	+	0.877	1
			P$MYBPH3.02	Myb-like protein of Petunia hybrida	ctttgtTTGTtgagtcc	226	242	234	-	0.766	0.817
					gaaaaactTTTAttccc	242	258	250	+	0.813	0.75
					taggttTAGTtactcat	315	331	323	-	0.775	1
					aaaaaactCTTAaaagc	462	478	470	+	0.817	0.75
			P$NTMYBAS1	Anther-specific	gtgaggttGTTAttaat	577	593	585	-	0.97	1

GmRemorin (continued)			.01	myb gene from tobacco						
	P$MYBS	MYB proteins with single DNA binding repeat	P$TAMYB80.01	MYB protein from wheat	ttatATATtccttcatt	354	370	362	+	0.847	1
			P$HVMCB1.01	Hordeum vulgare Myb-related CAB-promoter-binding protein 1	aaactcATCCactgcat	382	398	390	-	0.941	1
			P$OSMYBS.01	Rice MYB proteins with single DNA binding domains, binding to the amylase element (TATCCA)	ctttgTATCcaaactca	392	408	400	-	0.834	1
			P$ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	ccccatctcTATCatgt	527	543	535	+	0.834	1
	P$MYCL	Myc-like basic helix-loop-helix binding factors	P$ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	accacACACatgatagaga	532	550	541	-	0.961	0.955
GmSNK2	P$ERSE	ER stress-response elements	P$ERSE_I.01	ERSE I (ER stress-response element I)-like motif	ccaaatccagtggtcCACG	288	306	297	+	0.883	1

GmSNK2 (continued)	P$GBOX	Plant G-box/C-box bZIP proteins	P$GBF1.01	bZIP protein G-Box binding factor 1	ttggattgACGTggtccggcg	305	325	315	-	0.987	1
			P$UPRE.01	UPRE (unfolded protein response element) like motif	gccggaCCACgtcaatccaaa	306	326	316	+	0.929	1
	P$HEAT	Heat shock factors	P$HSFA1A.01	Arabidopsis thaliana class A heat shock factor 1a	atcaccCTTCatgaatc	7	23	15	+	0.776	1
			P$HSE.01	Heat shock element	agaaagagaccAGAAtg	258	274	266	-	0.821	1
					ccgaaattcccAGAAcc	544	560	552	-	0.86	1
	P$IBOX	Plant I-Box sites	P$IBOX.01	I-Box in rbcS genes and other light regulated genes	aaaaaAATAagacgaaa	428	444	436	-	0.814	0.75
	P$LREM	Light responsive element motif, not modulated by different light qualities	P$RAP22.01	RAP2.2, involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	gcATCTagaga	103	113	108	+	0.878	1
	P$MIIG	MYB IIG-type binding sites	P$NTLIM1.01	LIM domain protein binding to a PAL-box like sequence	tgagtttttgGTGGt	212	226	219	-	0.883	1

GmSNK2 (continued)			P$P_ACT.01	Maize activator P of flavonoid biosynthetic genes	ttttGGTTggtgacc	227	241	234	+	0.974	0.967
			P$MYBC1.01	Maize C1 myb-domain protein	agttgGTGGgtaaag	273	287	280	-	0.927	0.825
			P$ACTYP.01	AC-type motifs, MYB46/MYB83-responsive elements	ttggaGTTGgtgggt	277	291	284	-	1	1
	P$MYBL	MYB-like proteins	P$CARE.01	CAACTC regulatory elements, GA-inducible	tctgatgAGTTggatat	188	204	196	-	0.844	1
					gatttggAGTTggtggg	278	294	286	-	0.854	1
			P$MYBPH3.02	Myb-like protein of Petunia hybrida	ttcagcTTGTtgtttgg	390	406	398	+	0.767	0.817
					aaaaaaccGTTTttgca	411	427	419	-	0.837	0.75
					caaaaacgGTTTttttt	413	429	421	+	0.841	0.75
	P$MYBS	MYB proteins with single DNA binding repeat	P$HVMCB1.01	Hordeum vulgare Myb-related CAB-promoter-binding protein 1	gcacatATCCaactcat	184	200	192	+	0.952	1
	P$MYCL	Myc-like basic helix-loop-helix binding factors	P$PIL5.01	Phytochrome interacting factor3-like 5	ccggacCACGtcaatccaa	307	325	316	+	0.929	1
	P$TCPF	DNA-	P$ATTCP20.0	TCP class I	cctgGCCCaacac	476	488	482	-	0.95	1

GmSNK2 (continued)		binding proteins with the plant specific TCP-domain	1	transcription factor (Arabidopsis)							

[a] The element sequence is shown in “sequence”, whereby the capital letters represent the element core.

[bc] The “start” and “end” positions correspond to the position of the cis-element on the promoter sequence.

[d] The “anchor” position corresponds to the position of the element core.

[e] The “strand” corresponds to the DNA strand position.

[f] The “matrix similarity score” reflects the similarity between our sequence and the matrix sequence in which a value of 1 corresponds to sequences with the most conserved nucleotides at each position of the matrix. A matrix similarity >0.80 was used as the threshold.

APÊNDICE F

Artigo 1: Tabelas suplementares S3

Tables S3. ANOVA and Duncan's multiple range test for multiple comparisons among groups (5%) to evaluate the effects of water regimes (control, moderate and severe drought), time point (ZT0 to ZT20), and their interactions. Data regards qPCR analyses.

ANOVA: Drought stress genes						
Gene name	**Source**	**DF**	**SS**	**MS**	**F Value**	**Pr > F**
GmDREB1	treatment (water regime)	3	316.4295848	105.4765283	485.12	<.0001
	time	5	21.4888925	4.2977785	19.77	<.0001
	treatment*time	15	12.6147813	0.8409854	3.87	<.0001
	Error	116	25.2210223	0.2174226		
GmDREB2	treatment (water regime)	3	138.933642	46.311214	525.94	<.0001
	time	5	73.0266675	14.6053335	165.87	<.0001
	treatment*time	15	35.9684162	2.3978944	27.23	<.0001
	Error	116	10.2142857	0.0880542		
GmDREB3	treatment (water regime)	3	105.4570119	35.1523373	598.74	<.0001
	time	5	1.7477826	0.3495565	5.95	<.0001
	treatment*time	15	4.7380871	0.3158725	5.38	<.0001
	Error	116	6.810431	0.0587106		
GmbZIP	treatment (water regime)	3	73.13410869	24.37803623	112.27	<.0001
	time	5	83.32723866	16.66544773	76.75	<.0001
	treatment*time	15	89.16320765	5.94421384	27.37	<.0001
	Error	120	26.0575351	0.2171461		
GmABAR	treatment (water regime)	3	220.1605372	73.3868457	435.14	<.0001
	time	5	77.7442034	15.5488407	92.19	<.0001
	treatment*time	15	59.4573497	3.9638233	23.5	<.0001
	Error	119	20.0696904	0.1686529		
GmSnRK2	treatment (water regime)	3	125.23588	41.7452933	886.83	<.0001
	time	5	8.7116676	1.7423335	37.01	<.0001
	treatment*time	15	4.8845583	0.3256372	6.92	<.0001
	Error	107	5.0367645	0.0470726		
GmPP2C	treatment (water regime)	3	230.8127704	76.9375901	701.76	<.0001
	time	5	28.0913581	5.6182716	51.25	<.0001
	treatment*time	15	35.3695907	2.3579727	21.51	<.0001
	Error	113	12.388803	0.1096354		
GmRemorin	treatment (water regime)	3	775.4713674	258.4904558	1667.87	<.0001
	time	5	256.3941726	51.2788345	330.87	<.0001
	treatment*time	15	108.6796645	7.245311	46.75	<.0001
	Error	109	16.893087	0.154982		
GmGOLS	treatment (water regime)	3	792.9746649	264.3248883	3099.29	<.0001
	time	5	228.1929698	45.638594	535.13	<.0001
	treatment*time	15	25.3261683	1.6884112	19.8	<.0001

	Error	115	9.80785	0.085286		
GmRAB18	treatment (water regime)	3	5780.716389	1926.905463	5581.15	<.0001
	time	5	88.484436	17.696887	51.26	<.0001
	treatment*time	15	136.839222	9.122615	26.42	<.0001
	Error	117	40.39451	0.345252		

DF= Degress of Freedom; SS= Sum of Squares; MS= Mean square.

ANOVA: Circadian clocck genes

Gene name	Source	DF	SS	MS	F Value	Pr > F
GmTOC1	treatment (water regime)	3	29.5613073	9.8537691	24.26	<.0001
	time	5	120.1261935	24.0252387	59.15	<.0001
	treatment*time	15	27.0897981	1.8059865	4.45	<.0001
	Error	114	46.305408	0.4061878		
GmLUX	treatment (water regime)	3	40.7961043	13.5987014	86.74	<.0001
	time	5	716.1387663	143.2277533	913.62	<.0001
	treatment*time	15	54.7615116	3.6507674	23.29	<.0001
	Error	115	18.0284146	0.1567688		
GmPRR3	treatment (water regime)	3	97.4270071	32.475669	37.94	<.0001
	time	5	769.919129	153.9838258	179.89	<.0001
	treatment*time	15	42.5107255	2.8340484	3.31	0.0001
	Error	112	95.868713	0.855971		
GmPRR7	treatment (water regime)	3	261.466259	87.15542	706.32	<.0001
	time	5	3131.677092	626.335418	5075.91	<.0001
	treatment*time	15	780.599414	52.039961	421.74	<.0001
	Error	70	8.637555	0.123394		
GmPRR9	treatment (water regime)	3	69.617398	23.205799	59.18	<.0001
	time	5	1261.263761	252.252752	643.25	<.0001
	treatment*time	15	121.011074	8.067405	20.57	<.0001
	Error	111	43.529225	0.392155		
GmLCL1	treatment (water regime)	3	40.07466	13.35822	82.97	<.0001
	time	5	651.231215	130.246243	808.94	<.0001
	treatment*time	15	45.9773297	3.0651553	19.04	<.0001
	Error	120	19.3209402	0.1610078		
GmELF4	treatment (water regime)	3	44.8938398	14.9646133	124.82	<.0001
	time	5	742.5187438	148.5037488	1238.72	<.0001
	treatment*time	15	108.1830699	7.2122047	60.16	<.0001
	Error	114	13.6668825	0.1198849		
GmGI	treatment (water regime)	3	119.8279571	39.9426524	489.52	<.0001
	time	5	580.0271203	116.0054241	1421.7	<.0001
	treatment*time	15	36.440092	2.4293395	29.77	<.0001
	Error	109	8.893977	0.0815961		
GmZTL	dhidrico	3	28.22403406	9.40801135	118.72	<.0001
	pdia	5	14.00932881	2.80186576	35.36	<.0001
	dhidrico*pdia	15	5.39131956	0.3594213	4.54	<.0001
	Error	111	8.79651256	0.07924786		
GmCHE	treatment (water regime)	3	62.10921942	20.70307314	172.24	<.0001
	time	5	22.03684983	4.40736997	36.67	<.0001
	treatment*time	15	7.22357035	0.48157136	4.01	<.0001
	Error	118	14.1834486	0.1201987		
GmJumonji	treatment (water regime)	3	74.3930447	24.7976816	206.49	<.0001
	time	5	814.6386124	162.9277225	1356.73	<.0001
	treatment*time	15	126.5932572	8.4395505	70.28	<.0001
	Error	113	13.570019	0.120089		

DF= Degress of Freedom; SS= Sum of Squares; MS= Mean square.

		Duncan's test: Circadian clock genes																					
		GmTOC1		GmLUX		GmPRR3		GmPRR7		GmPRR9		GmLCL1		GmELF4		GmGI		GmZTL		GmCHE		GmJumonji	
Time point	**Treatment**	**ΔCt Mean**	**Duncan 5%**	**ΔCt Mean**	**Duncan 5%**	**ΔCt Mean**	**Duncan 5%**	**ΔCt Mean**	**Duncan 5%**	**ΔCt Mean**	**Duncan 5%**	**ΔCt Mean**	**Duncan 5%**	**ΔCt Mean**	**Duncan 5%**	**ΔCt Mean**	**Duncan 5%**	**ΔCt Mean**	**Duncan 5%**	**ΔCt Mean**	**Duncan 5%**	**ΔCt Mean**	**Duncan 5%**
ZT0	**1-Control-Moderate stress**	2.027	b	9.798	a	6.373	a	3.651	a	10.513	a	-2.112	b	7.900	a	3.959	a	2.202	a	2.079	b	10.284	a
	2-Control-Severe stress	3.532	a	8.831	b c	3.265	c	1.542	c	8.117	b	-3.630	d	6.373	b	1.563	b	0.708	c	0.917	c	10.443	a
	3-Drought-Moderate Stress	2.808	a	9.241	b	4.688	b	2.537	b	10.327	a	-2.857	c	7.622	a	3.969	a	2.341	a	0.665	c	9.285	b
	4-Drought-Severe stress	1.478	b	8.755	c	3.561	b c	0.089	d	6.936	c	-1.297	a	6.307	b	0.598	c	1.701	b	2.674	a	7.290	c
ZT4	**1-Control-Moderate stress**	3.946	a	9.422	a	1.947	a	0.692	a	6.129	a	-1.738	a	7.585	a	0.233	a	2.634	a	1.069	b	10.316	a
	2-Control-Severe stress	4.086	a	8.019	c	-0.432	c	-0.418	bc	2.836	c	-3.092	c	6.797	b	-1.565	d	1.236	c	0.560	c	9.957	a
	3-Drought-Moderate Stress	3.723	a	8.642	b	0.508	b c	-0.151	b	5.004	b	-2.298	b	6.248	c	-0.656	b	1.793	b	0.471	c	9.489	b
	4-Drought-Severe stress	1.538	b	5.268	d	1.354	a b	-0.771	c	2.569	c	-1.362	a	4.773	d	-1.230	c	1.991	b	1.718	a	4.234	c
ZT8	**1-Control-Moderate stress**	1.945	a	4.906	a	1.344	a	-0.967	a	-0.273	b	1.537	a	4.067	a	-1.401	a	2.550	a	0.518	b	3.268	a
	2-Control-Severe stress	1.924	a	3.794	b	-0.614	b	-1.826	b	-1.812	c	-0.702	c	2.464	c	-2.835	c	1.546	c	-0.451	c	2.699	b c
	3-Drought-Moderate Stress	2.212	a	4.606	a	0.912	a	-0.956	a	0.339	a b	1.741	a	3.986	a	-1.504	a	2.444	a	0.296	b	3.086	a b
	4-Drought-Severe stress	0.622	b	2.668	c	0.450	a	-0.896	a	0.577	a	0.469	b	3.489	b	-2.312	b	2.060	b	1.256	a	2.452	c
ZT12	**1-Control-Moderate stress**	0.868	a	2.883	b	3.183	a	0.116	a	1.998	b	3.853	a	-0.087	b	0.184	a	2.356	a b	0.697	b	3.654	a b
	2-Control-Severe stress	-0.308	b	1.985	c	1.138	b	-0.339	a	1.200	c	3.073	b	-1.694	c	-1.329	d	1.453	c	-0.714	d	3.176	c
	3-Drought-Moderate Stress	0.812	a	2.866	b	2.763	a	-0.269	a	1.668	b c	3.928	a	-0.320	b	-0.152	b	2.198	b	0.046	c	3.325	b c
	4-Drought-Severe stress	0.413	a b	3.532	a	2.102	a b	0.026	a	3.713	a	1.553	c	3.058	a	-0.517	c	2.540	a	1.678	a	3.941	a
ZT16	**1-Control-Moderate stress**	1.536	a	5.372	a	5.728	a	2.801	a	4.511	a	2.764	a	2.533	b	2.770	a	1.734	a b	1.147	b	5.626	b
	2-Control-Severe stress	0.685	b	4.634	b	4.168	b	1.372	b	4.665	a	1.628	c	1.492	c	0.389	c	0.651	c	0.337	c	7.277	a
	3-Drought-	1.204	a b	5.168	a	5.462	a	2.607	a	4.664	a	2.784	a	1.873	c	1.965	b	1.529	b	0.426	c	5.606	b

	Moderate Stress																						
	4-Drought- Severe stress	0.566	b	5.127	a	3.655	b	1.323	b	4.602	a	2.108	b	2.944	a	-0.063	d	1.933	a	1.855	a	5.685	b
ZT20	**1-Control- Moderate stress**	1.837	a	8.222	a	8.732	a	20.840	a	8.456	b	-0.704	a	4.817	c	6.093	a	1.407	a	0.782	b c	8.092	a
	2-Control- Severe stress	1.644	a b	7.452	b	6.974	b	18.669	b	6.764	c	-1.929	b	2.905	d	3.485	c	0.446	b	0.412	c	8.292	a
	3-Drought- Moderate Stress	2.124	a	8.473	a	8.860	a	20.684	a	9.907	a	-0.881	a	5.273	b	5.464	b	1.709	a	0.899	b	8.129	a
	4-Drought- Severe stress	0.845	b	7.040	b	5.401	c	1.910	c	6.181	c	-0.444	a	5.692	a	2.027	d	1.550	a	2.203	a	7.220	b

Means followed by different letters are significantly different at the 5% level by Duncan's test

		Duncan's test: Drought-responsive genes																			
		GmDREB1		GmDREB2		GmDREB3		GmbZIP		GmABAR		GmSnRK		GmPP2C			Gm Remorin	GmGOLS		GmRAB18	
Time point	Treatment	ΔCt Mean	Duncan 5%	ΔCt Mean	Duncan 5%	ΔCt Mean	Duncan 5%	ΔCt Mean	Duncan 5%	ΔCt Mean	uncan 5%	ΔCt Mean	ncan 5%	ΔCt Mean	Duncan 5%	ΔCt Mean	Duncan 5%	ΔCt Mean	Duncan 5%	ΔCt Mean	Duncan 5%
ZT0	1-Control-Moderate stress	4.819	a	2.063	a	1.987	a	0.992	b	-2.350	a	0.271	c	2.741	a	1.703	a	3.575	a	5.974	a
	2-Control-Severe stress	1.525	c	-0.560	b	-0.111	c	-0.785	c	-2.517	a	1.634	b	1.910	b	-1.104	b	3.199	b	4.878	b
	3-Drought-Moderate Stress	2.789	b	1.788	a	1.610	b	-1.305	c	-2.419	a	-0.145	d	2.863	a	1.487	a	1.735	c	3.199	c
	4-Drought-Severe stress	0.158	d	2.044	a	-0.805	d	2.354	a	2.235	a	2.964	a	-0.853	c	-1.268	b	-3.353	d	-9.385	d
ZT4	1-Control-Moderate stress	3.975	a	2.083	b	1.867	a	1.924	b	-2.215	b	1.191	c	4.864	a	3.463	a	5.428	a	6.429	a
	2-Control-Severe stress	1.374	c	0.412	d	0.742	c	-0.770	c	-2.223	b	2.220	b	3.058	b	-0.535	c	4.159	b	4.113	b
	3-Drought-Moderate Stress	2.660	b	1.091	c	1.155	b	-0.814	c	-2.476	b	1.179	c	3.342	b	1.089	b	2.754	c	0.231	c
	4-Drought-Severe stress	0.454	d	2.988	a	-0.211	d	2.694	a	1.910	a	3.470	a	0.044	c	-1.630	d	-0.212	d	10.751	d
ZT8	1-Control-Moderate stress	4.613	a	3.857	a	1.773	a	2.428	a	-2.136	c	1.155	c	3.102	a	6.017	a	5.729	a	6.757	a
	2-Control-Severe stress	1.575	b	1.235	c	0.828	b	-0.049	c	-3.563	d	2.567	b	3.407	a	2.685	c	5.049	c	3.776	b
	3-Drought-Moderate Stress	4.625	a	3.920	a	1.630	a	2.417	a	-1.482	b	0.809	d	2.091	b	5.408	b	5.583	b	6.299	a
	4-Drought-Severe stress	1.149	b	2.202	b	-0.398	c	1.681	b	-0.576	a	3.198	a	1.141	c	-1.227	d	0.178	d	-9.971	c
ZT12	1-Control-Moderate stress	5.031	a	3.733	a	2.023	a	3.288	a	-0.242	b	0.617	c	3.435	b	7.174	a	6.176	a	7.102	a

	2-Control-Severe stress	2.259	b	1.125	c	0.609	c	1.336	b	-1.114	c	2.326	b	4.100	a	2.829	c	5.828	b	4.811	c
	3-Drought-Moderate Stress	4.582	a	3.802	a	1.556	b	2.833	a	-0.024	b	0.553	c	2.586	c	6.275	b	5.452	c	6.283	b
	4-Drought-Severe stress	1.161	c	2.268	b	-0.160	d	1.733	b	1.218	a	2.983	a	0.881	d	-1.005	d	0.102	d	10.330	d
ZT16	**1-Control-Moderate stress**	4.164	a	3.325	a	1.922	a	3.389	a	-0.122	b	0.972	b	3.699	a b	5.965	a	6.916	a	6.644	a
	2-Control-Severe stress	1.593	c	0.278	d	0.413	b	2.034	bc	-0.136	b	2.496	a	3.331	b	0.014	b	6.864	a	5.164	b
	3-Drought-Moderate Stress	3.525	b	2.845	b	1.646	a	2.531	b	-0.272	b	0.971	b	3.805	a	6.220	a	6.505	b	5.506	b
	4-Drought-Severe stress	0.955	d	1.253	c	-0.155	c	1.727	c	1.145	a	2.754	a	0.522	c	-1.183	c	1.369	c	-7.460	c
ZT20	**1-Control-Moderate stress**	4.131	a	1.722	a	1.742	a	2.250	a	-1.450	b	0.660	c	3.087	a	3.208	a	6.412	a	6.022	a
	2-Control-Severe stress	1.207	b	-1.063	c	0.189	c	1.360	b	-1.195	b	1.962	b	2.115	b	-0.744	c	5.720	b	4.898	b
	3-Drought-Moderate Stress	3.864	a	1.514	a	1.421	b	1.434	b	-1.446	b	0.871	c	3.103	a	2.306	b	5.721	b	5.271	b
	4-Drought-Severe stress	0.658	c	0.923	b	-0.223	d	1.498	b	1.929	a	3.230	a	-0.468	c	-1.574	d	0.301	c	-8.487	c

Means followed by different letters are significantly different at the 5% level by Duncan's test

APÊNDICE G

Artigo 1: Tabelas suplementares S4

Tables S4: ANOVA and Tukey's HSD multiple comparison test (95% family-wise confidence level) to evaluate the effects of water regimes (control and moderate drought), time point (ZT0 to ZT20), and their interactions. p-values are shown; bold numbers represent significant gene expression differences (p-values ≤0.05). Data regards RNA-seq analyses.

ANOVA: Drought-responsive genes

Gene name	Source	DF	SS	MS	F Value	p-value
DREB1	Treatment (water regime)	1	1.1338	1.1338	9.836	**0.004480**
	Time	5	8.6980	1.7396	15.091	**0.000001**
	Treatment*Time	5	1.5119	0.3024	2.623	**0.049829**
	Error	24	2.7665	0.1153	-	-
DREB2	Treatment (water regime)	1	0.25749	0.257489	1.24120	0.276273
	Time	5	25.35463	5.070925	24.44394	**0.000000**
	Treatment*Time	5	1.81482	0.362964	1.74964	0.161691
	Error	24	4.97883	0.207451	-	-
DREB3	Treatment (water regime)	1	0.0358	0.0358	0.150	0.701661
	Time	5	8.2288	1.6458	6.902	**0.000403**
	Treatment*Time	5	0.5922	0.1184	0.497	0.775550
	Error	24	5.7229	0.2385	-	-
bZIP	Treatment (water regime)	1	8.9153	8.9153	40.445	**0.000001**
	Time	5	49.3781	9.8756	44.801	**0.000000**
	Treatment*Time	5	3.5888	0.7178	3.256	**0.021985**
	Error	24	5.2904	0.2204	-	-
GmABAR	Treatment (water regime)	1	0.026	0.026	0.18	0.674808
	Time	5	26.967	5.393	37.48	**0.000000**
	Treatment*Time	5	0.278	0.056	0.39	0.853476
	Error	24	3.454	0.144	-	-
GmSnRK2	Treatment (water regime)	1	0.8213	0.8213	6.393	**0.018445**
	Time	5	12.5584	2.5117	19.552	**0.000000**
	Treatment*Time	5	1.4576	0.2915	2.269	0.079898
	Error	24	3.0832	0.1285	-	-
GmPP2C	Treatment (water regime)	1	0.4712	0.4712	1.5599	0.223718
	Time	5	21.6466	4.3293	14.3329	**0.000002**
	Treatment*Time	5	3.6821	0.7364	2.4381	0.063718

	Error	24	7.2493	0.3021	-	-
GmRemorin	Treatment (water regime)	1	6.9229	6.92286	11.53214	**0.002381**
	Time	5	175.9056	35.18112	58.60490	**0.000000**
	Treatment*Time	5	2.6103	0.52205	0.86964	**0.515550**
	Error	24	14.4074	0.60031	-	-
GmGOLS	Treatment (water regime)	1	2.0779	2.0779	2.0323	0.166864
	Time	5	151.7961	30.3592	29.6935	**0.000000**
	Treatment*Time	5	16.6771	3.3354	3.2623	**0.021815**
	Error	24	24.5381	1.0224	-	-
GmRAB18	Treatment (water regime)	1	25.2066	25.2066	8.3749	**0.007973**
	Time	5	50.5254	10.1051	3.3574	**0.019353**
	Treatment*Time	5	50.4734	10.0947	3.3540	**0.019437**
	Error	24	72.2349	3.0098	-	-

DF= Degress of Freedom; SS= Sum of Squares; MS= Mean square; Bold numbers represent significant gene expression diferences (p-values ≤0.05).

Tukey's test:Drought-responsive genes			**Control**					**Drought**					
Gene Name	**Treament**	**Time point**	**ZT4**	**ZT8**	**ZT12**	**ZT16**	**ZT20**	**ZT0**	**ZT4**	**ZT8**	**ZT12**	**ZT16**	**ZT20**
GmDREB1	**Control**	**ZT0**	0.6360885	0.9785726	0.9999160	0.3108399	0.6280532	**0.0237752**	**0.0081460**	0.7436986	1.0000000	0.1579283	0.1559209
		ZT4	-	0.0885357	0.2638741	0.9999842	1.0000000	0.7733227	0.4909388	**0.0210955**	0.6894870	0.9975290	0.9973719
		ZT8	-	-	0.9999725	**0.0267348**	0.0862528	**0.0012459**	**0.0004079**	0.9999126	0.9656492	**0.0108225**	**0.0106503**
		ZT12	-	-	-	0.0943775	0.2583350	**0.0051194**	**0.0016866**	0.9782379	0.9997259	**0.0411635**	**0.0405504**
		ZT16	-	-	-	-	0.9999875	0.9713346	0.8223973	**0.0057569**	0.3545651	0.9999996	0.9999995
		ZT20	-	-	-	-	-	0.7803236	0.4987943	**0.0204879**	0.6816659	0.9978021	0.9976594
	Drought	**ZT0**	-	-	-	-	-	-	0.9999976	**0.0002530**	**0.0288808**	0.9983817	0.9984864
		ZT4	-	-	-	-	-	-	-	**0.0000839**	**0.0099809**	0.9591593	0.9605106
		ZT8	-	-	-	-	-	-	-	-	0.6925645	**0.0022431**	**0.0022063**
		ZT12	-	-	-	-	-	-	-	-	-	0.1850508	0.1827717
		ZT16	-	-	-	-	-	-	-	-	-	-	1.0000000

			Control					**Drought**					
Gene Name	**Treament**	**Time point**	**ZT4**	**ZT8**	**ZT12**	**ZT16**	**ZT20**	**ZT0**	**ZT4**	**ZT8**	**ZT12**	**ZT16**	**ZT20**
GmbZIP	**Control**	**ZT0**	0.2162191	**0.0000480**	**0.0000429**	**0.0001941**	**0.0061350**	**0.0252946**	0.5171790	**0.0001517**	**0.0039759**	**0.0018528**	0.9691368
		ZT4	-	**0.0418701**	**0.0376835**	0.1405699	0.8681209	**0.0000282**	**0.0014080**	0.1152010	0.7796460	0.5885179	0.9108064
		ZT8	-	-	1.0000000	0.9999721	0.6737627	**0.0000000**	**0.0000003**	0.9999960	0.7799473	0.9186281	**0.0011086**
		ZT12	-	-	-	0.9999394	0.6439923	**0.0000000**	**0.0000003**	0.9999896	0.7532710	0.9024840	**0.0009871**
		ZT16	-	-	-	-	0.9438364	**0.0000000**	**0.0000011**	1.0000000	0.9776245	0.9978218	**0.0045707**
		ZT20	-	-	-	-	-	**0.0000007**	**0.0000268**	0.9141921	1.0000000	0.9999947	0.1129139
GmbZIP (continued)	**Drought**	**ZT0**	-	-	-	-	-	-	0.8777738	**0.0000000**	**0.0000005**	**0.0000002**	**0.0011355**
		ZT4	-	-	-	-	-	-	-	**0.0000009**	**0.0000176**	**0.0000085**	0.0518877
		ZT8	-	-	-	-	-	-	-	-	0.9612235	0.9947538	**0.0035733**
		ZT12	-	-	-	-	-	-	-	-	-	1.0000000	0.0782153
		ZT16	-	-	-	-	-	-	-	-	-	-	**0.0397318**

			Control					**Drought**					
Gene Name	**Treament**	**Time point**	**ZT4**	**ZT8**	**ZT12**	**ZT16**	**ZT20**	**ZT0**	**ZT4**	**ZT8**	**ZT12**	**ZT16**	**ZT20**
GmRemorin	**Control**	**ZT0**	0.0730174	**0.0000009**	**0.0000000**	**0.0002916**	0.4577632	0.9824991	0.9999501	**0.0000054**	**0.0000018**	**0.0003912**	0.9943219
		ZT4	-	**0.0024363**	**0.0000434**	0.4631682	0.9941473	**0.004662**	0.2397694	**0.0174746**	**0.0055216**	0.5381852	0.455433
		ZT8	-	-	0.8595361	0.3599141	**0.0001925**	**0.0000001**	**0.0000034**	0.9992241	0.9999999	0.2985111	**0.0000087**
		ZT12	-	-	-	**0.0123109**	**0.0000039**	**0**	**0.0000001**	0.3763844	0.6748167	**0.009249**	**0.0000002**
		ZT16	-	-	-	-	0.0744017	**0.0000167**	**0.0013149**	0.8457742	0.5616606	1	**0.0036565**
		ZT20	-	-	-	-	-	0.0521422	0.8287686	**0.0014342**	**0.0004376**	0.095535	0.9673353
	Drought	**ZT0**	-	-	-	-	-	-	0.7831408	**0.0000004**	**0.0000001**	**0.0000221**	0.5253134
		ZT4	-	-	-	-	-	-	-	**0.0000223**	**0.0000072**	**0.0017660**	0.9999990
		ZT8	-	-	-	-	-	-	-	-	0.9999955	0.7845854	**0.0000601**
		ZT12	-	-	-	-	-	-	-	-	-	0.4857831	**0.0000191**
		ZT16	-	-	-	-	-	-	-	-	-	-	**0.0048972**

			Control					**Drought**					
Gene Name	**Treament**	**Time point**	**ZT4**	**ZT8**	**ZT12**	**ZT16**	**ZT20**	**ZT0**	**ZT4**	**ZT8**	**ZT12**	**ZT16**	**ZT20**
GmGOLS	**Control**	**ZT0**	0.1702691	**0.0031803**	**0.0003697**	**0.0026109**	**0.0011172**	0.7917922	1.0000000	**0.0014205**	**0.0029694**	**0.0000115**	**0.0035033**
		ZT4	-	0.8010852	0.2855287	0.7564840	0.5406631	**0.0030481**	0.1842964	0.6032421	0.7859800	**0.0140362**	0.8215496
		ZT8	-	-	0.9985675	1.0000000	0.9999988	**0.0000406**	**0.0035297**	0.9999999	1.0000000	0.4608370	1.0000000
		ZT12	-	-	-	0.9993997	0.9999978	**0.0000052**	**0.0004104**	0.9999838	0.9989277	0.9335924	0.9978912
		ZT16	-	-	-	-	0.9999999	**0.0000335**	**0.0028984**	1.0000000	1.0000000	0.5108190	1.0000000
		ZT20	-	-	-	-	-	**0.0000148**	**0.0012408**	1.0000000	0.9999994	0.7288781	0.9999969
	Drought	**ZT0**	-	-	-	-	-	-	0.7682134	**0.0000187**	**0.0000380**	**0.0000002**	**0.0000447**
		ZT4	-	-	-	-	-	-	-	**0.0015776**	**0.0032959**	**0.0000127**	**0.0038876**
		ZT8	-	-	-	-	-	-	-	-	1.0000000	0.6689447	0.9999997

		ZT12	-	-	-	-	-	-	-	-	-	0.4780439	1.0000000
		ZT16	-	-	-	-	-	-	-	-	-	-	0.4369779

			Control					**Drought**					
Gene Name	**Treament**	**Time point**	**ZT4**	**ZT8**	**ZT12**	**ZT16**	**ZT20**	**ZT0**	**ZT4**	**ZT8**	**ZT12**	**ZT16**	**ZT20**
GmRAB18	**Control**	**ZT0**	0.9999928	0.9977356	0.9979096	0.9985419	0.9986787	0.2300327	0.1157156	0.9999988	1.0000000	0.9999967	0.9770359
		ZT4	-	0.9999990	0.9999992	0.9501918	0.9524239	0.0869879	**0.0394763**	1.0000000	0.9999524	1.0000000	0.8161457
		ZT8	-	-	1.0000000	0.7843285	0.7896304	**0.0351322**	**0.0151102**	0.9999939	0.9944291	0.9999976	0.5646436
		ZT12	-	-	-	0.7889657	0.7942186	**0.0358106**	**0.0154153**	0.9999948	0.9947948	0.9999980	0.5700901
		ZT16	-	-	-	-	1.0000000	0.7299718	0.4993572	0.9665671	0.9995367	0.9582738	0.9999998
		ZT20	-	-	-	-	-	0.7241623	0.4932880	0.9682465	0.9995880	0.9602442	0.9999997
GmRAB18	**Drought**	**ZT0**	-	-	-	-	-	-	0.9999998	0.1025773	0.2748055	0.0938597	0.9046430
		ZT4	-	-	-	-	-	-	-	**0.0471727**	0.1421049	**0.0428492**	0.7251505
		ZT8	-	-	-	-	-	-	-	-	0.9999884	1.0000000	0.8548918
		ZT12	-	-	-	-	-	-	-	-	-	0.9999743	0.9880930
		ZT16	-	-	-	-	-	-	-	-	-	-	0.8344845

APÊNDICE H

Artigo 1: Tabelas suplementares S5

Tables S5. ANOVA and Duncan's multiple range test for multiple comparisons among groups (5%) to evaluate the effects of ABA treatment, time point (ZT4 to ZT24), and their interactions. Data regards qPCR analyses.

ANOVA: Circadian clocck genes						
Gene name	**Source**	**DF**	**SS**	**MS**	**F Value**	**p-value**
GmLUX	treatment (ABA)	1	0.747408	0.747408	8.2515	**0.00575**
	time	5	380.0231	76.00461	839.0983	**0.00001**
	treatment*time	5	1.056848	0.21137	2.3335	0.05234
	Error	60	5.434735	0.090579		
GmPRR3	treatment (ABA)	1	4.154668	4.154668	35.53	**0.00001**
	time	5	751.2648	150.253	1284.939	**0.00001**
	treatment*time	5	0.781684	0.156337	1.337	0.26033
	Error	60	7.016035	0.116934		
GmTOC1	treatment (ABA)	1	0.060523	0.060523	0.91	3.44E-01
	time	5	92.77595	18.55519	278.95	**<.0001**
	treatment*time	5	3.259519	0.651904	9.8	**<.0001**
	Error	58	3.858005	0.066517		
GmJumonji	treatment (ABA)	1	2.158273	2.158273	21.12	**<.0001**
	time	5	548.3737	109.6747	1073.09	**<.0001**
	treatment*time	5	7.474696	1.494939	14.63	**<.0001**
	Error	49	5.008029	0.102205		

ANOVA: Drought stress genes						
Gene name	**Source**	**DF**	**SS**	**MS**	**F Value**	**p-value**
GmDREB2	treatment (ABA)	1	1.653977	1.653977	21.82	**<.0001**
	time	5	64.87703	12.97541	171.19	**<.0001**
	treatment*time	5	3.057985	0.611597	8.07	**<.0001**
	Error	58	4.396042	0.075794		
GmbZIP	treatment (ABA)	1	9.940223	9.940223	187.29	**<.0001**
	time	5	63.42713	12.68543	239.02	**<.0001**
	treatment*time	5	1.923447	0.384689	7.25	**<.0001**
	Error	58	3.078219	0.053073		
GmABAR	treatment (ABA)	1	5.851341	5.851341	25.0126	**0.00005**
	time	5	231.3073	46.26146	197.7526	**0.00001**

GmABAR (continued)	treatment*time	5	0.817695	0.163539	0.6991	0.62856
	Error	60	14.03617	0.233936		
GmGOLS	treatment (ABA)	1	3.145292	3.145292	10.0085	**0.00279**
	time	5	54.56201	10.9124	34.7238	**0.00001**
	treatment*time	5	2.727234	0.545447	1.7356	0.13954
	Error	60	18.85578	0.314263		
GmRAB18	treatment (ABA)	1	123.1396	123.1396	498.33	<.0001
	time	5	99.38767	19.87753	80.44	<.0001
	treatment*time	5	36.44005	7.288011	29.49	<.0001
	Error	58	14.33213	0.247106		

Duncan's test		Circadian clock genes				Drought responsive genes					
		GmTOC1		GmJumonji		GmbZIP		GmDREB2		GmRAB18	
Time point	Treatment	ΔCt Mean	Duncan 5%	ΔCt Mean	Duncan 5%	ΔCt Mean	Duncan 5%	ΔCt Mean	Duncan 5%	ΔCt Mean	Duncan 5%
ZT4	**1-Control**	4.1282856	a	10.9112947	a	0.905001	a	0.9266677	a	3.364952	a
	2-ABA treatment	3.5081351	b	9.2648018	b	-0.1488664	b	0.2411336	b	2.161469	b
ZT8	**1-Control**	2.4251075	a	4.0936412	a	1.761054	a	2.0627207	a	5.421774	a
	2-ABA treatment	2.5684219	a	3.2167552	b	0.7981355	b	1.1581355	b	-0.14491	b
ZT12	**1-Control**	0.2988695	a	3.1284378	a	2.0230619	a	1.9680619	a	5.805536	a
	2-ABA treatment	-0.3826196	b	2.8943663	a	1.7247087	b	1.648042	b	3.615578	b
ZT16	**1-Control**	1.4734013	b	6.6674695	b	2.6464104	a	1.1867779	b	5.511735	a
	2-ABA treatment	1.984322	a	7.189322	a	2.0278461	b	1.6711795	a	2.760989	b
ZT20	**1-Control**	1.5975373	a	9.1909752	a	2.1283186	a	0.1566519	a	6.545871	a
	2-ABA treatment	1.8283366	a	9.4933366	a	1.7182006	b	-0.113466	a	5.05667	b
ZT24	**1-Control**	2.9083597	a	10.8178312	a	0.0846541	a	-0.7884558	a	6.610026	a
	2-ABA treatment	2.9698734	a	10.4282068	b	-1.1222165	b	-0.9488831	a	3.793207	b

Means followed by different letters are significantly different at the 5% level by Duncan's test