

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
CENTRO DE BIOTECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

BIOSSÍNTESE E SINALIZAÇÃO MEDIADA POR ETILENO EM SOJA: USO DE  
ABORDAGENS *IN SILICO* E IMPLICAÇÕES NA RESPOSTA À SECA

**DISSERTAÇÃO DE MESTRADO**

**FABRÍCIO BARBOSA MONTEIRO ARRAES**

**PORTO ALEGRE – 2014**

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
CENTRO DE BIOTECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

BIOSSÍNTESE E SINALIZAÇÃO MEDIADA POR ETILENO EM SOJA: USO DE  
ABORDAGENS *IN SILICO* E IMPLICAÇÕES NA RESPOSTA À SECA

**Fabício Barbosa Monteiro Arraes**

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

Orientadora: Dra. Maria Fátima Grossi de Sá

PORTO ALEGRE  
DEZEMBRO – 2014

“Todas as vitórias ocultam uma abdicação”.

Simone de Beauvoir

Este trabalho foi concluído no Laboratório de Interação Molecular Planta-Patógeno (EMBRAPA - Recursos Genéticos e Biotecnológica - CENARGEN), Laboratório de Biotecnologia Vegetal (EMBRAPA/SOJA) e no Laboratório de Química e Bioquímica de Alimentos na Faculdade de Ciências Farmacêuticas da Universidade de São Paulo. O projeto foi subsidiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pela Fundação de Amparo à Pesquisa do Distrito Federal (FAPDF).



## AGRADECIMENTOS

À minha orientadora a Dra. Maria Fátima Grossi de Sá, pela oportunidade, confiança no meu trabalho e pela compreensão no período final da Dissertação.

Aos órgãos de fomento CAPES, CNPq, EMBRAPA e FAPDF pelo apoio financeiro, estrutura e pela bolsa de estudos, essenciais para a conclusão deste projeto.

Ao PPGBCM, mais especificamente aos professores do programa que contribuíram tanto com a minha formação acadêmica.

Aos membros da minha Comissão de Acompanhamento e aos membros da Banca Examinadora pela contribuição com o trabalho.

Ao Dr. Eduardo Purgatto (USP), Dr. Alexandre Nepomuceno (EMBRAPA/Soja), Silvana Rockenbach (EMBRAPA/Soja) e Claudinei Toledo (EMBRAPA/Soja) que auxiliaram em experimentos que não tínhamos como realizar em nosso laboratório.

Aos amigos Dra. Magda Beneventi, Dra. Maria Eugênia Lisei de Sá e Joaquin Roca Paixão que tanto contribuíram para a conclusão do artigo científico originado a partir desta Dissertação.

À Sílvia Centeno e ao Luciano Saucedo pelo acolhimento no período do Mestrado. Sem a alegria e ajuda de vocês a conclusão do mesmo teria sido muito mais difícil.

Aos meus amados pais Maria de Fátima Barbosa Monteiro e Joel Arraes que sempre acreditaram e investiram em mim.

À minha amiga e irmã de consideração Raquel Sampaio que tanto me deu forças nos momentos mais difíceis, iluminando meu caminho com seu coração puro.

À minha “filha científica” Júlia Costa que com seu jeito único de ser me ensinou mais do que imagina.

Aos amigos Roberta Coelho, Djair Jr., Danila Cabral, Ana Gabriela Borges e Hudson Moura que tanto me deram força e energia positiva, principalmente na conclusão da Dissertação.

E aos demais amigos, colegas e familiares, incluindo Dr. Nivaldo Pereira Alves, Dra. Elisiane Lacerda e Rossy Melo que tanto me apoiaram durante este período.

Agradeço imensamente a todos!

## ÍNDICE

LISTA DE ABREVIATURAS E SIGLAS .....	008
RESUMO .....	013
ABSTRACT .....	014
1. INTRODUÇÃO .....	015
1.1. UMA BREVE INTRODUÇÃO SOBRE FITORMÔNIOS .....	015
1.2. ETILENO (Et) .....	017
1.2.1. Biossíntese .....	018
1.2.2. Sinalização .....	023
1.2.3. Transporte .....	032
1.2.4. Catabolismo .....	032
1.2.5. Etileno em Situações de Estresse .....	032
1.3. <i>CROSTALK</i> ENTRE ROTAS SINALIZADAS POR FITORMÔNIOS .....	034
1.4. ELEMENTOS REGULADORES EM PROMOTORES GÊNICOS .....	040
1.5. RESPOSTAS À DÉFICIT HÍDRICO EM PLANTAS .....	050
1.5.1. Fitormônios e Tolerância à Estresse Hídrico .....	051
1.6. A SOJA .....	054
1.6.1. Principais Características .....	054
1.6.2. Estabelecimento e Conjuntura Atual da Soja .....	055
1.6.3. Principais Pragas .....	059
1.6.4. Tolerância à Seca .....	062
1.6.5. Projeto GENOSOJA .....	064
2. JUSTIFICATIVA .....	066
3. OBJETIVOS .....	067
3.1. OBJETIVOS GERAIS .....	067
3.2. OBJETIVOS ESPECÍFICOS .....	067
4. MATERIAL E MÉTODOS .....	068
4.1. EXPERIMENTOS <i>IN SILICO</i> .....	068
4.1.1. Identificação dos Genes Alvo .....	068
4.1.2. Determinação de Domínios Proteicos .....	070
4.1.3. Alinhamentos, Ideogramas, Dendogramas e Mapas Metabólicos .....	070

4.1.4.	Determinação de Possíveis Sítios de Fosforilação .....	071
4.1.5.	Determinação de Motivos Grupo-Específicos Comuns .....	071
4.1.6.	Modelagem Molecular dos Domínios AP2 .....	072
4.1.7.	Experimentos de Melhor “Hit” Bidirecional (BBH) .....	073
4.1.8.	Identificação de Elementos <i>cis</i> -Atuantes em Possíveis Promotores .....	073
4.1.9.	Análise de Expressão em Transcriptomas de Soja .....	073
4.2.	EXPERIMENTOS <i>IN VITRO/IN VIVO</i> .....	075
4.2.1.	Crescimento das Plantas .....	075
4.2.2.	Determinação de Parâmetros Fisiológicos .....	079
4.2.3.	Determinação da Produção de Etileno .....	079
4.2.4.	Determinação de ACC Livre .....	080
4.2.5.	Extração e Quantificação de RNA Total .....	081
4.2.6.	Tratamento das Amostras de RNA Total com DNase I .....	081
4.2.7.	Síntese de DNA Complementar (cDNA) .....	082
4.2.8.	PCR Quantitativa em Tempo Real (qPCR) .....	083
5.	RESULTADOS .....	086
5.1.	IDENTIFICAÇÃO <i>IN SILICO</i> E ONTOLOGIA DOS GENES .....	086
5.2.	EXPERIMENTO DE BBH .....	089
5.3.	PREDIÇÃO DE MODELOS <i>IN SILICO</i> EM SOJA .....	091
5.3.1.	Biossíntese de Etileno .....	091
5.3.2.	Transdução de Sinal Mediada por Etileno .....	096
5.4.	FATORES DE TRANSCRIÇÃO DA SUPERFAMÍLIA AP2/ERF .....	102
5.5.	ANÁLISE E VALIDAÇÃO DE TRANSCRIPTOMAS DE SOJA .....	105
5.6.	ANÁLISE DE POSSÍVEIS PROMOTORES GÊNICOS DE SOJA .....	113
5.7.	NÍVEIS DE PRODUÇÃO DE ETILENO E DE ACC LIVRE EM SOJA .....	117
6.	DISCUSSÃO .....	120
7.	CONCLUSÕES FINAIS .....	132
8.	REFERÊNCIAS BIBLIOGRÁFICAS .....	134
APÊNDICES ( <i>CURRICULUM VITAE</i> ) <sup>1</sup>		

---

<sup>1</sup> O material contido neste item também encontra-se no CD em anexo para melhor visualização.

## LISTA DE ABREVIATURAS E SIGLAS

**°C:** Grau Célsius

**A:** Taxa fotossintética

**AAT:** Aminoácido transferase

**ABA:** Ácido abscísico

**ABRE:** Elementos de resposta à ABA

**ACC:** 1-aminociclopropano-1-ácido carboxílico

**ACD:** 1-aminociclopropano-1-ácido carboxílico deaminase

**ACO:** 1-aminociclopropano-1-ácido carboxílico oxidase

**ACS:** 1-aminociclopropano-1-ácido carboxílico sintase

**ACT:** Aciltransferase; N-malonil transferase

**AdoMet (SAM):** S-adenosil-metionina

**ANT:** Aintegumenta

**AOA:** Ácido aminoxacético

**AP2:** *APETALA 2*

**ARD:** Acidoreductona oxigenase

**AREF:** Elementos de resposta à auxina

**Asc:** Ascorbato

**ASP (AAT):** Aspartato aminotransferase

**ATAF:** Elementos de interação com proteínas NAC semelhante à ATAF

**ATP:** Adenosina trifosfato

**AVG:** Aminoetóxi-vinil-glicina

**BBH:** Melhor "hit" bidirecional

**BR:** Brassinoesteróides

**BRRE:** Elementos de resposta à brassinoesteróides

**CAAT:** Elementos *CCAAT-box*

**CDC5:** Elementos homólogos ao CDC5 de *Arabidopsis thaliana*

**cDNA:** DNA complementar

**CE1F:** Elementos de acoplamento 1

**cGMP:** Guanosina monofosfato cíclica

**ChiP:** Imunoprecipitação de cromatina

**C:** Concentração interna de CO<sub>2</sub>

**CK:** Citocinina

**cm:** Centímetro

**CNAC:** Elementos de interação com fatores NAC regulados por cálcio

**CONAB:** Companhia Nacional de Abastecimento

**CPK (CDPK):** Proteína cinase cálcio dependente

**CTR:** Tripla resposta constitutiva

**DEP:** Complexo desidratase-enolase-fosfatase

**DHAsc:** Diidroascorbato

**DHKMP:** 1,2-diidroxi-3-ceto-5-metil-tio-penteno

**DNA:** Ácido Desoxirribonucleico

**dNTP:** Desoxirribonucleotídeo trifosfato

**DPBF:** Elementos de interação em promotores *Dc3*

**DRE:** Elemento de Interação com fatores de transcrição DREB

**DREB:** Proteínas com elementos de interação responsivos à desidratação

**DTT:** Ditioneitol

**E:** Taxa de transpiração

**EAR:** Motivo anfílico de repressão associado à ERF

**EBF:** Proteína *F-box* que interage com EIN3

**EFE:** Enzima de formação de etileno

**EIL:** Semelhante à EIN3

**EIN:** Insensível à etileno

**EINL:** Elementos de interação com fatores semelhantes à EIN3

**EOL:** Semelhante à ETO

**ER:** Retículo endoplasmático

**EREF:** Elementos responsivos à etileno

**ERF:** Fator de transcrição responsivo à etileno

**ERS:** Sensor responsivo à etileno

**Et:** Etileno

**ETO:** Superprodução de etileno

**ETP:** Proteína alvo de EIN2

**ETR:** Receptores de etileno

**FLO2:** Elementos de interação com proteína homeótica floral *APETALA 2*

**FPKM:** Fragmentos por quilobase de transcrito por milhão de fragmentos mapeados

**g/mol:** gramas por mol (medida de quantidade de substância)

**g:** grama

**GA:** Giberelina

**GACC:** 1-glutamil-ACC

**GAF:** Fosfodiesterase dependente de cGMP, adenil ciclase e *Fh1A*

**GARP:** Elementos de interação com DNA relacionados à MYB - Golden2, ARR, Psr

**Gb:** Gigabase

**GBOX:** Elementos *G-box/C-box* de interação com bZIP de plantas

**GCCF:** Elementos *GCC-box*

**GENOSOJA:** Consórcio Brasileiro do Genoma da Soja

**GGT:**  $\gamma$ -glutamil transpeptidase

**gs:** Condutância estomática

**h:** Hora

**HCN:** Cianeto

**HEAT:** Elementos de choque térmico

**IAA:** Auxina

**IBGE:** Instituto Brasileiro de Geografia e Estatística

**ISR:** Resistência sistêmica induzida

**JA:** Ácido jasmônico

**JARE:** Elementos de resposta à jasmonato

**KMTB:** 2-ceto-4-metil-tio-butirato

**LREM:** Elementos responsivos à luz, não modulados por diferentes tipos de luz

**m:** Metro

**M:** Molar

**MACC:** 1-malonil-ACC

**MAT:** Metionina-adenosil transferase

**Mb:** Megabase

**MIIG:** Sítios de interação com MYB IIG

**MKK (MAPKK):** Proteína cinase cinase ativada por mitógeno

**MKKK (MAPKKK):** Proteína cinase cinase cinase ativada por mitógeno

**mL/min:** Mililitro de gás por minuto  
**mL:** Mililitro  
**mm:** Milímetro  
**mM:** Milimolar  
**MPK (MAPK):** Proteína cinase ativada por mitógeno  
**mRNA:** RNA mensageiro  
**MTA:** 5-metil-tio-adenosina  
**MTI:** MTR-1-P isomerase  
**MTK:** 5-metil-tio-ribose cinase  
**MTN:** 5-metil-tio-adenosina nucleosidase  
**MTR:** 5-metil-tio-ribose  
**MTR-1-P:** 5-metil-tio-ribose-1-fosfato  
**MTRu-1-P:** 5-metil-tio-ribulose-1-fosfato  
**MYBL:** Elementos de interação com proteínas semelhantes à MYB  
**MYBS:** MYB com elementos únicos de interação com DNA  
**MYCL:** Elementos de interação com proteínas semelhantes à MYC  
**NACF:** Elementos de interação com fatores NAC específicos de plantas  
**ng g<sup>-1</sup>:** Nanograma de ACC por grama de massa fresca  
**ng:** Nanograma  
**nL g<sup>-1</sup> h<sup>-1</sup>:** Nanolitro de etileno por grama de massa fresca por hora  
**NLS:** Sinal de localização nuclear  
**nm:** Nanômetro  
**NRQ:** Quantificações relativas normalizadas  
**ORF:** Fase aberta de leitura  
**OXB:** 2-oxobutanoato ou oxobutirato  
**pb:** Pares de base  
**PLP:** Piridoxal-5'-fosfato  
**PTBP:** *TATA-box* de plantas  
**PUFA:** Ácido graxo poliinsaturado  
**qPCR:** Reação em cadeia da polimerase em tempo real  
**RAN:** Responsivo à agonista  
**RAV:** Relacionado com ABI3/VP1

**RAV3:** Porção 3' do sítio de interação com RAV1

**RAV5:** Porção 5' do sítio de interação com RAV1

**RD:** Domínio de repressão

**RNA:** Ácido ribonucleico

**ROS (EROS):** Espécie reativa de oxigênio

**RTE:** Reversão da sensibilidade de etileno

**SA:** Ácido salicílico

**SALT:** Elementos responsivos à seca e variação de salinidade

**SAMS:** S-adenosil-metionina sintetase

**SAR:** Resistência sistêmica adquirida

**SNP:** Polimorfismo de base única

**SPAC:** Sistema solo-planta-atmosfera contínuo

**SWNS:** Elementos de resposta à NACS - parede secundária

**T<sub>ar</sub>:** Temperatura do ar

**T<sub>folha</sub>:** Temperatura da folha

**T<sub>m</sub>:** Temperatura de derretimento

**TSS:** Sítio de início da transcrição

**U/ $\mu$ L:** Unidade de atividade enzimática por microlitro

**WUE:** Eficiência do uso da água

$\Delta T$ : Variação de temperatura

$\mu\text{g}/\text{mL}$ : Micrograma por microlitro

$\mu\text{g}$ : Micrograma

$\mu\text{L}/\text{L}$ : Microlitro de etileno por litro de ar

$\mu\text{L}$ : Microlitro

$\mu\text{M}$ : Micromolar

$\mu\text{mol m}^2 \text{ s}^{-1}$ : Micromolar por metro quadrado por segundo



## RESUMO

O etileno (Et) é um fitormônio importante na chamada *resposta tripla* em plântulas, compreendendo a inibição do alongamento caulinar, o espessamento do caule e o hábito de crescimento horizontal (perda da sensibilidade gravitrópica), além de participar na indução da abscisão foliar e na resposta à estresse. Estudos sobre a resposta da soja aos mais diversos tipos de estresse têm emergido dentre a comunidade científica brasileira, uma vez que esta cultura é de grande importância econômica ao país. Tendo em vista o papel descrito das respostas desencadeadas por etileno em diversas espécies vegetais, este trabalho descreveu *in silico* a rota de biossíntese deste fitormônio e da sinalização por ele mediada em soja, utilizando recursos disponibilizados por banco de dados públicos e pelo Projeto GENOSOJA (*Consórcio Brasileiro do Genoma da Soja*), que disponibilizou dados moleculares relacionados à genômica, transcriptômica e proteômica desta leguminosa sob diversas condições de estresse. Com auxílio de ferramentas de bioinformática foram identificados 502 genes em soja, divididos em três grupos: (i) biossíntese de etileno, (ii) transdução de sinal por ele mediada e (iii) fatores de transcrição da superfamília AP2/ER. Análises posteriores possibilitaram a avaliação da possível ortologia das proteínas de soja identificadas por experimentos *in silico* de BBH (*Best Bidirectional Hit*), mostrando uma maior proximidade filogenética com *Arabidopsis thaliana*. A associação destes dados propiciou a elaboração de modelos *in silico* da biossíntese de etileno e da sinalização por ele mediada em soja, onde foram identificados elementos conservados principalmente com *A. thaliana*. Além disso, a análise do banco de dados fornecido pelo Projeto GENOSOJA possibilitou identificar em transcriptomas de raízes e folhas de cultivares de soja cultivadas em sistema hidropônico e submetidas à déficit hídrico a expressão diferencial de cerca de 24,00 % dos genes caracterizados. A regulação transcricional destes genes foi estudada pela análise de seus possíveis promotores, onde foram identificadas *in silico* 29 famílias de elementos *cis*-atuantes importantes na regulação mediada por etileno e em possíveis interconexões com a sinalização mediada por outros fitormônios, como ácido jasmônico, ácido abscísico, auxinas e brassinoesteróides. Estes dados, associados com a avaliação dos níveis de ACC livre e produção de etileno, possibilitaram melhor compreensão em nível molecular da importância deste fitormônio em soja nas condições analisadas e a sua possível participação em respostas desta *commodity* à seca.

## ABSTRACT

The ethylene (Et) is the phytohormone known for the seedlings triple response, comprising the inhibition of shoot elongation, the stem thickness and the habit of horizontal growth (loss of gravitropic sensitivity). Ethylene also participates in the induction of leaf abscission and in response to stress. Studies about soybean response to various types of stress have emerged among the Brazilian scientific community since this crop is extremely relevant to the country. Considering the described role triggered by ethylene responses in several plant species, this work described *in silico* the biosynthesis and signaling pathways mediated by this gas in soybean, using resources provided by public databases and GENOSOJA (*Brazilian Soybean Genome Consortium*) project. The study expanded the molecular data related to genomics, transcriptomics and proteomics known of this plant under several stress conditions. With the aid of bioinformatics tools, we identified 502 genes in soybean, divided into three groups: (i) ethylene biosynthesis, (ii) signal transduction mediated by this gas, and (iii) AP2/ERF transcriptional factors superfamily. Subsequent analyzes of BBH (*Best Bidirectional Hit*) experiments enabled the evaluation of the possible protein orthology encoded by these genes, showing phylogenetic proximity with *Arabidopsis thaliana*. The combination of these data propitiated the development of soybean ethylene biosynthesis and signaling *in silico* models, which were identified preserved elements mainly with *A. thaliana*. Furthermore, GENOSOJA database analysis allowed to identify the differential expression of 24, 00% of these genes in soybean transcriptomes, as the crop was initially grown in hydroponic system and then submitted to water deficit. In addition, the transcriptional regulation of these genes was studied by analyzing its putative promoters, in which were identified, *in silico*, 29 families of *cis*-acting elements. These elements are important for ethylene-mediated regulation and its possible interconnections with other signaling pathways mediated by other plant hormones, such as jasmonic acid, abscisic acid, auxin and brassinosteroids. Thereby, this study, combined with assessing the abundance of free ACC and ethylene production, allowed a better understanding of ethylene importance at a molecular level in this crop, and its possible participation in responses to drought.

## 1. INTRODUÇÃO.

### 1.1. UMA BREVE REVISÃO SOBRE FITORMÔNIOS.

Os hormônios em plantas possuem um papel crucial no controle de vias importantes no crescimento e desenvolvimento das mesmas. Enquanto o metabolismo fornece energia e elementos estruturais para a sobrevivência da planta, os hormônios integram a velocidade do crescimento de regiões específicas através de respostas moleculares. Adicionalmente, eles possuem papel no controle de processos reprodutivos (Gerashchenkov & Rozhnova, 2013) e em resposta a diversos tipos de estresse (Forcat et al., 2008; Bari & Jones, 2009; Kaya et al., 2009; Peleg & Blumwald, 2011; Erb et al., 2012; Kohli et al., 2013). Desta forma, pode-se definir como *fitormônio* (hormônios de plantas) o grupo de compostos orgânicos que ocorrem naturalmente e, mesmo em baixas concentrações, influenciam processos fisiológicos importantes (Gardner, 2009).

Desta forma, é necessário considerar como os sinais hormonais podem atuar eficientemente mesmo em baixas concentrações. Em qualquer sistema de sinalização existe um sinal que é captado por um alvo. Na sinalização hormonal em plantas é difícil identificar cada um destes componentes, mas algumas generalizações podem ser consideradas: (i) cada hormônio pode ser sintetizado em mais de um local na planta. De fato, todas as células vivas de um vegetal podem produzir todos os hormônios, mas somente algumas podem produzi-los em altas quantidades; (ii) cada hormônio pode apresentar muitas funções, fato este que é comprovado por deduções a partir de experimentos específicos com hormônios, fenótipos hormônio deficientes e mutantes insensíveis a hormônios; (iii) fitormônios são pequenas moléculas com grande mobilidade e difusão celular; (iv) vários tipos celulares respondem a cada classe hormonal; (v) muitos hormônios podem ter atividade nas mesmas células ou tecidos em que são sintetizados (Went & Thimann, 1937).

Charles Darwin foi um dos primeiros cientistas que investigou o papel dos fitormônios em plantas. Suas observações originais do fototropismo dos coleóptilos de grama o levaram a postular a existência de um sinal que era transportado da extremidade do coleóptilo até zonas mais basais da planta (Darwin & Darwin, 1881). Após diversos

autores caracterizarem o modo pelo qual este sinal era transferido, Went e colaboradores (1927) finalmente foram capazes de isolar a substância química por difusão em blocos de ágar, a partir de extremidades de coleótilos (Went, 1927). Quando estas substâncias foram administradas nas extremidades dos coleótilos decapitados, foi observada a estimulação do crescimento dos mesmos, seguida de eventual curvatura quando a substância era adicionada assimetricamente sobre estas extremidades. Esta substância identificada foi denominada originalmente por Went como *Wuchsstoff*, nome que foi mudado posteriormente para *auxina*. Após algumas falsas identificações, o material foi finalmente caracterizado como *ácido indol-3-acético*, conhecido mundialmente como IAA (ou AIA).

Outras linhas de investigação levaram à identificação de outros hormônios: pesquisas envolvendo patogênese em plantas levaram à descoberta das *giberelinas* (GAs) (MacMillan & Suter, 1958); as *citocininas* (CKs) foram detectadas em estudos de padronização metodológica de cultura de tecidos vegetais (Kieber & Schaller, 2010); o controle da abscisão foliar e dormência culminaram na descoberta do *ácido abscísico* (ABA) (Addicott & Lyon, 1969); e o *etileno* (Et) foi detectado pela primeira vez ao se avaliar os efeitos da fumaça e da iluminação à gás em plantas (Doubt, 1917).

Vale a pena ressaltar que além dos fitormônios presentes nas cinco principais classes citadas acima, diversas outras moléculas foram caracterizadas em plantas desempenhando funções hormonais. Dentre elas, destacam-se o *ácido jasmônico* (JA), o *ácido salicílico* (SA) e os *brassinosteróides* (BRs). O *ácido jasmônico* e seus principais metabólitos são importantes no crescimento (inibição) e desenvolvimento de plantas (senescência, formação de gavinhas e tubérculos, desenvolvimento de flores e abscisão foliar), assim como em resposta tanto à estresse biótico (principalmente ferimentos e resistência adquirida sistêmica) quanto abiótico (seca, alta salinidade e temperaturas adversas) (Delker et al., 2006; Worrall et al., 2012; Pirbalouti et al., 2014). Já o *ácido salicílico* é um fitormônio com papel importante na fotossíntese e transpiração (induzindo mudanças na anatomia foliar e estrutura do cloroplasto), transporte e captação de íons, assim como na defesa contra patógenos (sinalizando na produção de proteínas específicas) (Vlot et al., 2009). Por fim, os *brassinosteróides* promovem a expansão, alongamento e divisão celular (em parceria com as auxinas), regeneração da parede celular, diferenciação vascular e formação do tubo polínico. Além disso, aceleram a senescência em cultura de

tecidos e podem apresentar um papel protetor sob influência de frio e seca (Clouse & Sasse, 1998; Cano-Delgado et al., 2004; Nemhauser et al., 2004).

## 1.2. ETILENO (Et).

O etileno (nome IUPAC: eteno) é um hidrocarboneto gasoso incolor, inflamável, com odor doce e almiscarado (quando puro) e com fórmula molecular  $C_2H_4$  (Booth & Campbell, 1929). Este alceno (hidrocarboneto com ligação dupla entre carbonos) forma quatro ligações com átomos de hidrogênio com conformação coplanar, onde o ângulo de ligação H-C-H é de  $117,4^\circ$ , ideal para hibridização  $sp^2$  no carbono. A molécula é relativamente rígida devido à dupla ligação presente entre os carbonos. A necessidade de processos altamente energéticos para rompimento da ligação  $\pi$  é a responsável pela reatividade do composto, principalmente por ser uma região de alta densidade eletrônica e suscetível à ataques por eletrófilos (Booth & Campbell, 1929). Muitas reações de metais com o gás são catalisadas por transições metálicas, podendo ser transientes pelo uso de orbitais  $\pi$  e  $\pi^*$ . Além disso, por ser uma molécula simples, o método mais utilizado para sua detecção ainda é a espectroscopia normal no espectro UV-visível (Booth & Campbell, 1929).

Como fitormônio, estudos empíricos do efeito deste gás em plantas são estudados, mesmo que indiretamente, desde a Antiguidade. Os antigos egípcios e chineses já haviam observado que a fumaça de incenso acelerava a maturação de frutos. Uma prática antiga refere-se à utilização de fogueiras junto a plantações de manga (nas Filipinas) e de abacaxi (nas Antilhas), no sentido de ajudar a iniciar a sincronização da floração nestas culturas. Em 1864, Girardin observou que o gás de iluminação (produto da combustão de carvão) provocava o desfolhamento de árvores (Doubt, 1917). Em 1901, Neljubow foi o primeiro a constatar que o etileno era o componente do gás que afetava o crescimento das plantas (Neljubow, 1901). Observou ainda, que o etileno causava a resposta tripla em plântulas de ervilha, compreendendo a inibição do alongamento caulinar, o espessamento do caule e o hábito de crescimento horizontal (perda da sensibilidade gravitrópica). Em 1917 estudos mostraram a participação deste fitormônio na indução da abscisão foliar, mas somente em meados de 1934 que foi comprovada a sua biossíntese em plantas e a sua associação com o

amadurecimento dos frutos e senescência de uma variedade de tecidos vegetais (folhas, flores e células do xilema) (Doubt, 1917; Nath et al., 2006).

A partir daí, inúmeras funções fisiológicas foram atribuídas ao etileno (Davies, 1995). Dentre elas destacam-se a indução da germinação de sementes, crescimento de raízes adventícias sob condições de alagamento (aumentando a eficiência de absorção de água e minerais), estimulação da epinastia (curvatura descendente de folhas, resultado do crescimento assimétrico do pecíolo) e da curvatura nutricional, inibição do crescimento de brotos e do fechamento de estômatos (exceto em plantas aquáticas e em espécies que se desenvolvem normalmente em condições de alagamento, como arroz), além da indução da floração em *Ananas comosus* (L.) Merrill (abacaxi) (Trusov & Botella, 2006; Wilmowicz et al., 2008). Além disso, existem estudos que correlacionam a liberação de etileno em resposta a uma grande variedade de estresses (bióticos e abióticos) (Morgan & Drew, 1997).

### 1.2.1. Biossíntese.

Para entender o papel do etileno em plantas, é importante saber como o fitormônio gasoso é sintetizado. A produção de etileno varia de acordo com o tecido e com a espécie vegetal e sua biossíntese é realizada a partir da metionina, que é disponibilizada pela rota metabólica de reciclagem denominada de ciclo de Yang, ciclo da metionina ou ainda ciclo da 5-metil-tio-adenosina (MTA) (Miyazaki & Yang, 1987). Muitas das reações deste ciclo foram previstas pela identificação dos intermediários esperados, mas muitos dos passos enzimáticos foram deduzidos de dados obtidos em animais, leveduras e bactérias. Hoje é bem estabelecido o fato de que o ciclo de Yang em plantas não é essencial somente para a biossíntese de etileno, mas também para reações biossintéticas de poliaminas e nicotinaminas/fitosideróforos, sendo que todas essas rotas utilizam S-adenosil-metionina (SAM) e produzem MTA, que é reciclado (Kende, 1993; Roje, 2006).

Rzewuski e colaboradores caracterizaram a primeira enzima do ciclo, a *5-metil-tio-adenosina nucleosidase* (MTN; EC 3.2.2.16) que é a responsável pela clivagem da adenina, presente no MTA, produzindo 5-metil-tio-ribose (MTR) (Rzewuski et al., 2007). A próxima enzima do ciclo de Yang encontrada em plantas, a *5-metil-tio-ribose cinase*

(MTK; EC 2.7.1.100), foi identificada inicialmente em arroz e *A. thaliana*, e catalisa a conversão dependente de ATP da 5-metil-tio-ribose (MTR) em 5-metil-tio-ribose-1-fosfato (MTR-1-P) (Sauter et al., 2004). Como *A. thaliana* possui somente um gene MTK, Pommerrenig e colaboradores identificaram e caracterizaram em *A. thaliana* e *Plantago* as duas próximas enzimas do ciclo, a *MTR-1-P isomerase* (MTI; EC 5.3.1.23), que converte MTR-1-P em 5-metil-tio-ribulose-1-fosfato (MTRu-1-P), e o complexo *desidratase-enolase-fosfatase* (DEP; EC 4.2.1.109 e 3.1.3.77), que catalisa a conversão de MTRu-1-P em 1,2-diidroxi-3-ceto-5-metil-tio-penteno (DHKMP) (Pommerrenig et al., 2011). Vale a pena ressaltar que a mesma conversão catalisada por estas duas enzimas em plantas é realizada em bactérias, fungos e animais em quatro diferentes passos enzimáticos. Subsequentemente, a quinta reação é catalisada por uma *acidoreductona oxigenase* (ARD; EC 1.13.11.53 e 1.13.11.54), que converte DHKMP em 2-ceto-4-metil-tio-butirato (KMTB) (Sauter et al., 2005). Por fim, pouco se sabe sobre a conversão de KMTB em metionina, mas estudos *in silico* demonstram que em *A. thaliana* os produtos dos genes referentes à *aspartato aminotransferase* (AAT ou ASP; EC 2.6.1.1) podem desempenhar tal função metabólica (Pommerrenig et al., 2011). Dados experimentais demonstraram que em plantas superiores o ciclo de Yang ocorre preferencialmente em células do floema e sugerem também que em folhas de *A. thaliana* e *Plantago* esta rota recicla MTA derivado principalmente da biossíntese de poliaminas (Pommerrenig et al., 2011).

O primeiro metabólito importante na biossíntese de etileno seria S-adenosil-metionina (SAM ou AdoMet), que é sintetizado a partir de ATP e metionina (fornecida pelo ciclo de Yang), por ação da enzima *metionina-adenosil transferase* (MAT), conhecida também como *S-adenosil-metionina sintetase* (SAMS; EC 2.5.1.6) (Roje, 2006). Esta enzima é extremamente conservada entre os organismos vegetais e apresenta-se quase sempre como tetrâmeros, tendo seus monômeros organizados em três domínios, formados por trechos não consecutivos de sequências, contendo cada um duas  $\alpha$ -hélices e quatro folhas  $\beta$ . Estes domínios são denominados respectivamente de N-terminal (importante na interação com substrato), central (com função não caracterizada completamente, mas sabe-se que influencia na catálise) e C-terminal (essencial para a localização subcelular da enzima no citoplasma) (Takusagawa et al., 1996; Garrido et al., 2011).

A S-adenosil-metionina, produzida na reação catalisada por MAT, serve como substrato da principal enzima da biossíntese de etileno, a *1-aminociclopropano-1-ácido*

*carboxílico sintase* (ACS; EC 4.4.1.14), que produz 1-aminociclopropano-1-ácido carboxílico (ACC) e MTA, que é reciclado pelo ciclo de Yang (Argueso et al., 2007). As ACSs pertencem a família de proteínas que requerem piridoxal-5'-fosfato (PLP) como cofator e são associadas com reações de transaminação, deaminação, carboxilação e eliminação ou substituição dos carbonos  $\beta$  e  $\gamma$  em uma variedade de aminoácidos. Embora diverjam na sequência e conseqüentemente na atividade catalítica, as estruturas cristalográficas das enzimas dependentes de PLP revelam uma impressionante conservação da estrutura do núcleo catalítico, indicando mecanismos comuns de catálise. As ACSs possuem similaridade de sequência peptídica com enzimas da classe das *aminotransferase* (aspartato e tirosina) e convertem SAM em ACC através de uma reação de eliminação dos carbonos  $\beta$  e  $\gamma$  (Capitani et al., 1999; Li et al., 2005). Estas enzimas, que podem ser homo ou heterodímeros citoplasmáticos, atingem atividade ótima em condições de pH entorno de 8,5 e são induzidas por diversos fatores, tais como situações de estresse (ferimentos, alagamento e seca), outros fitormônios (IAA e citocininas, por exemplo), metais citotóxicos (cobre, ferro, prata, lítio e zinco), compostos inorgânicos (amônia e bissulfito) e orgânicos (ácido ascórbico), herbicidas, pesticidas e desfolhante (cianeto de potássio). No entanto, as ACSs podem ser inibidas por diversos componentes, destacando-se S-adenosil-etionina, que compete com o substrato da enzima (AdoMet), além de aminoetóxi-vinil-glicina (AVG) e ácido aminoxacético (AOA), inibidores competitivos de muitas enzimas dependentes de PLP. Vale ressaltar que intermediários do ciclo de Yang (MTA e KMTB) também podem inibir a atividade destas enzimas (Acaster & Kende, 1983; Yip et al., 1992; Lewis et al., 2011). Na maioria das espécies de plantas, as ACSs são codificadas por famílias multigênicas, que são diferencialmente reguladas por diversos fatores ambientais e do desenvolvimento. Em *A. thaliana*, existem oito genes que codificam ACSs ativas, e pelo menos um que codifica uma ACS com inatividade catalítica (AtACS1) (Liang et al., 1992; Liang et al., 1995; Yamagami et al., 2003). Estas proteínas podem ser divididas em três grupos, baseados na sequência presente em sua região C-terminal: *tipo I*, que possui nesta porção da proteína resíduos de serina (Ser) que são alvos de fosforilação pela *proteína cinase ativada por mitógeno 3 ou 6* (AtMPK3-6; EC 2.7.11.24) (Liu & Zhang, 2004), assim como resíduos de aminoácido que são fosforilados por *proteína cinase cálcio dependente* (AtCDPK2; CDPK ou CPK; EC 2.7.11.1); *tipo II*, que possuem um C-terminal reduzido e apresentam sítios de fosforilação somente para CPK (Tatsuki & Mori, 2001;



Sebastia et al., 2004); e *tipo III*, que possuem a porção C-terminal muito reduzida e não apresentam sítios de fosforilação para ambas as cinases (Tucker et al., 2010).

O metabólito ACC é o elemento chave da biossíntese de etileno e pode seguir por várias rotas metabólicas. A principal delas é a catálise final da biossíntese deste fitormônio gasoso na qual ACC, associado à ascorbato (Asc) e oxigênio (O<sub>2</sub>), produz etileno, diidroascorbato (DHAsc), cianeto (HCN), gás carbônico (CO<sub>2</sub>) e água (H<sub>2</sub>O). A enzima responsável por esta reação é a *1-aminociclopropano-1-ácido carboxílico oxidase* (ACO ou EFE - *ethylene forming enzyme*; EC 1.14.17.4), pertencente à família das oxidoreduases ou oxigenases (mononucleares férricas não-hêmicas), que funcionam com doadores pareados, onde o O<sub>2</sub> atua como oxidante para posteriormente ser incorporado ou reduzido (Hegg & Que-Jr, 1997). Estas enzimas são caracterizadas por possuírem motivos 2-histidina-1-ácido carboxílico de interação com ferro [2OG-Fe(II) oxigenase] e são capazes de catalisar uma variedade de reações, tais como hidroxilações, insaturações, expansão e fechamento oxidativo de anéis. No caso da enzima ACO, que pode ser inibida por altas concentrações de CO<sub>2</sub> e temperaturas elevadas (acima de 35 °C), o ACC é convertido em etileno por modificações nos carbonos C-2 e C-3 do ACC, enquanto C-1 é convertido em cianeto e o grupamento carboxil é eliminado na forma de dióxido de carbono (Peiser et al., 1984; Hegg & Que-Jr, 1997). Com relação à localização subcelular desta enzima, os estudos a respeito do assunto são conflitantes, uma vez que alguns autores postulam que esta proteína seja integral de membrana, devido à dificuldade em sua purificação e a inibição da conversão de ACC em etileno em condições de choque frio e osmótico, ou ainda quando as amostras são previamente tratadas com detergentes (Apelbaum & Yang, 1981; Mayne & Kende, 1986). Por outro lado, existem evidências experimentais de que a enzima ACO de tomate e maçã sejam citoplasmáticas, uma vez que as mesmas não possuem em seu N-terminal o peptídeo sinal que direcione sua síntese para o retículo endoplasmático rugoso e sua subsequente exportação para a membrana plasmática (Reinhardt et al., 1994; Chung et al., 2002).

Em 1981, Apelbaum e colaboradores descobriram que quantidades de ACC em folhas de trigo sob condições de déficit hídrico sempre foi muito maior do que a quantidade produzida de etileno, o que sugere a existência outra rota metabólica envolvendo ACC (Apelbaum & Yang, 1981). Posteriormente, estudos com marcação radioativa e espectrometria de massas demonstraram a existência da rota de N-malonização

de ACC em diversos tecidos vegetais, onde o conjugado 1-malonil-ACC (MACC) seria formado pela ação da enzima *1-aminociclopropano-1-ácido carboxílico-N-malonil transferase*, enzima esta que já foi purificada a partir de extratos proteicos vegetais, mas sem a identificação de seu respectivo gene (Hoffman et al., 1982; Amrhein et al., 1984). Esta rota metabólica ocorre frequentemente em plantas sob vários tipos de estresse, como o crescimento de folhas de trigo no claro, grão de bico pré-aquecido ou crescimento de sementes de musgo, além de ser regulada pela concentração de auxinas (IAA), ABA e etileno (Liu et al., 1985; Nath et al., 2006). Experimentos com ACC marcado radioativamente em protoplastos de *Acer pseudoplatanus*, a síntese de MACC ocorre no citosol, sendo posteriormente transportado para os vacúolos, com auxílio de proteínas carreadoras dependentes de ATP, presentes no tonoplasto, onde é sequestrado para o espaço apoplásmico através da membrana plasmática (Bouzayen et al., 1988; Bouzayen et al., 1989). O conjugado de ACC pode ainda se difundir livremente por variação do pH vacuolar (Pedreno et al., 1991). Geralmente, a reação de formação de MACC é considerada irreversível, sendo o metabólico um produto final inativo em vez de uma forma de estoque de ACC (Amrhein et al., 1984). Esta hipótese foi inicialmente comprovada por Hoffman e colaboradores, que mostraram níveis constantes MACC em folhas de trigo sob estresse hídrico, enquanto que os níveis de ACC e etileno mudavam bruscamente (Hoffman et al., 1982). Estes dados contrastaram com as observações apresentadas por Amrhein e colaboradores, onde a recuperação de ACC a partir de MACC foi detectada em hipocótilos de trigo mesmo que em níveis muito baixos (Amrhein et al., 1984). Acredita-se que para tais fins MACC seria substrato da enzima *MACC hidrolase*, que catalisaria esta reação de recuperação de ACC em tecidos onde o conjugado estivesse presente em altos níveis (Yang & Hoffman, 1984; Jiao et al., 1986; Fluhr et al., 1996).

Sabe-se que além da rota metabólica de formação de MACC, foi observado em tomates que ACC também pode ser conjugado na forma de 1-glutamil-ACC (GACC), reação que é catalisada pela enzima *γ-glutamil transpeptidase* (GGT; EC 2.3.2.2). Pouco se sabe sobre esta rota metabólica de conjugação de ACC uma vez que, semelhante à biossíntese de MACC, a enzima que catalisa esta reação foi identificada somente em extratos proteicos vegetais, sem identificação de seu respectivo gene. Ainda assim, a formação do conjugado 1-malonil-ACC é considerada a principal via secundária envolvendo o precursor de etileno, possuindo um importante papel na regulação dos seus

níveis endógenos e conseqüentemente na regulação da biossíntese do fitormônio gasoso em tecidos vegetais (Martin et al., 1995).

Outra possível rota metabólica com a participação de ACC seria a que envolve a deaminação irreversível do metabólito. Este processo, descrito inicialmente em bactérias, é catalisado pela enzima *1-aminociclopropano-1-ácido carboxílico deaminase* (ACD; EC 3.5.99.7), proteína pertencente à classe das hidrolases que, na presença de água, converte ACC em oxobutirato (OXB ou 2-oxobutanoato) e amônia (NH<sub>3</sub>) (Glick, 2005). A expressão heteróloga de ACDs em plantas mostra ser possível gerar plantas transgênicas que tenham seus níveis de etileno reduzidos pela redução dos níveis de ACC disponível (Klee et al., 1991). Posteriormente, foram identificados em plantas genes que codificam proteínas homólogas à ACDs bacterianas em *A. thaliana* e *Populus*, genes esses que se mostram ativos em diversas condições do desenvolvimento vegetal (McDonnell et al., 2009). Além disso, estudos em tomate mostram que a atividade inerente de ACD varia durante o amadurecimento do fruto, sendo que seus picos de atividade coincidem com as quedas na produção de etileno (Plett et al., 2009).

Desta forma, a maioria dos tecidos vegetais possui a capacidade de sintetizar etileno, mesmo que em níveis baixos. A sua produção aumenta drasticamente durante diversos eventos do desenvolvimento, tais como germinação, amadurecimento de frutos, senescência e abscisão de flores e folhas, resposta a outros fitormônios como auxinas, brassinoesteróides, citocininas e o próprio etileno (auto estimulação e auto inibição). Finalmente, uma grande variedade de estresses, incluindo ferimentos, ataque de patógenos, inundação, seca, hipóxia, carga física, mudanças bruscas de temperatura e reagentes químicos nocivos (ozônio e dióxido de enxofre) podem induzir a produção deste fitormônio gasoso (Yang & Hoffman, 1984; Bleeker, 1999).

### **1.2.2. Sinalização.**

Diversos estudos descrevem mecanismos moleculares das rotas de metabólicas sinalizadas por etileno (Johnson & Ecker, 1998; Lin et al., 2009). Nestes trabalhos foram identificadas proteínas capazes de perceber a presença deste fitormônio gasoso, proteínas estas que possuem grande similaridade com o sistema histidina cinase de dois componentes

identificado em bactérias. Esta família de receptores vegetais é caracterizada por um domínio de interação com etileno presente na porção N-terminal, seguido por dois outros domínios, um de interação entre receptores, denominado de GAF (*cGMP-specific phosphodiesterases*, *adenylyl cyclases* e *FhlA*) (Aravind & Ponting, 1997; Xie et al., 2006; Grefen et al., 2008), e outro semelhante à proteína histidina cinase (Bleecker, 1999). Além desses, alguns membros dessa família possuem o domínio *coiled coil*, que representa a segunda metade do sistema de dois componentes. Após a caracterização em diversas plantas, estes receptores de etileno (ETR; EC 2.7.13.-) foram divididos em duas subfamílias: a *subfamília I* possui membros com atividade histidina cinase [com exceção do receptor ERS1 (*ethylene response sensor*) de *A. thaliana* que possui tanto atividade histidina quanto serina/treonina cinase] e três domínios transmembrana em seu N-terminal, enquanto que a *subfamília II* possui membros com atividade serina/treonina cinase e quatro domínios transmembrana (Hall et al., 2007). Em *A. thaliana* foram caracterizados cinco receptores de etileno que contribuem de forma distinta na sinalização mediada por este fitormônio, sendo dois pertencentes à subfamília I (ETR1 e ERS1) (Chang et al., 1993; Hua et al., 1995) e três a subfamília II [ETR2, *ethylene insensitive 4* (EIN4) e *ethylene response sensor 2* (ERS2)] (Hua et al., 1998; Sakai et al., 1998). Os receptores de etileno podem estar presentes como homodímeros em mais de um sistema de membrana. Em *A. thaliana*, ETR1 e ETR2 foram localizados no retículo endoplasmático (ER) a partir de dados de fracionamento em gradiente de sacarose (Hall et al., 2007) e posteriormente todos os receptores desta planta modelo tiveram a mesma citolocalização quando foram expressos em células epidermais de folhas de tabaco (Grefen et al., 2008). Em contraste, experimentos de imunohistoquímica com células radiculares de *A. thaliana* demonstraram que homodímeros de ETR1 podem estar localizados primariamente no complexo de Golgi (Dong et al., 2008). Já em protoplastos de tabaco, foi relatado que o receptor de etileno localiza-se na membrana plasmática (Hall et al., 2007). Mesmo com dados contrastantes, a citolocalização dos receptores de etileno mais aceita na atualidade é a referente ao retículo endoplasmático.

Estudos bioquímicos e genéticos mostraram que a interação do etileno com os homodímeros de receptores requer como cofator o íon cobre ( $\text{Cu}^{2+}$ ), o que está de acordo com dados de que a interação reversível etileno-receptor requer um metal de transição (Burg & Burg, 1967; Rodriguez et al., 1999). Em *A. thaliana*, o papel do cobre na

percepção do fitormônio gasoso está diretamente associado com o produto do gene RAN (*responsivo à agonista*), um transportador ativo (dependente de ATP), presente na membrana do complexo de Golgi, e que possui como função primordial gerar receptores de etileno funcionais por fornecimento intracelular de cobre (Bleecker, 1999; Hirayama et al., 1999; Binder et al., 2010).

Além disso, em *A. thaliana*, a proteína integral de membrana RTE1 (*reversion to ethylene sensitivity 1*) está envolvida com a regulação das funções do receptor de etileno ETR1 (Resnick et al., 2006). Esta proteína é altamente conservada em plantas, animais e alguns protistas, mas sua função molecular ainda não é conhecida, mas em plantas possui a mesma localização que ETR1 (retículo endoplasmático e complexo de Golgi) (Dong et al., 2008). RTE1 foi identificada com base na análise de mutantes nocaute para o gene *rte1* que suprimem a insensibilidade a etileno ocasionada pelas mutações *etr1-2* (Resnick et al., 2006). Ainda em *A. thaliana*, a insensibilidade à etileno ocasionada pela alta expressão de RTE1 é altamente dependente de ETR1, mas não de outros receptores para o gás, sugerindo que possivelmente esta proteína atuaria em processos de reciclagem de receptores ETR1 nesta planta modelo. Além disso, a expressão de RTE1 é induzida por etileno, sugerindo a participação desta proteína em mecanismos de retroalimentação negativa na sinalização mediada pelo gás (Resnick et al., 2006).

Geneticamente, na ausência de etileno, os receptores ETR são reguladores negativos da resposta mediada por este fitormônio, reprimindo elementos a jusante na cascata de sinalização, através da proteína cinase semelhante à RAF, denominada de CTR (*constitutive triple response*; EC 2.7.11.1) (Kieber et al., 1993; Hua et al., 1998; Chen et al., 2005; Hall et al., 2007; Qu et al., 2007). Esta proteína cinase foi identificada inicialmente em mutantes que apresentaram evidências morfológicas da resposta tripla (inibição do alongamento caulinar, o espessamento do caule e perda da sensibilidade gravitrópica) na ausência de etileno exógeno (Kieber et al., 1993). Mutações de perda de função *ctr1* resultam na ativação constitutiva da resposta mediada por etileno em mudas e plantas adultas, o que indica que CTR atua como um regulador negativo da sinalização mediada por este fitormônio gasoso (Kieber et al., 1993). A porção C-terminal desta proteína cinase é similar ao domínio catalítico da família RAF de proteínas cinases, que geralmente atuam como reguladores de outras cinases, por auto regulação e/ou outros elementos regulatórios (Jouannic et al., 1999; Ichimura et al., 2002). A respeito da

localização subcelular da proteína MKKK (*MAP cinase cinase cinase*) CTR, estudos demonstram que dímeros da proteína cinase estão associadas à face externa da membrana do retículo endoplasmático devido a interações com receptores da família ETR (Gao et al., 2003; Ju et al., 2012; Mayerhofer et al., 2012).

Na ausência do fitormônio gasoso, a porção N-terminal de CTR interage com os membros das duas subfamílias de receptores (Clark et al., 1998; Gao et al., 2003). Nesta conformação, CTR inibe as respostas sinalizadas por etileno pela fosforilação de elementos importantes nesta rota metabólica. Já na presença de etileno, ocorre a formação do híbrido etileno-receptor, devido à alta afinidade que existe entre este fitormônio e os domínios transmembrânicos presentes nos homodímeros de ETR (formados por ponte de dissulfeto entre monômeros), na presença de cobre (Schaller & Bleecker, 1995). A formação deste híbrido induz por sua vez uma mudança conformacional no receptor, e conseqüentemente em CTR, o que reduz drasticamente a atividade desta cinase, induzindo assim a ativação da sinalização mediada pelo fitormônio gasoso.

Uma das rotas vegetais reguladas por CTR em respostas à etileno seria a representada pela proteína EIN2 (*ethylene insensitive 2*), um importante regulador positivo das rotas de sinalização mediadas por este fitormônio gasoso, que atua a jusante de CTR (Roman et al., 1995; Alonso et al., 1999). Esta proteína localiza-se no retículo endoplasmático, próxima ou associada à ETRs (Bisson et al., 2009; Bisson & Groth, 2010). Além disso, EIN2 é formada por uma porção N-terminal, com alta similaridade à transportadores Nramp de íons metálicos, apresentando doze hélices transmembrana preditas, além de um domínio C-terminal hidrofílico e citoplasmático, com sinal de localização nuclear (NLS) (Alonso et al., 1999).

Estudos *in vitro* e *in vivo* demonstram que, na ausência de etileno, os níveis de EIN2 são reduzidos devido à eventos envolvendo proteínas *F-box* e subsequente degradação via proteassoma 26S (McClellan & Chang, 2008). Estes eventos podem ser propiciados pela fosforilação de sítios específicos, presentes no domínio C-terminal de EIN2 (Chen et al., 2011; Ju et al., 2012). Em *A. thaliana*, ETP1 e ETP2 (*EIN2 targeting proteins*) possuem um papel primordial nesta proteólise de EIN2. Estas proteínas com domínios *F-box* interagem com a porção C-terminal de EIN2 (que é extremamente conservada), previamente fosforilada por CTR. Desta forma, na ausência de etileno, EIN2 com o C-terminal fosforilado é então ubiquitinada, para posteriormente ser degradada pelo

proteassoma 26S (Qiao et al., 2009). Já na presença do fitormônio gasoso, a porção C-terminal de EIN2 não é fosforilada por CTR e, desta forma, este domínio da proteína é clivado e, por mecanismos ainda não elucidados, é translocado para o núcleo, onde influencia a expressão e atividade de importantes fatores de transcrição (Ju et al., 2012).

Sabe-se que na ausência de etileno, os fatores de transcrição EIN3 (*ethylene insensitive 3*) e seus equivalentes funcionais EILs (*EIN3-like*) sofrem degradação pelo proteassoma 26S (Qiao et al., 2009) e acredita-se que EIN2 influencia, através de sua porção C-terminal, a atividade desses fatores no núcleo. Sabe-se que EIN3/EIL são importantes elementos que regulam a expressão de outros fatores de transcrição, que por sua vez interagem com elementos *cis*-atuantes de promotores (*GCC-box*) de genes responsivos à etileno (Solano et al., 1998). Estas proteínas são encontradas com alta homologia dentro do reino vegetal, principalmente com relação à sua porção N-terminal, que é responsável pela interação com o DNA, através de um domínio composto por cinco hélices (Solano et al., 1998; Yamasaki et al., 2005). A estabilidade de EIN3/EIL é mediada por outra classe de proteínas *F-box* associadas à *E3 ubiquitina ligase*, denominadas em *A. thaliana* de EBF1 e EBF2 (*EIN3 binding F-box protein 1 and 2*) (Guo & Ecker, 2003; Stepanova & Alonso, 2009), uma vez que na ausência de etileno, os fatores de transcrição em questão são direcionados para degradação por estas proteínas *F-box* (An et al., 2010). Na presença de etileno, CTR e os receptores ETR são inativados, e na ausência da fosforilação de EIN2, o C-terminal desta proteína pode então ser clivado e ser direcionado para o núcleo, onde induz, por mecanismos ainda desconhecidos, a degradação das proteínas EBF via proteassoma 26S (An et al., 2010; Ju et al., 2012).

Outro mecanismo descrito para regulação da atividade de EIN3/EIL também seria mediada pela exoribonuclease citoplasmática 5'-3' XRN4, denominada de EIN5 (EC 3.1.13.-). Estudos em *A. thaliana* demonstram que mutações no gene desta nuclease são capazes de suprimir parcialmente a resposta constitutiva mediada por etileno tanto em brotos quanto em plantas adultas. Estes dados sugerem que EIN5 está envolvido na cascata de transdução de sinal mediada por este fitormônio, atuando a jusante de CTR (Olmedo et al., 2006). Adicionalmente, na presença do gás, os mRNAs de EBF se acumulam em mutantes nocaute para o gene *ein5*, gerando fenótipos de insensibilidade à etileno. Desta forma, acredita-se EIN5 atue na degradação dos mRNAs de EBF, diminuindo assim os

respectivos níveis de sua proteína livre, o que regula positivamente a atividade de EIN3/EIL e propicia a indução da resposta sinalizada por etileno (Olmedo et al., 2006).

Estudos mais antigos supunham que a rota de ativação de EIN3/EIL também podia ser influenciada por CTR, independentemente de EIN2, pela atividade de *proteínas cinases ativadas por mitógeno* (MAP cinases ou MAPKs). Dados mostravam que em *A. thaliana*, a proteína MKK9 (*MAP cinase cinase*; EC 2.7.12.2), em associação com as MAPKs MPK3 e MPK6, seriam reguladas positivamente pela inibição de CTR na presença de etileno, induzindo assim respostas à jusante de CTR (Yoo et al., 2008; Hahn & Harter, 2009; Stepanova & Alonso, 2009). Desta forma, a cascata de fosforilação MKK9 → MPK3/MPK6 culminaria na fosforilação e subsequente estabilização de EIN3/EIL, o que estaria de acordo com o que foi observado em mutantes nocaute de *A. thaliana* para o gene *mkk9*, que apresentaram fenótipo de insensibilidade parcial à etileno (Yoo et al., 2008).

Em paralelo, estudos independentes demonstraram que as MPKs de *A. thaliana* MPK3/MPK6, juntamente com as MKKs MKK9 ou MKK4/5 regulam a biossíntese deste fitormônio gasoso (Liu & Zhang, 2004; Joo et al., 2008; Xu et al., 2008). Foi demonstrado que membros dos dois principais grupos das enzimas ACS de biossíntese de etileno (tipo I e tipo II), representadas em *A. thaliana* por ACS5/ACS9 e ACS6/ACS2 respectivamente, são reguladas por diferentes proteínas cinases. Estudos com esta planta modelo em situações de estresse mostram que a estabilidade das ACSs do tipo I é regulada pelas proteínas cinases MPK3 e MPK6 a partir da cascata de fosforilação MKK4/MKK5/MKK9 → MPK3/MPK6 (Ecker, 2004; Liu & Zhang, 2004). Estas enzimas da biossíntese de etileno possuem uma pequena vida útil e, na ausência do fitormônio gasoso, são alvos de degradação proteolítica. Já na presença do sinal de estresse, a porção C-terminal das ACSs do tipo I é fosforilada pela cinases MPK3/MPK6, o que aumenta a estabilidade destas proteínas. As enzimas ACS do tipo II possuem um mecanismo similar de regulação, uma vez que, semelhante às ACSs do tipo I são continuamente alvos de degradação na ausência de estímulos, que neste caso seria de natureza endógena, como por exemplo, altos níveis de citocininas (Chae et al., 2003). Esta degradação é mediada pelas proteínas ETO (*ethylene overproducer*) e EOL (*ETO like*), membros da pequena família de proteínas específicas de plantas que possuem o domínio BTB/TPR, importante em interações entre proteínas, além de possuírem atividade de *E3 ubiquitina ligase* (Wang et al., 2004). Na ausência de um sinal endógeno, as proteínas ETO/EOL interagem com a exterminada C-terminal das ACSs



do tipo II, o que as torna passíveis de degradação via proteossoma 26S. Existem evidências de que na presença de um sinal, a porção C-terminal desta classe de ACSs seja fosforilada pela *proteína cinase dependente de cálcio 2* (CPK2 ou CDPK2), evento esse que estabiliza estas enzimas inibindo sua degradação (Ecker, 2004; Sebastia et al., 2004; Wang et al., 2004). Vale a pena ressaltar que a biossíntese de etileno também pode ser regulada em nível transcricional, uma vez que em *A. thaliana* sob situações de estresse, os níveis de mRNAs de ACSs do tipo II são elevados (Cara & Giovannoni, 2008; Tsuchisaka et al., 2009; Vandenbussche et al., 2012).

Dados mais recentes demonstraram que a cascata de fosforilação mediada por MKK9 pode sim influenciar a atividade de EIN3/EIL, mas de forma independente de CTR e EIN2 (An et al., 2010). Com base nestes estudos foi proposto um modelo em que estas proteínas cinases participam em respostas mediadas por etileno, mas principalmente pela modulação de sua biossíntese. No entanto, EIN2 e EBFs são transdutores de sinais essenciais, requeridos para a percepção do fitormônio gasoso na estabilização de EIN3/EIL, atuando em cooperação, mas de forma diferencial (Kendrick & Chang, 2008; Stepanova & Alonso, 2009; An et al., 2010).

A sinalização mediada por etileno produz grandes mudanças na expressão gênica em células vegetais. Análises de regiões promotoras de genes induzidos pelo fitormônio gasoso propiciaram a identificação de elementos *cis*-atuantes responsivos ao gás, assim como a família de proteínas *trans*-atuantes de interação com o DNA, denominada de EREBP (*ethylene responsive element binding proteins*) ou ERF (*ethylene response factors*) (Ohme-Takagi & Shinshi, 1995; Deikman et al., 1998; Leubner-Metzger et al., 1998). Estudos menos recentes demonstraram que EIN3/EIL são ativadores do gene ERF1 (*ethylene response factor 1*), membro da família ERF, estabelecendo uma hierarquia na sinalização mediada por etileno (Solano et al., 1998). Os homodímeros de EIN3/EIL interagem com elementos *cis*-atuantes na região promotora do gene ERF1, que uma vez transcrito e traduzido, interage com outros elementos *cis*-atuantes presentes na região promotora de genes alvo (Solano et al., 1998). Hoje se sabe que EIN3 pode induzir a transcrição não somente de ERF1, mas também de outros membros da superfamília de fatores de transcrição AP2/ERF (Vandenbussche et al., 2012).

Historicamente, as proteínas da superfamília AP2/ERF contêm pelo menos um domínio de interação com o DNA, denominado de AP2 (*APETALA 2*), e podem ser

divididas em três famílias, nomeadas de AP2, ERF (*ethylene response factor*) e RAV (*related to ABI3/VPI*) (Nakano et al., 2006). A família AP2 é formada por membros caracterizados por repetições *in tandem* de dois domínios AP2, sendo somente um pequeno número de proteínas que apresentam uma cópia deste domínio, que por sua vez apresenta maior similaridade com membro que possuem duplo AP2 do que com proteínas da família ERF, que também apresenta somente uma cópia deste domínio (Shigyo & Ito, 2004). A terceira família, nomeada de RAV (*related to ABI3/VPI*), possui dois elementos de interação com o DNA, sendo um deles o AP2 (similar ao da família ERF), e outro denominado de B3, domínio identificado inicialmente na proteína de milho VP1 (*viviparous 1*), composto por duas  $\alpha$ -hélices e sete folhas  $\beta$  (Yamasaki et al., 2004; Swaminathan et al., 2008). A porção conservada do domínio AP2 é composta de aproximadamente 60 resíduos de aminoácidos que conferem uma conformação tridimensional típica organizada em uma porção com três folhas  $\beta$  antiparalelas seguida de uma  $\alpha$ -hélice (Allen et al., 1998).

Mesmo que o acrônimo original da família ERF tenha sido mantido, a resposta mediada por estes fatores de transcrição não é exclusiva do fitormônio etileno. Além disso, a nomenclatura dos membros desta família é baseada em dois grandes esforços em organizá-las em grupos filogenéticos que permitissem facilitar o seu reconhecimento funcional (Sakuma et al., 2002; Nakano et al., 2006). Sakuma e colaboradores (2002) dividiram em dois grupos (ou subfamílias) as 121 proteínas ERF de *A. thaliana*, de acordo com a similaridade na sequência de resíduos de aminoácidos do domínio AP2: *grupo A*, onde estavam as proteínas com elementos de interação responsivos à desidratação (DREBs); e *grupo B* os ERFs, sendo que cada grupo foi ainda subdividido em outros 6 (A1-A6; B1-B6). Já Nakano e colegas (2006) exploraram o sequenciamento do genoma do arroz (*Oryza sativa*) e *A. thaliana* para refinar a classificação de Sakuma, explorando a estrutura dos genes ERF além de levar em consideração a ocorrência de motivos adicionais. Nesta classificação, as proteínas ERF foram subdivididas em doze grupos (I até X, VI-L e Xb-L), que coincidiam substancialmente com a classificação feita por Sakuma. Hoje em dia, ambas as classificações são utilizadas independentemente (Licausi et al., 2013).

A especificidade da interação do domínio AP2 com o DNA se dá pela identidade de resíduos de aminoácidos em posições específicas (Sakuma et al., 2002). Muitas proteínas

DREB interagem com o elemento *cis*-atuante A/GCCGAC denominados de *DRE* (*drought response element*), presentes em promotores, estando frequentemente associados com genes regulados também em resposta à ABA, seca e frio (Stockinger et al., 1997). Já as proteínas ERF reconhecem *in vitro* o elemento *cis*-atuante AGCCGCC, denominado de *GCC-box*, que está relacionado principalmente com respostas à etileno, patógenos e ferimentos (Ohme-Takagi & Shinshi, 1995). Apesar da generalização, alguns membros dos dois grupos podem reconhecer tanto os elementos *DRE* quanto *GCC-box* ou ainda elementos que divergem significativamente destes dois (Welsch et al., 2007; Shaikhali et al., 2008; Sun et al., 2008). Surpreendentemente, membros da família AP2 que possuem o domínio duplicado não reconhecem repetições *in tandem* dos elementos *cis*-atuantes de interação *DRE* ou *GCC-box* (Dinh et al., 2012). Por fim, o domínio AP2 das proteínas RAV reconhecem motivos CAACA, enquanto que o domínio B3 interage com sequências CACCTG (Kagaya et al., 1999; Ohta et al., 2000; Yamasaki et al., 2004).

Seguindo a regra geral, os fatores de transcrição contendo o domínio AP2 podem ser classificados como ativadores ou repressores, dependendo se os mesmos induzem ou suprimem a transcrição de genes alvo específicos. Os repressores podem ainda ser classificados como *ativos*, quando possuem domínios distintos de repressão (RD), ou *passivos*, que são aqueles que não possuem RD, mas suprimem a transcrição por competição com ativadores transcricionais (Licausi et al., 2013). Em geral, os domínios de ativação em planta não possuem sequências específicas, mas apresentam pelo menos uma região rica em aminoácidos ácidos, o que contrasta com os domínios de repressão, que são formados por sequências bem definidas (Ohta et al., 2000; Tiwari et al., 2012). Três RDs, denominados respectivamente de *motivo anfífilico de repressão associado à ERF* (EAR; LxLxLx ou DLNxxP) (Ohta et al., 2001; Hiratsu et al., 2003), *motivo TLLLFR* (Matsui et al., 2008), e *domínio de repressão B3* (BRD; RLFGV) (Ikeda & Ohme-Takagi, 2009) são específicos de plantas e estão presentes em alguns membros da superfamília de fatores de transcrição AP2/ERF.

Por fim, a expressão e atividade destes fatores de transcrição podem ser reguladas em vários níveis. Estudos demonstram que ERF1 (sob condições de estresse biótico) (Lorenzo et al., 2003), CBF1 (estresse abiótico) (Jaglo-Ottosen et al., 1998), SMZ (durante o crescimento) (Mathieu et al., 2009) e PLTs (no decorrer do desenvolvimento) (Aida et al., 2004) são regulados em diferentes plantas durante a transcrição. Já o DREB2A de

grama, cevada, trigo e milho apresentam sua expressão regulada por mecanismos pós-transcricionais (processamento de mRNA) (Egawa et al., 2006). Além disso, modificações pós traducionais, como fosforilação, sumoilação e ubiquitinação são mecanismos comuns na ativação/inativação de membros desta superfamília de fatores de transcrição (Bethke et al., 2009; Bailey-Serres et al., 2012).

### **1.2.3. Transporte.**

A maioria dos fitormônios é transportada através de mecanismos conhecidos, associados com proteínas específicas de transporte e/ou conjugados. Para pequenos deslocamentos, a transmissão citoplasmática e a difusão lenta são os mais utilizados. Já para longas distâncias, os hormônios podem ser transportados pelo sistema vascular, sendo pelo floema juntamente com carboidratos, partindo das folhas tendo como destino raízes e flores, ou pelo xilema, com água e minerais a partir das raízes até as folhas. Como o etileno é um gás, o seu transporte ocorre por difusão entre as células, sem necessidade de um sistema de transporte específico, mas o seu precursor, o ACC, pode ser transportado por longas distâncias pelo xilema (Finlayson et al., 1991; Hoad, 1995).

### **1.2.4. Catabolismo.**

Mesmo sem relevância fisiológica, o rastreamento onde o precursor de etileno foi marcado com isótopos radioativos propiciou a identificação de dióxido de carbono, óxido de etileno, etileno glicol e glicose conjugada à etileno glicol como produtos do catabolismo do gás (Raskin & Beyer, 1989).

### **1.2.5. Etileno em Situações de Estresse.**

Talvez as primeiras observações da elevação dos níveis de produção de etileno sob condições de estresse seriam os estudos onde tecidos vegetais são cortados, perfurado,

esmagado, espremido ou bloqueados por barreiras físicas. Posteriormente, foi observada sua presença em estresses mecânicos que não envolvem cortes ou feridas, como o vento, que pode induzir a produção de etileno, inibindo o alongamento da haste (Emery et al., 1994).

Outra característica comum está no fato de que tanto ferimentos quanto infecções microbianas podem estimular a produção deste gás pela indução e regulação da atividade de proteínas específicas (Bol et al., 1996; Schweizer et al., 1996). Em tomates, a indução da expressão de *inibidores de proteinases* (PINs) por ataques de insetos é sinalizada por etileno e ácido jasmônico (O'Donnell et al., 1996). A expressão do gene ETR1 de *A. thaliana* em tabaco conferiu suscetibilidade a fungos não patogênicos de solo (Knoester et al., 1998). Em tomate, mutantes nocaute para genes da biossíntese de etileno tiveram respostas secundárias de resistência sistêmica adquirida (SAR) substancialmente reduzidas em lesões pontuais e infecções por bactérias (Lund et al., 1998).

Mais evidências de que o etileno participa em rotas de defesa sinergicamente com o ácido jasmônico foram comprovadas, uma vez que análises de mutantes nocaute de *A. thaliana* para os genes *etr1* e *jar1* (*jasmonate response deficient 1*) mostram deficiência em eventos de resistência sistêmica induzida (ISR), quando comparadas com plantas selvagens (Pieterse et al., 1998; Pieterse & Van Loon, 1999). Experimentos semelhantes mostraram que o sinergismo entre as rotas sinalizadas pelos dois fitormônios se dá pela atividade do fator ERF1, que apresenta sua expressão elevada na presença de ambos (Lorenzo et al., 2003). Além disso, a sinalização mediada por estes dois fitormônios é necessária para a elevação dos níveis de expressão do gene PDF1.2 (*plant defensin 1.2*), que codifica uma defensina antimicrobiana (Penninckx et al., 1996; Alonso et al., 1999). Estudos de indução com mutantes nocaute para os genes *etr1* e *coi1* (*jasmonic acid insensitive*) indicam que tanto o etileno quanto o ácido jasmônico são importantes nesta rota, uma vez que plantas sob ISR não apresentam indução do gene da defensina (Penninckx et al., 1998). Acredita-se também que a ação sinérgica destes dois fitormônios induza a expressão do fator de transcrição ORA59 da superfamília AP2/ERF, que por sua vez é um regulador positivo da expressão de PDF1.2 (Pre et al., 2008).

Dados obtidos em estudos similares com mutantes de soja (*Glycine max*) insensíveis à etileno apresentaram dados contrastantes quando sujeitos à infecções por uma variedade de patógenos. Em algumas situações apresentaram uma ligeira vantagem com

relação ao estabelecimento destas infecções, enquanto que em outras conferiram grande suscetibilidade. Estes dados sugerem que este fitormônio possui diferentes efeitos dependentes do tipo de patógeno e conseqüentemente do mecanismo de infecção (Hoffman et al., 1999).

A expressão de ERF1, induzida por etileno e/ou ácido jasmônico, também é importante em respostas à estresses abióticos. A transcrição de TaERF1 de trigo (*Triticum aestivum*) é induzida por alta salinidade, seca e baixas temperaturas (Xu et al., 2007). Além disso, a especificidade desse fator de transcrição no reconhecimento de elementos *cis*-atuantes em promotores de genes alvo varia de acordo com o tipo de estresse em que a planta é submetida. Ensaio de imunoprecipitação de cromatina (ChIP) com ERF1 de *A. thaliana* mostraram que este fator de transcrição aumentou os níveis de expressão de um conjunto de genes, em resposta à diferentes tipos de estresses abióticos, por interação específica com elementos *DRE*. Já em resposta à estresses bióticos, ERF1 interagiu somente com elementos do tipo *GCC-box*, o que evidencia um reconhecimento estresse específico de elementos *cis*-atuantes em promotores (Cheng et al., 2013).

O papel antagônico entre ABA etileno em resposta à déficit hídrico tem sido relatado. Acredita-se que a biossíntese de ABA, em raízes de *A. thaliana*, reduza a biossíntese de etileno e, conseqüentemente, o crescimento tanto da parte aérea quanto o das raízes (Spollen et al., 2000). Outro aspecto a ser considerado é de que o estresse hídrico é frequentemente acompanhado por elevação de temperatura, o que pode promover a produção de etileno até o limite de 35 °C, momento em que a produção do fitormônio é inibida. Os efeitos observados em altas temperaturas podem ser devido à influência sobre as enzimas de biossíntese, principalmente ACO (Lurie et al., 1996; Morgan & Drew, 1997).

### **1.3. CROSSTALK ENTRE ROTAS SINALIZADAS POR FITORMÔNIOS.**

O crescimento e desenvolvimento vegetal são regulados por uma complexa rede de interações hormonais, onde as auxinas encontram-se como um fator comum na maioria destas interações. Auxinas e citocininas apresentam tanto ações sinérgicas quanto antagonistas para o desenvolvimento do broto e de raízes, respectivamente (Swarup et al.,

2002; Dello Ioio et al., 2008). Outros fitormônios possuem interações clássicas e complexas na regulação de vários processos do desenvolvimento, como a formação do gancho apical, alongamento e gravitropismo de raízes e broto, crescimento lateral radicular, fototropismo do hipocótilo e abscisão foliar (Muday & Rahman, 2008). Foi demonstrado que giberelinas e auxinas coordenam o alongamento do caule e paternocarpia em ervilhas (Swarup et al., 2002), crescimento lateral de raízes em *Populus* (Gou et al., 2010), dominância apical, inflorescência e desenvolvimento radicular em *A. thaliana* (Silverstone et al., 1997; Willige et al., 2011). Estudos ainda sugerem que o ácido abscísico (ABA) e auxinas interagem na regulação do desenvolvimento lateral de raízes e na formação do eixo embriogênico (Belin et al., 2009; Shkolnik-Inbar & Bar-Zvi, 2010).

Estas interações podem ser reguladas em nível biossintético, como *auxinas-etileno* (Muday & Rahman, 2008) ou *auxinas-giberelinas* (Chandler, 2009), ou em nível de sinalização, como *auxinas-citocininas* (Moubayidin et al., 2009), sendo todas essas rotas interligadas pelo transporte de auxinas. Por exemplo, etileno aumenta o transporte de auxinas em zonas de alongamento radicular por estimulação de AUX1, PIN1, PIN2 e PIN4, levando a uma redução do crescimento das raízes e do alongamento celular (Ruzicka et al., 2007). Por outro lado, em regiões maduras das raízes, o etileno inibe a expressão de AUX1, mas estimula PIN3 e PIN7, resultando em um bloqueio da formação do gradiente de auxinas requerido para o desenvolvimento lateral deste órgão (Lewis et al., 2011). Em *A. thaliana*, altos níveis de expressão dos genes WIE2, WIE7 e WIE8 (subunidades da *antranilato sintase*, enzima da biossíntese de triptofano), induzidos por etileno, estão associados com o acúmulo de auxinas, fato este que não ocorre em mutantes insensíveis ao gás (Stepanova et al., 2005; Ruzicka et al., 2007; Swarup et al., 2007). O etileno induz um aumento na curvatura apical, que também é atribuído a um aumento dos níveis de auxinas, modulado pelo reforço na expressão de PIN3 e AUX1 (Muday & Rahman, 2008).

As citocininas modulam o gradiente celular de auxinas em raízes pela regulação negativa da expressão de PIN1, PIN2 e PIN3 e positiva de PIN7 (Ruzicka et al., 2009). Na expansão do hipocótilo e em cultura de células de tabaco, as citocininas regulam negativamente o efluxo de auxinas (Pernisova et al., 2009).

Em *Populus*, as giberelinas estimulam o transporte de auxinas pela indução da expressão de PIN9, o que também está relacionado com o desenvolvimento lateral radicular desta planta (Bjorklund et al., 2007; Gou et al., 2010). Em *A. thaliana*, a

deficiência na produção de giberelinas reduz a expressão de PIN1 e PIN2, além de direcionar a proteína PIN2 para degradação vacuolar, o que afeta a diferenciação dos cotilédones e o gravitropismo radicular (Willige et al., 2011).

Sabendo que hipocótilos e raízes de mutantes de *A. thaliana* para os genes *aux1* e *pin2* são insensíveis à repressão dependente de ácido abscísico (Belin et al., 2009) e que a expressão de PIN1 é alterada em mutantes *abi4* (*abscisic acid insensitive 4*), sugere-se que o fitormônio ABA é importante na modulação do transporte de auxinas (Shkolnik-Inbar & Bar-Zvi, 2010).

Todos estes estudos indicam que o transporte intracelular e o gradiente de auxinas possuem um papel primordial e central no controle de diversas rotas envolvendo as principais classes de fitormônios, regulando assim o desenvolvimento e crescimento vegetal.

O modelo de ação das giberelinas em plantas ainda está longe de ser completamente compreendido devido às diversas interações possíveis com fatores endógenos e ambientais, direta ou indiretamente correlacionadas com a atividade de outros fitormônios (Weiss & Ori, 2007).

As auxinas afetam principalmente a biossíntese de giberelinas, mas já existem estudos em *A. thaliana* que relatam também o seu efeito em rotas de sinalização destes diterpenóides, uma vez que IAA promove a desestabilização de proteínas DELLA (Ross et al., 2000; Wolbang & Ross, 2001; Fu & Harberd, 2003; Frigerio et al., 2006; O'Neill et al., 2010).

Giberelinas e citocininas exercem efeitos antagônicos em numerosos processos, incluindo o alongamento de raízes e brotos, diferenciação celular, regeneração em cultura de tecidos e atividade meristemática (Greenboim-Wainberg et al., 2005; Jasinski et al., 2005). Em *A. thaliana*, dois componentes principais controlam o balanço entre estas duas classes de hormônios vegetais: (i) as proteínas KNOXI (*proteína homeobox de classe I semelhante a Knotted 1*), importantes fatores no desenvolvimento apical de brotos por serem capaz de induzir a produção de citocininas, além de inibirem diretamente a biossíntese de giberelinas e promover indiretamente a sua inativação; (ii) a proteína SPY (*spindly*), que suprime a sinalização desencadeada por GAs promovendo a resposta desencadeada por citocininas (Chen et al., 2004; Brenner et al., 2005; Silverstone et al., 2007).



O ácido abscísico e as giberelinas possuem papéis antagonistas em diversos eventos do desenvolvimento em *A. thaliana*. Enquanto GA está associada com a promoção do crescimento, germinação e floração, o ABA inibe estes processos pela supressão de vias dependentes e independentes de proteínas DELLA (Achard et al., 2004; Achard et al., 2006; Xie et al., 2006).

A interação entre giberelinas e o gás etileno é bastante complexa, uma vez que efeitos positivos e negativos de regulação são demonstrados. Sabe-se que em *A. thaliana*, pelo menos parte do efeito inibitório do etileno no crescimento e sua interação com GAs é mediada por estabilização de proteínas DELLA (Achard et al., 2003). Em contrapartida, GAs promovem respostas mediadas por etileno em crescimento de brotos (formação do gancho apical no escuro e alongamento do hipocótilo na presença de luz), além de que, em condições de alagamento, a biossíntese de ambos os fitormônios em arroz é estimulada, promovendo o alongamento dos internódulos (Vriezen et al., 2004; Benschop et al., 2006; Achard et al., 2007).

A expressão de elementos importantes para a homeostase sinalizada por CKs também é regulada por outros fitormônios como auxinas (Hwang et al., 2012), ácido abscísico (ABA) e etileno. Em raízes de *A. thaliana*, o acúmulo dos transcritos de AtIPT5 e AtIPT7 (importantes enzimas da biossíntese de citocininas) é promovida por auxinas, enquanto que os níveis de expressão dos transcritos de AtIPT1, AtIPT3, AtIPT5 e AtIPT7 são negativamente regulados por CKs (Miyawaki et al., 2004). Por outro lado, a expressão dos genes CYP735A1 e CYP735A2 em raízes é elevada por CKs, mas é reduzida por auxinas ou ABA (Takei et al., 2004). A expressão dos genes CKX (*CK oxidase/desidrogenase*, importante no catabolismo de citocininas) em *Zea mays* é aumentada por CKs e ABA (Brugiere et al., 2003). Estes padrões de regulação sugerem que estas enzimas regulam antagonicamente os níveis celulares de CKs.

Citocininas possuem um profundo efeito no desenvolvimento de plântulas de *A. thaliana*. Benziladenina (BA) inibe o alongamento radicular (influenciado pelo claro ou escuro) e do hipocótilo (somente no escuro), intensificando a curvatura do gancho apical. Estas últimas são respostas características de etileno e, conseqüentemente comprovam o possível envolvimento deste fitormônio no crescimento desta plântula. Foi observado que os efeitos inibitórios de BA no alongamento de raízes e do hipocótilo foram parcialmente bloqueados pela ação dos inibidores de etileno ou em mutantes resistentes a este

fitormônio (*ein1-1* e *ein2-2*). A produção deste gás é estimulada por baixas concentrações de BA e poderia explicar, pelo menos em parte, a inibição do alongamento de raízes e do hipocótilo. Foi demonstrado também que BA não afeta a sensibilidade das plântulas ao gás. Assim, o efeito das citocininas no alongamento de raízes e do hipocótilo em *A. thaliana* parece ser em grande parte mediado pela produção de etileno. O acoplamento entre as repostas mediadas por citocininas e etileno é ainda apoiada pela descoberta de que mutantes de *A. thaliana* tolerantes à citocininas (*ckr1*) também são tolerantes e alélicos à mutantes *ein2* (Cary et al., 1995).

A maior dificuldade na compreensão das interações hormonais em plantas reside no fato de que as rotas de biossíntese e sinalização estão acopladas. Isto é ainda mais complicado pela capacidade da maioria dos tecidos vegetais serem capazes de sintetizar hormônios e a planta por um todo ser capaz de redistribuí-los local e sistemicamente. Um exemplo bem estudado é a conexão existente entre as rotas de biossíntese de auxinas, etileno e ácido abscísico, compreendida em larga escala pela dosagem de herbicidas e análogos de auxinas (Grossmann, 2000). Em baixas concentrações, auxinas estimulam a divisão e o alongamento celular, mas em altas concentrações reduzem a taxa de crescimento celular, abertura estomática, transpiração e assimilação de carbono. Estas mudanças são seguidas por uma rápida senescência foliar e morte da planta (Grossmann, 2000). Entre os primeiros efeitos da ação das auxinas está em estimular a biossíntese de etileno por aumentar a expressão dos genes que codificam as ACS (*1-aminociclopropano-1-ácido carboxílico sintase*), enzimas limitantes da biossíntese de etileno (Abel et al., 1995; Abel & Theologis, 1996). A expressão de ACSs dependente de auxinas é seguida por um pico de liberação de etileno, que por sua vez é acompanhado por um aumento da síntese de ABA devido à produção de H<sub>2</sub>O<sub>2</sub> (Hansen & Grossmann, 2000; Grossmann et al., 2001). Inibidores da ACC oxidase, enzima que catalisa a conversão de ACC (*1-aminociclopropano-1-ácido carboxílico*) em etileno, inibem a produção de ABA, o que culmina na conclusão de que etileno pode induzir a síntese de ABA (Hansen & Grossmann, 2000). Os resultados de análises da rota biossintética do ácido abscísico sugerem que o etileno eleva os níveis de ABA elevando a atividade de enzimas que clivam epoxicarotenóides em xantoxina, aumentando assim a ocorrência de eventos de retroalimentação positiva (Hansen & Grossmann, 2000).

Interconecções entre as rotas de sinalização mediadas por etileno e ABA vieram à tona por meio de análises de mutações que afetam ambas as respostas hormonais. As mutações que tornam plantas insensíveis à etileno (*etr1* e *ein2*) aumentam a sensibilidade à ABA durante a germinação, além de intensificarem a dormência, enquanto que mutações no gene *ctr1* diminuem esta sensibilidade ao ácido abscísico e diminuem levemente a dormência (Beaudoin et al., 2000; Ghassemian et al., 2000). Mutantes nocaute para os genes *ein2* e *ctr1* foram identificados como supressores e potenciadores respectivamente da mutação *abil-1* (insensível à ABA) (Beaudoin et al., 2000). Este fato é consistente com a interpretação de que a sinalização mediada por etileno é negativamente regulada pela MAPKKK CTR1 (Kieber et al., 1993). Além disso, estudos sugerem que a sinalização mediada pelo fitormônio gasoso inibe respostas em sementes, induzidas por fatores de transcrição específicos da sinalização mediada por ABA. Estes fatores de transcrição estão ativos no estabelecimento da dormência e são degradados logo após a germinação, a menos que sejam estabilizados pela sinalização mediada por ácido abscísico (Arc et al., 2013). Em *A. thaliana*, a sensibilidade à ABA durante a germinação é aumentada pelo aumento da expressão da proteína AtABI5 e AtMPK3, assim como pela mutação *hyl11*, que aumenta a quantidade de MAPKs ativas (Lu & Fedoroff, 2000). Contudo, a relação entre a sinalização mediada por etileno e a mediada por ABA pode ser diferente em tecidos vegetativos distintos das sementes, como é sugerido pelo fato da biossíntese de etileno ao desencadear a biossíntese de ABA, além das respostas fisiológicas associadas com a sinalização mediada por ácido abscísico. Mutações no gene *ein2* não alteram a regulação da abertura estomática mediada por ABA ou a resposta de vários genes induzidos por este fitormônio (como RD29A e RAB18) em tecidos vegetativos, mas estudos com mutantes nocaute para o gene *eto1*, importante na biossíntese de etileno, mostram que a produção deste gás pode sim inibir o fechamento estomático induzido por ácido abscísico (Beaudoin et al., 2000; Tanaka et al., 2005). Além disso, o crescimento de raízes de mutantes nocaute para os genes *ein2* e *etr1* são tão insensíveis à ABA quantos os mutantes *abil* (Beaudoin et al., 2000). O fato do precursor de etileno, ACC, ser capaz de restaurar a sensibilidade de raízes de mutantes *abil* à ABA, mas não em mutantes *ein2* ou *etr1*, sugere que a sinalização mediada por etileno atua a jusante da mutação *abil* para potencializar a resposta vegetativa ao ácido abscísico.

Sendo assim, o sinergismo e/ou antagonismo coordenados entre a produção e a sinalização mediada pelos diversos fitormônios possuem um papel primordial no desenvolvimento e na adaptação das plantas nas mais diversas condições ambientais.

#### **1.4. ELEMENTOS REGULADORES EM PROMOTORES GÊNICOS.**

O rápido progresso em pesquisas relacionadas com os genomas de plantas, como *A. thaliana*, *Oryza sativa* e até *Glycine max*, geraram muitas sequências de nucleotídeos no decorrer dos anos. A maior parte delas é considerada como *não codante*, podendo atuar em diversos processos celulares que não sejam relacionados diretamente com a produção de proteínas, dentre eles, diversos processos regulatórios. A expressão de um gene é controlada pela interação de proteínas (denominadas de fatores de transcrição) com elementos *cis*-atuantes presentes no DNA, representados principalmente por promotores, *enhancers* (ou potenciadores) e silenciadores (ou atenuadores).

Dentre as famílias de elementos *cis*-atuantes identificadas em plantas, duas compõem a maioria dos *promotores mínimos* ou *basais* (menor sequência de DNA suficiente para o início da transcrição, mediada pela maquinaria de transcrição associada à RNA polimerase) de genes vegetais, sendo elas *PTBP* (*TATA-box* de plantas) e *CAAT* (elementos *CCAAT-box*) (Dai et al., 1996). A família *PTBP* está envolvida na formação no complexo de iniciação da transcrição, e encontra-se próxima ao sítio de início da transcrição *TSS* (*transcription start site*). Já elementos *cis*-atuantes da família *CAAT* possuem aproximadamente 70 nucleotídeos, contendo a sequência consenso GGCCAATCT, localizando-se 75 a 80 pares de bases antes do sítio de início da transcrição e, semelhante a família *PTBP*, também são conhecidos por sinalizar o sítio de início da transcrição. Em plantas, estudos mostram também a participação de elementos *CAAT* na regulação da transcrição de diversos genes. A expressão do gene *AtpC* de espinafre, que codifica a subunidade  $\gamma$  da ATP sintetase do cloroplasto, é regulada por fitormônios e luminosidade, através do fator de transcrição que interage com elementos *CAAT* (*CAAT binding fator - CBF*) (Kusnetsov et al., 1999). Haralampidis e colaboradores (2002) mostraram que a deleção de elementos *CAAT* em promotores de *A. thaliana* diminui de forma significativa a atividade do promotor do gene referente à proteína HSP90-1, que é

conhecida por ser expressa em situações de estresses abióticos (como calor) e metais pesados (Haralampidis et al., 2002). Já Lopes-Caitar e colaboradores (2013) identificaram numerosos elementos *CAAT* na família de genes GmHsp20 de soja, associados a diversos tipos de estresses abióticos (Lopes-Caitar et al., 2013).

Estudos genéticos e moleculares em *A. thaliana* mostram que fatores de transcrição do tipo bZIP (*AtbZIP*) regulam diversos processos biológicos, tais como defesa contra patógenos, sinalização mediada pela luz e por estresses, além do desenvolvimento das sementes e flores (Dai et al., 1996). Elementos *cis*-atuantes da família *GBOX*, encontrados em vários promotores de genes de planta, são conhecidos por interagir com fatores de transcrição do tipo bZIP, mediando respostas à diferentes estímulos. A maioria dos elementos desta família possuem sítios palindrômicos (CCACGTGG), sendo o núcleo (*core*) ACGT o responsável por interagir com vários fatores bZIP. Estudos mostram que as sequências que flanqueiam o núcleo dos elementos desta família afetam a especificidade dos fatores de transcrição que o reconhecem, podendo gerar três outras possíveis subfamílias: *ABOX*, *CBOX* e *TBOX* (Schindler et al., 1992; Izawa et al., 1993). Estudos realizados por Mehrotra e colaboradores (2013), comparando os padrões e a evolução do *core* ACGT em *O. sativa*, *Sorghum bicolor*, *Arabidopsis thaliana* e *Glycine max*, sugerem que a família de elementos mais conservada entre as espécies é a *GBOX*, sendo responsável pela regulação de estresses exógenos, principalmente estresses hídrico e salino (Mehrotra et al., 2013). Outra importante família de elementos *cis*-atuantes relacionada com respostas associadas com bZIP seria a denominada de *DPBF*, que é reconhecida pelos fatores de transcrição DPBF1 e DPBF2, que por sua vez reconhecem sequências contendo o *core* ACACNNG. Estudos mostraram que a interação desses fatores com promotores *Dc3* de cenoura, ocorre possivelmente em forma de homodímeros, sendo importante na expressão do referido gene na embriogênese somática e zigótica, podendo ser induzido por ABA (Kim et al., 1997).

A família de fatores de transcrição MYB é ampla, funcionalmente diversa e representada em todos os organismos eucarióticos. A maioria dos membros desta família de fatores de transcrição atua com um variado número de repetições de domínios específicos, conferindo-lhes a habilidade de interagir com DNA. Em plantas, a família de fatores MYB expandiu seletivamente através da família R2R3-MYB e participa de uma vasta rede metabólica, controlando o metabolismo, desenvolvimento e diversas respostas à

estresses bióticos e abióticos (Dubos et al., 2010). Os R2R3-MYB são associados principalmente com a regulação da síntese de diferentes fenilpropanóides, assim como na sinalização mediada por ácido salicílico, ABA e giberelinas. Estes fatores de transcrição interagem especificamente com certas famílias de elementos *cis*-atuantes chamadas de I, II e IIG, sendo a família *MIIG* representada pela sequência consenso GKTWGGTR (onde K=G ou T; R=A ou G; W=A ou T) (Romero et al., 1998). Além da família *MIIG*, fatores de transcrição da família MYB interagem com diversas outras famílias de elementos *cis*-atuantes, dentre as quais destacam-se principalmente *MYBL*, *MYBS*, *GARP* e *CDC5*. Os elementos *MYBL* interagem com fatores de transcrição do tipo MYB, tais como o elemento repressor AS1/AS2, vários R2R3-MYB, elementos CAACTC reguladores, elementos de resposta à giberelinas, dentre outros (Gubler & Jacobsen, 1992; Sutoh & Yamauchi, 2003; Guo et al., 2008; Li et al., 2009). Já os elementos *MYBS* interagem com proteínas de tipo MYB com repetição única do domínio de interação com DNA. Diversas proteínas foram descritas na literatura como mediadoras da resposta à carboidratos e giberelinas, regulando a expressão de genes da  $\alpha$ -amilase, que possuem em seus promotores o elemento *cis*-atuante *amilase-box* (TATCCA) (Gubler & Jacobsen, 1992; Lanahan et al., 1992; Toyofuku et al., 1998; Lu et al., 2002). Os fatores de transcrição GARP são proteínas moduladas pela sinalização *His/Asp phosphorelay*, principalmente em resposta a hormônios como etileno e citocininas. Estas proteínas apresentam um motivo MYB, chamado de *motivo B*, com estrutura e função pouco conhecidas (Hosoda et al., 2002). O gene *AtCDC5* de *A. thaliana* codifica uma proteína importante na regulação do ciclo celular em plantas, com alta similaridade com proteínas relacionadas à MYB, similar a *CDC5* de *Schizosaccharomyces pombe* (Hirayama & Shinozaki, 1996).

Fatores de transcrição com domínios básicos de interação com DNA do tipo hélice-volta-hélice (*basic helix-loop-helix; bHLH*), onde se destacam os membros da família MYC, são de extrema importância na regulação de diversos eventos em células vegetais (Heim et al., 2003). Dentre estes eventos, Zhu e colaboradores (2003) mostram em arroz a interação entre uma proteína MYC (OsBP5) e uma proteína da família EREBP (OsEBP89) regulando a expressão do gene *WX*, destacando desta forma o sinergismo entre as rotas de sinalização mediadas por etileno e ácido jasmônico, além da complexidade das mesmas em resposta a diferentes tipos de estímulo em plantas (Zhu et al., 2003). Boter e colaboradores (2004) isolaram dois fatores de transcrição MYC de tomate (JAMYC2 e JAMYC10) que

interagem com elementos de resposta ao ácido jasmônico em promotores de genes responsivos à injúrias mecânicas, reconhecendo elementos *T/G-box* (Boter et al., 2004). Estudos com tabaco sugerem que a sequência consenso de interação do fator NtMYC2 seria CA(C/T)(G/T)T(G/T), sendo CACGTG a preferencial (Shoji & Hashimoto, 2011), pertencente à família de elementos *cis*-atuantes *MYCL*. Por fim, Wang e colaboradores (2014) identificaram sete genes semelhantes à GmMYC2 de soja, entre os quais GmMYC1 que estimula mecanismos de defesa em resposta à infecções causadas por *Spodoptera litura* (Wang et al., 2014).

A família de fatores de transcrição NAC, específica de plantas, possui como principais membros os fatores NAM, ATAF1-2 e CUC2 que participam de diversos processos celulares como desenvolvimento embrionário floral e vegetativo, além de resposta a diversos tipos de estresses. Estudos mostram que os fatores de transcrição NAC1, AtNAM e ANAC019 de *A. thaliana* interagem com o promotor *CaMV35S*, assim como alguns ANAC reconhecem elementos específicos presentes no promotor do gene ERD1 (*early responsive to dehydration stress 1*). O principal elemento *cis*-atuante reconhecido por estes fatores de transcrição pertence à família *NACF* e possui a sequência CAGG como núcleo principal (Olsen et al., 2005). Em soja, estudos mostraram que membros da família NAC (GmNAC019, GmNAC043, GmNAC062, GmNAC085, GmNAC095 e GmNAC101) são induzidos por déficit hídrico, assim como por injúrias, alta salinidade, baixas temperaturas, mas não por ácido abscísico (GmNAC5) (Jin et al., 2013; Thu et al., 2014). Além desta família de elementos *cis*-atuantes, existem diversas outras que respondem aos fatores de transcrição da família NAC, dentre as quais destacam-se as famílias *ATAF*, *CNAC* e *SWNS*. O fator de transcrição ATAF2 é conhecido como um depressor de genes relacionados à patogênese, assim como está associado com respostas sinalizadas por hormônios, infecção por patógenos e injúrias. Estudos mostraram que ATAF2 interage diretamente com elementos no promotor do gene NIT2, importante na biossíntese de auxinas e na indução de respostas mediadas por este fitormônio (Huh et al., 2012). Já os fatores de transcrição do tipo CBNAC (proteína NAC de interação com calmodulina) são conhecidos por regular a expressão de diversos genes em protoplastos de *A. thaliana*, em resposta à variações na concentração de  $Ca^{2+}$ , interagindo com elementos *cis*-atuantes da família *CNAC*, como por exemplo *CBNACBS* (Kim et al., 2007). Por fim, fatores de transcrição que interagem com elementos família *SWNS* (*secondary wall NAC*

*binding elements*) são reguladores da biossíntese da parede celular secundária em plantas. Fatores desta classe em *A. thaliana* (SND1, VND6, VND7, NST1 e NST2) interagem com a sequência palindrômica imperfeita *SNBE* presente em promotores de genes alvo (Zhong et al., 2010).

Fitormônios são moléculas muito importantes no processo de regulação da transcrição. Com base neste fato, podem-se destacar algumas famílias de elementos *cis*-atuantes, identificadas em promotores de diversas plantas, que respondam direta ou indiretamente à vias de sinalização mediadas por estas moléculas. Dentre elas, destacam-se as famílias de elementos responsivos à ácido abscísico, auxinas, brassinoesteróides, etileno e ácido jasmônico. Com relação às respostas mediadas por ABA, estudos demonstram que elementos *cis*-atuantes pertencentes à família *ABRE* (*abscisic acid responsive elements*) foram identificados em regiões promotoras de genes induzidos por este fitormônio, sendo representados pela sequência consenso PyACGTGG/TC. A presença destes elementos foi correlacionada também com a regulação de diversos processos vegetais, como o fechamento dos estômatos, dormência de sementes e botões florais e respostas fisiológicas à adversidade ambientais, tais como frio, seca e alta salinidade. Vários trabalhos revelam que a expressão de genes dependentes de ABA requer mais de uma sequência, ou uma combinação de elementos da família *ABRE* para compor um promotor funcional. Em *A. thaliana*, essa expressão é controlada por fatores de transcrição da família AREB/ABF, fatores de tipo bZIP que interagem com estes elementos (Maruyama et al., 2012; Fujita et al., 2013; Nakashima & Yamaguchi-Shinozaki, 2013). Embora a importância funcional dos elementos *ABRE* seja de extrema relevância para a transcrição dependente de ABA, promotores contendo estes elementos têm sido caracterizados em processos regulatórios mediados por outros tipos de sinais, tais como progressão do ciclo celular, intensidade e qualidade luminosa e altos níveis de ácido jasmônico. Para isso, diferentes tipos de elementos flanqueiam elementos da família *ABRE*, compondo complexos funcionais regulatórios. Dentre os elementos que atuam em associação com membros da família *ABRE* destacam-se principalmente os elementos *CE1*, *CE3* e *CE3-like* (cevada), componentes da família *CE1F*, comumente chamada de *G/ABRE* e *C/ABRE* (Hobo et al., 1999). O primeiro fator de transcrição descrito por interagir diretamente com um elemento *CE1* foi o fator *ZmABI4* de milho, que interage com a sequência CACCG. Este fator de transcrição regula a expressão de genes responsivos à ABA e à carboidratos,



principalmente durante o desenvolvimento das sementes (Niu et al., 2002). É importante destacar que estudos mostraram que fatores da família ERF de resposta à etileno, como o TSRF1 de tomate, interagem fisicamente com elementos de tipo *CEI* ao mesmo tempo que promovem a expressão de genes em resposta ao ABA (Zhang et al., 2008). Da mesma forma, o fator de transcrição JERF3, também da família ERF, modula os elementos *cis*-atuantes *DRE* (elemento responsivo à seca) e *CEI*, aumentando a tolerância a estresse osmótico, seca e frio, quando expresso em plantas transgênicas de tabaco (Wu et al., 2008).

Já elementos da família *AREF* (elementos de resposta à auxina) são importantes componentes da via de regulação da transcrição mediada por auxinas. Esta família de elementos *cis*-atuantes é composta principalmente por dois elementos: *ARE* (ou *AuxRE* - elemento responsivo à auxina) e *SEBF* (elemento de silenciamento - repressores transcricionais). Os promotores dos genes de soja GmGH3 e GmSAUR15A foram caracterizados em vários estudos, mostrando uma rápida ativação em resposta à auxinas, sendo que os promotores destes genes apresentaram elementos *cis*-atuantes do tipo *AuxRE*, elementos estes que pode atuar em cópia única ou repetições, com sequências flangeadoras dispostas de diversas formas (sobrepostas, invertidas ou adjacentes) (Guilfoyle et al., 1998). Antagonicamente à *AuxRE*, a literatura relata que elementos *cis*-atuantes do tipo *SEBF* regulam negativamente a expressão de genes, apesar de apresentar similaridade com o elemento *AuxRE*. O gene PR-10a de batata, relacionado com patogênese é ativado em resposta à infecção por patógenos ou tratamento elicitor. A presença de elementos *SEBF* na região promotora do gene PR1-10a foi caracterizada como sendo responsável pelo silenciamento do referido gene e consequente redução das respostas à patógenos das quais o produto deste gene participa (Boyle & Brisson, 2001).

Os brassinoesteróides são hormônios que desempenham papéis essenciais no crescimento e desenvolvimento nas plantas. A homeostase desses fitormônios dentro das células deve ser monitorada de forma estreita, pois além de possuírem atividade no mesmo tecido ou célula em que são sintetizados, eles não são transportados à longas distâncias. Desse modo, a biossíntese dos brassinoesteróides é autorregulada pela sinalização mediada por eles mesmos, por condições ambientais e possivelmente por outros fitormônios. Desta forma, o fator de transcrição BZR1, induzido por altos níveis deste fitormônio, interage com o elemento *cis*-atuante da família *BRRE*, presente no promotor do gene CPD,

importante na biossíntese de brassinoesteróides, atuando como repressor e retroalimentador negativo da rota (He et al., 2005).

Os fatores de transcrição EIN3 (*ethylene insensitive 3*) e EIL (semelhante à EIN3) atuam na via de sinalização mediada por etileno em diversas espécies vegetais, mas possui sua função amplamente descrita em *A. thaliana*. Estudos realizados sobre o homólogo de EIN3 em tabaco (TEIL) mostraram que o elemento *cis*-atuante reconhecido por este fator nesta espécie vegetal é A(A/C)G(A/T)A(A/C)CT. Ensaaios com genes repórteres regulados por esse elemento (pertencente à família *EINL*) indicaram que TEIL atua como ativador transcricional em resposta ao etileno (Kosugi & Ohashi, 2000). Diferentes trabalhos mostraram o efeito sinérgico e ao mesmo tempo antagônico entre o etileno e o ácido jasmônico (jasmonato) na defesa de plantas contra fungos necrotróficos. Foi mostrado que EIN3/EIL1 interagem com fatores de transcrição JAZ, importantes em inibir MYC2, o principal fator de transcrição da sinalização mediada por ácido jasmônico. Em presença de etileno e jasmonato, a atividade das EIN3/EIL1 seria maior do que quando comparada com os dois hormônios separados, permitindo uma maior expressão de defensas de plantas. Da mesma forma, os dois hormônios possuem interação antagônica, uma vez que MYC2 também interage com EIN3 regulando-o negativamente. Estudos mostram mais uma vez o quanto a sinalização em plantas pode ser complexa e diversa dependendo de vários fatores tanto ambientais como do desenvolvimento (Zhu, 2014). Efeitos fisiológicos antagônicos foram também apresentados entre os fitormônios ABA e etileno *A. thaliana*, uma vez que Tanaka e colaboradores (2005) mostraram que a sinalização mediada por etileno inibe o fechamento de estômatos, que por sua vez é induzido por ABA (Tanaka et al., 2005).

Estudos com o promotor do gene JAZ2 (*jasmonate SIM domain 2 protein*), que está evoluído na resposta ao ácido jasmônico, permitiram identificar três cópias de um elemento *cis*-atuante semelhante à *G-box* (pertencente à família *JARE*), que confere atividade autônoma e dependente deste fitormônio. Além disso, os elementos responsivos à MYC2, MYC3, e MYC4 *trans*-ativam a transcrição do gene JAZ2. Uma seqüência rica em timidinas que flanqueia a posição 3' do elemento semelhante à *G-box* é necessária para a total ativação do promotor, e por isso, este elemento da família *JARE* seria composto por esta combinação de seqüências, sendo esta ativação provavelmente regulada por intermédio dos fatores MYC2, MYC3 e MYC4 (Figueroa & Browse, 2012).

Uma importante classe de fatores de transcrição à ser considerada em mecanismos de regulação transcricional em plantas é a representada por membros da superfamília AP2/ERF. Estes fatores reconhecem inúmeras famílias de elementos *cis*-atuantes em promotores vegetais, destacando-se as famílias *GCCF* (elementos *GCC-box*), *ERFEF* (elementos responsivos à etileno), *DREB* (proteínas com elementos de interação responsivos à desidratação), *FLO2* (elementos de interação com proteína homeótica floral *APETALA 2*), *RAV3* (porção 3' do sítio de interação com RAV1) e *RAV5* (porção 5' do sítio de interação com RAV1). A família *GCCF* é composta pelos elementos *ERE* (elemento responsivo à etileno) e *JERE* (elemento responsivo à jasmonato e eliciador). A identificação da sequência do elemento *ERE* (TAAGAGCCGCC), denominada também de *GCC-box*, foi determinada em estudos com o gene PRB-1b, cuja função está relacionada com a resposta à ataques por patógenos induzida por etileno, assim como em genes induzidos por ozônio (Sessa et al., 1995; Grimmig et al., 2003). Por outro lado, a indução do gene de biossíntese do metabólito secundário que atua como eliciador em células de *Catharanthus roseus* requer ácido jasmônico como segundo mensageiro. Uma pequena porção da região promotora deste gene é necessária e suficiente para a expressão do mesmo, e após análises de interação deste elemento com o fator de transcrição ORCA2, que possui domínio AP2 em sua estrutura, observou-se a presença de um elemento semelhante à *GCC-box*, denominado posteriormente de *JERE* (Menke et al., 1999).

Já as famílias de fatores de transcrição do tipo ANT e EREBP (proteínas que interagem com elemento responsivo à etileno) estão envolvidas em diferentes processos celulares vegetais, tais como o desenvolvimento, transdução de sinal mediada por fitormônios, além de respostas a estresses bióticos e abióticos. Estas proteínas contêm um (EREBP) ou dois (ANT) domínios AP2, que interagem com o DNA principalmente através dos elementos *cis*-atuante *ERE* ou de outros elementos semelhantes à *GCC-box* (pertencentes à família *ERFEF*), encontrados em genes induzidos pelo etileno em resposta a patógenos. Estas proteínas também interagem com elementos denominados de *c-repeat/DRE*, conhecidos por estarem presentes em promotores de genes importantes em resposta à baixas temperaturas e déficit hídrico. A expressão dos fatores de transcrição GmERF3 e GmERF7 em soja foi induzida em resposta à alta salinidade, seca, ABA, ácido salicílico, ácido jasmônico e etileno. Foi observada a interação destes fatores com elementos do tipo *GCC-box* e *DRE/CRT*, e quando expresso ectopicamente em tabaco,

GmERF3 induz a expressão de genes de resistência à patógenos, além de e conferir maior tolerância à déficit hídrico e alta salinidade, enquanto GmERF7 confere maior tolerância somente à alta salinidade (Zhang et al., 2009; Zhai et al., 2013). Os fatores homeóticos com duas repetições do domínio AP2 denominados de *AINTEGUMENTA* (ANT) interagem com inúmeros elementos específicos (com sequências distintas das reconhecidas pelos demais domínios AP2), como também com elementos *c-repeat/DRE* (Nole-Wilson & Krizek, 2000).

O elemento *cis*-atuante designado como *DRE/CTR* (A/GCCGAC), principal componente da família *DREB* de elementos *cis*-atuantes, foi inicialmente relacionado a expressão de genes independente de ABA. Em *A. thaliana*, DREB1/CBF e DREB2, (além de DDF, HDR, RAP dentre outros) são fatores de transcrição de tipo AP2/ERF, que interagem especificamente com estes elementos, regulando principalmente a expressão de genes em resposta à seca e baixas temperaturas (Liu et al., 1998). A alta expressão de genes DREB/CBF tem sido correlacionada com o aumento da tolerância ao estresse hídrico em diferentes plantas de interesse agrícola, incluindo arroz e soja (Datta et al., 2012; Paiva Rolla et al., 2014). Da mesma forma, ortólogos de DREB/CBF têm sido estudados nas mais diversas espécies agrícolas. Estudos realizado por Maruyama e colaboradores (2013) relatam como a proteína DEAR1 de tipo DREB de *A. thaliana* interage com a região promotora do gene *AtERF9*, estimulando a defesa da planta a um fungo chamado *Botrytis cinerea* (Maruyama et al., 2013). A ação de fatores de transcrição conhecidos comumente por responder a estresses abióticos ativando defesas para estresses bióticos mostra o quanto a sinalização molecular em plantas é complexa e pode resultar de interconecções entre os dois tipos de estresse.

A especificação do destino celular no desenvolvimento vegetal requer fatores de transcrição com a propriedade de regular a expressão de genes. Em *A. thaliana*, quatro classes de fatores de transcrição (A, B, C e E) agem em diferentes combinações com a finalidade de coordenar a identidade floral. Os genes de classe A (*APETALA 2*; AP2), que semelhantes aos fatores *AINTEGUMENTA* possuem dois domínios AP2 de interação com o DNA, promovem a identidade entre sépalas e pétalas, ao mesmo tempo que inibem a expressão do gene *AGAMOUS*, através do reconhecimento de elementos *cis*-atuantes da família *FLO2* (TTTGTT ou AACAAA) em regiões promotoras específicas (Dinh et al., 2012).

Fatores de transcrição da família RAV (relacionado à ABI3/VPI1) são conhecidos por possuírem dois domínios de interação com DNA, AP2 em sua extremidade N-terminal e B3 em seu C-terminal, atuando como repressores transcricionais (regulador negativo do desenvolvimento) em processos como florescimento, senescência e respostas à diferentes tipos de estresses (Matias-Hernandez et al., 2014). Estes domínios interagem com elementos *cis*-atuantes específicos pertencentes às famílias *RAV3* (interação com domínio B3) e *RAV5* (reconhecido pelo domínio AP2), sendo representados respectivamente pelas sequências consenso CACCTG e CAACA. Devido à grande flexibilidade estrutural dos membros desta família de fatores de transcrição, cada um dos dois domínios pode reconhecer regiões regulatórias independentemente, apresentadas com diferentes espaçamentos e orientações (Kagaya et al., 1999).

Dentre as famílias de elementos *cis*-atuantes estudadas em plantas, também destacam-se algumas que respondem à estímulos e/ou estresses abióticos ambientais específicos. A regulação transcricional dos genes de resposta à choque térmico (*heat shock*) tem sido relacionada à presença de diversos elementos *cis*-atuantes. Dentre eles, destacam-se os elementos *HSE* (*heat shock element*), que é composto por repetições de um pentanucleotídeo (nGAAn), *STRE* (*stress responsive element*), e *ARE* (*API responsive element*). O gene *AtHsp90-1* de *A. thaliana*, por exemplo, apresenta em sua região promotora diferentes sequências consenso de tipo *HSE* e *STRE*, assim como *AP-1*, responsáveis pela regulação em resposta à diferentes tipos de estresses, incluindo altas temperaturas (Aranda et al., 1999). Em girassol, a região promotora do gene que codifica uma pequena proteína relacionada à choque térmico (*sHSP*) também contém diferentes elementos *HSE* e *c-repeat*, sugerindo uma possível resposta a estresses abióticos (Haralampidis et al., 2002; Diaz-Martin et al., 2005).

Já elementos responsivos à luz foram caracterizados em estudos com *A. thaliana* onde a resposta à luz monocromática com diferentes qualidades é mediada por um elemento próximo ao *TATA-box* (-300 pb). Da mesma forma, o mesmo estudo demonstrou que regiões distantes do *TATA-box* são responsáveis pela ativação não modulada pela qualidade da luz (família *LREM*). Estas regiões possuem o elemento *cis*-atuante ATCTA repetido *in tandem*, que é decisivo para a descrição deste mecanismo de regulação. Além disso, este elemento foi relacionado com a expressão de genes relacionados à fotossíntese, reforçando sua importância na regulação em resposta à luz (Welsch et al., 2003).

Por fim, a regulação mediada pelo fator de transcrição Alfin1 em alfafa (*Medicago sativa* L.) está associada com a tolerância a alta salinidade, devido à interação com elementos *cis*-atuantes compostos por *triplets* ricos em G, pertencentes à família *SALT* de elementos *cis*-atuantes. Em soja, foram identificados homólogos ao fator Alfin1 (*GmPHD*) que respondem a déficit hídrico, alta salinidade, baixas temperaturas e ABA devido à interação com o elemento GTGGAG (Winicov & Bastola, 1999; Wei et al., 2009).

### **1.5. RESPOSTAS À DÉFICIT HÍDRICO EM PLANTAS.**

Em situações naturais e agricultáveis, as plantas são frequentemente expostas ao estresse ambiental. Alguns fatores dessa natureza, como a temperatura do ar, podem se tornar estressantes em poucos minutos, enquanto outros, como o conteúdo de água do solo, necessitam de dias ou até semanas para se manifestar (Boyer, 1982). Em face da escassez de recursos hídricos, a seca é a ameaça mais crítica para a segurança alimentar mundial, sendo a catalisadora de grandes surtos de fome em um passado não tão distante. Devido ao fato do suprimento mundial de água ser limitado, a demanda futura de alimentos necessária para a manutenção da crescente população mundial é o fator mais prejudicado pelos efeitos da seca (Somerville & Briscoe, 2001). A severidade do déficit hídrico é imprevisível e depende de muitos fatores, tais como ocorrência e distribuição de chuvas, taxas de evaporação e umidade, associadas com capacidade de armazenamento de água pelo solo (Wery et al., 1993).

A capacidade das plantas para atender a demanda de água e, assim, evitar o déficit hídrico depende de sua *maquinaria hidráulica*, que envolve a capacidade de reduzir o *albedo* (razão entre a radiação refletida e a incidida), transportar água suficiente do solo para a atmosfera através do estômato (que assimila CO<sub>2</sub>), com a finalidade de fornecê-la para processos fisiológicos como transpiração, resfriamento e assimilação de carbono. A água é transportada através do sistema SPAC (contínuo solo-planta-atmosfera ou *soil-plant-atmosphere continuum*), que é amplamente controlado por resistências, determinadas pelas raízes, caule, folhas, estômatos e cutícula hidráulica, que por sua vez podem variar de acordo com as respostas da planta e condições ambientais. A força inicial que direciona a água contra as resistências vegetais é o gradiente do potencial hídrico existente entre o solo

e as folhas, que reduz este potencial nas folhas, propiciando o ajuste do potencial osmótico foliar e subsequentemente sua hidratação (Blum, 2011).

As plantas respondem à seca de diversas formas, tanto pelo retardo da desidratação (capacidade de manutenção da hidratação do tecido), pela tolerância à seca (capacidade de sobreviver mesmo desidratadas), ou ainda pelo escape da seca, onde o ciclo de vida da planta é completado durante a estação úmida. Dentre as características fisiológicas e estratégias de retardo da desidratação, observadas em plantas em condições de seca moderada e severa, destacam-se a redução da área foliar, com a finalidade de reduzir a perda de água, podendo até culminar na abscisão, dependendo da intensidade do déficit hídrico; a estimulação do aprofundamento das raízes no solo, uma vez que em regiões mais profundas permanecem úmidas por mais tempo; o fechamento estomático induzido por ácido abscísico, com o objetivo de reduzir a transpiração; a limitação da fotossíntese, ocasionada pela redução da condutância estomática ( $g_s$ ) e posterior desidratação do mesófilo; a redução do potencial osmótico; o aumento no depósito de cera na superfície foliar e na camada interna da cutícula; e o aumento da temperatura da folha pelo acúmulo de energia não dissipada (Levitt, 1980).

### **1.5.1. Fitormônios e Tolerância à Estresse Hídrico.**

Fitormônios são essenciais na habilidade das plantas em se adaptarem a uma variedade de estresses abióticos por mediar uma grande variedade de respostas adaptativas (Santner & Estelle, 2009). Eles alteram rapidamente a expressão gênica por indução ou prevenção da degradação de reguladores transcricionais (Santner & Estelle, 2010).

Estudos em arroz (*O. sativa*) mostram o papel das auxinas na tolerância à seca, uma vez que a expressão da enzima *IAA-amido sintase* (TLD1/OsGH3.13), importante na conjugação de IAA com amido, induz o aumento da expressão do gene *LEA* (*late embryogenesis abundant*), que por sua vez está correlacionado com o aumento da tolerância à seca em mudas de arroz (Zhang et al., 2009). A expressão de muitos outros genes associados com a biossíntese, percepção e ação de auxinas pode ser regulada por etileno (Stepanova & Alonso, 2009). Dentre eles, destacam-se os fatores responsivos à

auxinas ARF2 e ARF19 (Li et al., 2004; Li et al., 2006), transportadores do fitormônio PIN1, PIN2, PIN3, PIN4, PIN7 e AUX1 (Ruzicka et al., 2007; Lewis et al., 2011), além de enzimas relacionadas biossíntese de IAA (WEI2, WEI7 e WEI8) (Stepanova et al., 2008). Reciprocamente, as auxinas influenciam também a produção de etileno em resposta à estresse hídrico, uma vez que as enzimas ACS sofrem influências de IAA (Tsuchisaka et al., 2009). Além disso, estudos mais recentes demonstram o papel protetor da auxina endógena e exógena na tolerância à seca em *A. thaliana*, podendo participar na regulação positiva deste tipo de adversidade ambiental, através da modulação da arquitetura radicular, da regulação de genes responsivos à ABA, do metabolismo de espécies reativas de oxigênio e homeostase metabólica, mesmo que parcialmente (Shi et al., 2014).

O ácido abscísico é um dos fitormônios mais estudados em respostas à diversos tipos de estresses abióticos. Este fitormônio regula a expressão de genes específicos em uma diversidade de adaptações fisiológicas relacionadas com déficit hídrico, dentre as quais se destacam o ajuste osmótico, a compartimentalização iônica, a condutividade hidráulica em raízes, a indução do fechamento de estômatos, a limitação da transpiração e do processo de murcha, além da regulação do crescimento de brotos e raízes (Verslues & Zhu, 2005; Pospisilova et al., 2009). Embora muitos genes sejam induzidos por adição exógena de ABA em resposta à seca e à baixas temperaturas, existem muitos outros que não respondem a tal tratamento (Zhu, 2002; Shinozaki et al., 2003; Yamaguchi-Shinozaki & Shinozaki, 2005). Com base nestes dados, é proposto que existam duas possíveis rotas de sinalização, uma dependente e outra independente de ácido abscísico. Estudos apontam pelo menos dois sistemas na sinalização dependente de ABA: (i) AREB/ABF (*ABA responsive element binding protein/ABA binding factor*); e (ii) MYC/MYB (Abe et al., 1997; Busk & Pages, 1998). Por outro lado, os sistemas independentes de ácido abscísico são: (i) CBF/DREB (*cold binding factor/ drought response element binding protein*); e (ii) NAC e ZF-HD (*zinc finger homeobox family protein*) (Saibo et al., 2009). Além disso, estudos prévios identificaram a existência de rotas que são ao mesmo tempo dependentes e independentes de ABA, trazendo evidências de interconexões entre as duas vias de sinalização (Yamaguchi-Shinozaki & Shinozaki, 1994; Kizis & Pages, 2002; Lata et al., 2011).

As rotas de sinalização independentes de ácido abscísico são reguladas principalmente por dois membros da superfamília AP2/ERF de fatores de transcrição,



denominados de DREB1 e DREB2 (Gilmour et al., 1998; Liu et al., 1998). A expressão dos seis representantes da subfamília DREB1 de *A. thaliana* é fortemente e transientemente induzida em condições de baixas temperaturas, enquanto que os oito representantes da subfamília DREB2 (principalmente DREB2A e DREB2B) são induzidos por déficit hídrico, alta salinidade e altas temperaturas (Shinwari et al., 1998; Nakashima & Yamaguchi-Shinozaki, 2010).

Tendo em vista as vias de resposta à déficit hídrico, DREB2A e DREB2B foram os primeiros fatores de transcrição isolados em *A. thaliana* relacionados diretamente com este tipo de resposta, sendo que genes ortólogos a estes fatores foram identificados posteriormente em diversas cultivares de interesse econômico, como arroz, trigo, cevada e milho (Liu et al., 1998; Sakuma et al., 2002; Nakashima & Yamaguchi-Shinozaki, 2010). Estudos demonstraram que a regulação da expressão e da atividade dos membros desta subfamília de fatores de transcrição ocorre em dois níveis: pós transcricional, por processamento (ou *splicing*) alternativo de mRNA; e pós traducional, pela necessidade de fosforilação de alguns membros para interação com o elemento *DRE* presente em promotores de genes alvo (Agarwal et al., 2007). É interessante ressaltar que, em arroz, o transcrito do gene OsDREB2A é induzido com 24 horas de déficit hídrico e 250,0 mM de estresse salino, mas também responde à ABA e à baixas temperatura, diferentemente do padrão estabelecido em outras plantas (Dubouzet et al., 2003). Adicionalmente, estudos com todos os cinco OsDREB2s de arroz mostram que eles se apresentam em níveis mais elevados em situações normais (não estressadas) e sofrem um ligeiro aumento em condições de alta temperatura, seca e alta salinidade, mas não em baixas temperaturas (Matsukura et al., 2010). Desta forma, a identificação destes e de muitos outros ortólogos de DREB1 e DREB2 propiciaram uma melhor compreensão dos mecanismos de sinalização decorrentes de estímulos ambientais, além de abrirem caminhos para obtenção de plantas geneticamente modificadas com maior tolerância aos mais diferentes tipos de estresse.

## 1.6. A SOJA.

### 1.6.1. Principais Características.

Na atualidade, a soja é um dos principais agronegócios mundiais, sendo responsável pela movimentação de bilhões de dólares em todo o mundo. Esta importante *commodity* pertence à classe *Dicotyledoneae*, subclasse *Archichlamydae*, ordem *Rosales*, subordem *Leguminosinae*, família *Leguminosae*, subfamília *Papilionaceae*, tribo *Phaseoleae*, gênero *Glycine* L., subgênero *Glycine* (Moench) e à espécie *Glycine max* (L.) Merrill (Singh et al., 1987). *Glycine* é o único gênero da tribo *Phaseoleae* que possui espécies diplóides com número de cromossomos de 40 a 80 e não de 20, conforme é o caso do feijão (*Phaseolus vulgaris*), sendo então classificada como uma planta tetraplóide diploidizada ( $2n = 40$ ). Esta característica particular é provavelmente derivada de ancestrais diplóides com número base de 11 cromossomos, que sofreram processo de aneuploidia e passaram a ter um número base de 10 cromossomos. Desta forma, a soja cultivada tem sido considerada um tetraplóide estável com genomas diploidizados (Gurley et al., 1979; Skorupska et al., 1989). Com base nestas informações, Schmutz e colaboradores (2010) sequenciaram 1,1 gigabases (Gb) correspondentes aos 20 cromossomos da soja cultivada na atualidade (*Glycine max* [L.] Merrill, cultivar William82), que foram posteriormente integrados com mapas genéticos de alta densidade, podendo assim gerar modelos de estrutura física correspondentes a cada cromossomo da planta. Foram preditos 46.430 genes codificadores de proteínas, 70,0 % a mais que *A. thaliana* e similar ao genoma do álamo (*Populus nigra*), que semelhante à soja, é um ancestral poliploide (paleopoliploide). O sequenciamento do genoma desta cultura possibilitou estabelecer a data aproximada da duplicação de seu material genético (59,0 - 13,0 milhões de anos atrás), evento este que resultou na multiplicidade de cópias de aproximadamente 75,0 % de seus genes (Schmutz et al., 2010).

Quanto à composição do grão, cerca de 35,0 a 40,0 % do peso seco da soja é composto por proteínas, sendo que 90,0 % delas correspondem a duas globulinas de estoque (glicina-11S e  $\beta$ -conglicina-7S) (Liu, 1997). Além disso, apresentam aproximadamente 19,0 % de óleo em sua estrutura, onde os triglicerídeos são o maior

componente, possuindo também em sua composição uma grande amostragem de ácidos graxos poliinsaturados (PUFA), (55,0 % de ácido linoleico e 8,0 % de ácido  $\alpha$ -linoleico do total de ácidos graxos) (Liu, 1997), assim como 350,0 mg (aproximadamente) de esteróis vegetais (54,5 % de  $\beta$ -sitosterol, 21,5 % de campesterol e 19,0 % de estigmasterol) e cerca de 1,0 - 3,0 % de fosfolipídios (35,0 % de fosfatidilcolina, 25,0 % de fosfatidiletanoamina, 15,0 % de fosfatidilinositol e 5,0 - 10,0 % de ácido fosfatídico) (Liu, 1997). Grande parte da composição da soja também é representada por carboidratos (cerca de 35,0 %), sendo a maioria polissacarídeos distintos do amido (4,0 % de estaquiose e 1,1 % de rafinose). Quando comparado com cereais, este grão também é uma grande fonte de vitaminas, principalmente do complexo B (com exceção da vitamina B-12) e vitamina E ( $\alpha$ -tocoferol,  $\beta$ -tocoferol,  $\gamma$ -tocoferol e  $\delta$ -tocoferol) (Liu, 1997), podendo também ser rica em minerais (aproximadamente 5,0 % do peso seco), como potássio, fósforo, cálcio e ferro (estocado em ferritinas) (Liu, 1997). A soja também é uma das fontes mais abundantes de isoflavonóides (acima de 3,0 mg/g de peso seco), representados neste grão pela daidzeína, genisteína e gliciteína, além de saponinas (aproximadamente 2,0 % do peso seco), glicosídeos triterpênicos com grande importância no tratamento de hiperlipidemia, câncer e doenças infecciosas (Liu, 1997; Friedman & Brandon, 2001).

O germoplasma de soja possui grande diversidade quanto ao ciclo (número de dias da emergência à maturação), variando de 70 dias, para as mais precoces, até 200 dias para as mais tardias. De modo geral, as variedades brasileiras têm ciclo entre 100 e 160 dias e, para determinadas regiões, podem ser classificadas em grupos de maturação precoce, semiprecoce, médio, semitardio e tardio. O ciclo total da planta pode ser dividido em duas fases: *vegetativa* e *reprodutiva*. A fase vegetativa é o período da emergência da plântula até a abertura das primeiras flores, e a fase reprodutiva compreende o período do início da floração até a maturação (Boerma & Specht, 2004).

### **1.6.2. Estabelecimento e Conjuntura Atual da Soja.**

Acredita-se que a soja cultivada (*Glycine max* [L.] Merrill) originou-se da soja selvagem (*Glycine soja* Sieb. e Zuccna) na região leste da China, onde sofreu domesticação por volta do século XI a.C. A partir daí, foi introduzida em outras regiões e

países do Oriente, como Manchúria, Coréia, Japão, União Soviética e países do sudeste da Ásia, que são considerados centros secundários e/ou terciários de diversidade (Chung & Singh, 2008). No Ocidente, sua introdução se deu a partir do século XVIII, quando foram plantadas na Europa, sementes recebidas pelo Jardim Botânico de Paris. No continente americano, o primeiro relato envolvendo a soja data de 1765, com registro de plantio na Geórgia, EUA (Hymowitz & Newell, 1981).

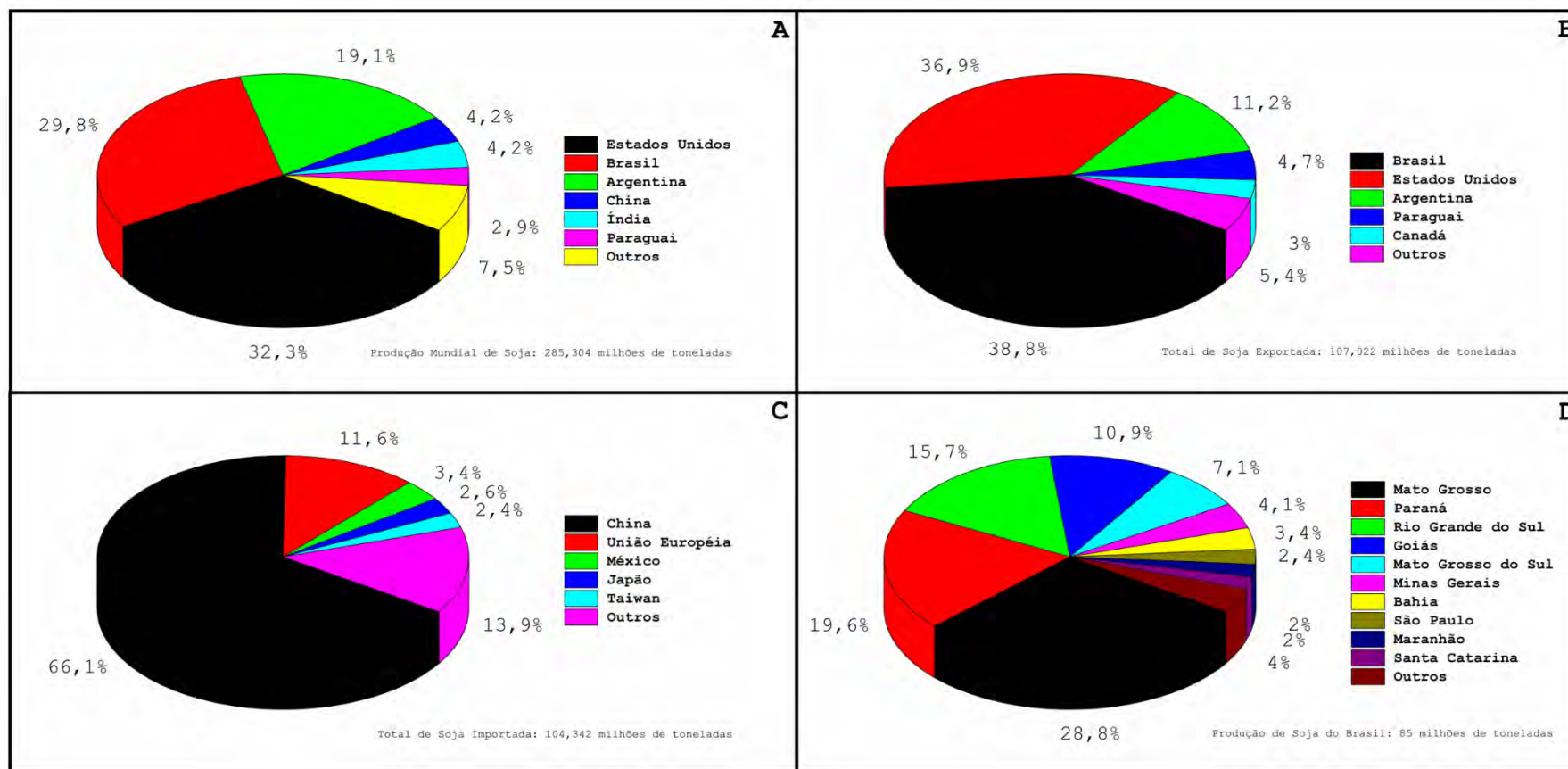
No Brasil, o primeiro registro da introdução da soja data de 1882, na Bahia, por Gustavo Dutra (Sediyama, 1985), que realizou em 1891 experimentos com a soja amarela na Estação Agronômica de Campinas (SP), para teste como planta forrageira. Em 1900, esta estação paulista distribuiu sementes de soja para produtores do Estado e, em 1914, têm-se o registro o primeiro cultivo de soja no Rio Grande do Sul (RS), no município de Santa Rosa, onde apresentou condições ideais de cultivo (Miyasaka & Medina, 1981).

Desta forma, as primeiras estatísticas revelando a utilização de soja para produção de grãos no Brasil datam de 1941, no estado do Rio Grande do Sul, que produziu, em 640 hectares (ha) de área cultivada, 450 toneladas do grão, com um rendimento médio de 700 kg/ha (Schnepf et al., 2001). No estado do Paraná, a soja foi introduzida por agricultores gaúchos que emigraram para o sudoeste e então para o oeste do Estado. A produção no Paraná cresceu rapidamente até atingir 60.000 toneladas na década de 60 (Schnepf et al., 2001). Diversos fatores contribuíram para o estabelecimento da soja como uma cultura de destaque na região Sul brasileira, dentre eles destacam-se: semelhança de ecossistema com a região de maior plantio nos EUA; correção do solo através do processo de calagem; incentivos fiscais e na área de pesquisa, associados com a alta do agronegócio no mercado internacional; substituição da gordura animal por óleos vegetais; estabelecimento de um importante parque industrial de processamento de soja, juntamente com a facilidade de mecanização total da cultura; surgimento de um sistema cooperativista dinâmico e eficiente; e melhorias nos sistemas viário, portuário e de comunicações, o que facilitou o transporte e as exportações do grão (Schnepf et al., 2001; Ortega et al., 2005; Cavalett & Ortega, 2009). A partir da década de 70, a cultura da soja evoluiu significativamente nos estados produtores, não só no Sul, mas também nos estados do Centro-Oeste do Brasil. Esta região era considerada a maior e principal fronteira do plantio da soja no Brasil, e dentre as causas do espetacular crescimento deste agronegócio na região destacam-se a construção de Brasília, culminando no estabelecimento de diversas obras de infraestrutura

regional; incentivos fiscais; desenvolvimento agroindustrial da região; baixo valor da terra quando comparado com o da região Sul na mesma época; topografia altamente favorável à mecanização, associada com as boas condições físicas dos solos da região; melhorias no sistema de transporte da produção regional; e regime pluviométrico da região altamente favorável aos cultivos de verão, em contraste com o da região Sul, destacadamente no Rio Grande do Sul (Schnepf et al., 2001).

Com o desenvolvimento de novas cultivares adaptadas às diferentes regiões agroclimáticas do país, o Brasil tornou-se o segundo maior produtor mundial de soja (perdendo somente para os EUA), com área cultivada na safra de 2013/2014 de aproximadamente 29,5 milhões de hectares e contribuindo com 29,6 % da produção mundial do grão (Figura 1-A). Além disso, o Brasil é o maior exportador mundial de soja, sendo responsável por 38,8 % das 107,0 milhões de toneladas exportadas no mundo neste período (Figura 1-B). Deste montante, a China, a União Europeia e o México são responsáveis por mais de 80,0 % da soja importada em todo o globo (Figura 1-C). No Brasil, a região Centro-Oeste, representada pelos estados do Mato Grosso e Goiás, juntamente com alguns estados da região Sul (Rio Grande do Sul e Paraná) são responsáveis por 75,0 % de toda a produção brasileira de soja (Figura 1-D) (CONAB, 2014 - <http://www.conab.gov.br/>; IBGE, 2014 - <http://www.ibge.gov.br/>).

A principal utilidade da soja é fornecer óleo e farelo, sendo estes os produtos mais produzidos, comercializados e utilizados (Boerma & Specht, 2004). Como o farelo de soja é rico em proteínas, o mesmo é a principal fonte do nutriente em rações animais. O óleo de soja é um dos óleos vegetais mais produzidos no mundo e sua demanda, além da alimentícia, vem aumentando com o passar dos anos devido ao seu uso na produção de combustíveis renováveis, como biodiesel (Vasudevan & Briggs, 2008). De acordo com dados da Companhia Nacional de Abastecimento (CONAB), até junho de 2013, o Brasil foi responsável pela produção de mais de 5,8 milhões de toneladas de farelo de soja, de aproximadamente 580,0 mil toneladas de óleo proveniente do grão e mais de 2,0 bilhões de litros de biodiesel, movimentando assim mais de 3,5 bilhões de dólares no período (CONAB, 2014 - <http://www.conab.gov.br/>).



**Figura 1. A Soja no Brasil e no Mundo.** Os gráficos apresentam os principais índices agropecuários relacionados à soja no período de 2013/2014. **A** - produção mundial e principais países produtores; **B** e **C** - dados de exportação e importação mundial, respectivamente; **D** - principais estados brasileiros produtores e suas respectivas produções. Fonte: Companhia Nacional de Abastecimento (CONAB) e Instituto Brasileiro de Geografia e Estatística (IBGE).

### 1.6.3. Principais Pragas.

Entre os diversos fatores que limitam a produtividade que afetam a qualidade da soja, as doenças estão entre os mais importantes e de difícil controle. Com a expansão da cultura para novas fronteiras agrícolas, a intensificação da monocultura, a adoção de práticas inadequadas de manejo e a mudança do quadro de cultivares, têm aumentado o número de doenças e os níveis de danos causados por microrganismo e parasitas, principalmente fungos e nematoides. Além dos agentes patogênicos que foram introduzidos como contaminantes das regiões ou países de origem das primeiras sementes, nos últimos anos, novos patógenos, nativos ou introduzidos, incluindo diversas anormalidades de causa ainda desconhecidas, têm preocupado os produtores, a assistência técnica e os pesquisadores nos últimos anos.

Dentre as diversas doenças de soja causadas por fungos, a ferrugem asiática é uma das piores pragas relacionadas a esta *commodity* no mundo. Tendo como agente etiológico o fungo *Phakopsora pachyrhizi*, esta infecção pode ser avaliada pela sua rápida expansão, virulência e pelo montante de perdas causadas. No Brasil, a maior intensidade de seus danos é mais comum nos Cerrados, onde o clima favorece a doença, dificulta o controle e as grandes extensões das lavouras representam um desafio à mais nas pulverizações. No começo da doença, pequenas lesões castanhas são observadas nas vias foliares. Estas lesões aumentam com cinco ou seis dias e as protuberâncias com aspecto de ferrugem (urédias) são observadas, compreendendo as estruturas de reprodução do fungo. As urédias desenvolvem-se mais frequentemente em lesões na superfície inferior das folhas, e ao abrirem-se geram um ostíolo por onde os esporos são liberados, principalmente pelo vento. Quanto ao controle, se a doença já estiver se estabelecido, o controle químico com fungicida é, até o momento, a principal medida de controle. O número e a necessidade de novas aplicações dependem do estágio inicial em que foi identificada a doença na lavoura e do período residual dos produtos. Vale a pena ressaltar que algumas cultivares de soja são resistentes à infecção por este fungo, o que as torna alvos de estudos para melhor compreensão da doença e dos mecanismos de defesa que a planta dispõe para combatê-la (Goellner et al., 2010).

O fungo *Sclerotinia sclerotiorum* (Lib.), agente causal do mofo branco, encontra-se amplamente disseminado e afeta mais de 400 espécies de plantas, dentre elas a cultura da

soja, causando danos consideráveis no rendimento. A transmissão do patógeno ocorre por meio de sementes contendo micélio dormente ou pela presença de escleródios no lote, os quais podem permanecer viáveis por um longo período, dificultando o manejo da doença. Os sintomas iniciais desta infecção são lesões encharcadas nas folhas ou qualquer outro tecido da parte aérea que normalmente tenham tido contato com as flores infectadas. As lesões espalham-se rapidamente para as hastas, ramos e vagens. Nos tecidos infectados surge uma eflorescência branca que lembra algodão, constituindo os sinais característicos da doença. Até a cultura chegar ao florescimento, dificilmente a doença torna-se importante. Após este período, a doença é disseminada rapidamente devido ao fato da flor ser fonte de nutrição para o fungo iniciar novas infecções. Não há disponibilidade de cultivares de soja geneticamente resistentes a *Sclerotinia sclerotiorum*, podendo sim haver diferenças de suscetibilidade entre elas. Diversidades no dossel, porte ou arquitetura das plantas podem influenciar no desenvolvimento da doença. As formas de prevenção do mofo branco constituem no uso de sementes sadias; racionalização do volume de água na lavoura; fuga de épocas muito favoráveis ao estabelecimento da infecção com alta umidade e temperaturas mais baixas; incremento de microrganismos antagônicos no solo como o *Trichoderma* spp.; cobertura do solo com *Brachiaria*, visando uma barreira física à germinação dos escleródios presentes no solo; rotação de cultura com gramíneas e uso de fungicidas em tratamento de sementes e parte aérea (Bolton et al., 2006)

Os nematoides parasitas de plantas causam perdas de aproximadamente 100,0 bilhões de dólares anualmente nos EUA (Koenning et al., 1999). As espécies *Meloidogyne javanica* e *Meloidogyne incognita* são caracterizados por formarem estruturas semelhantes à galhas nas raízes e destacam-se pelos danos que causam à soja (Caillaud et al., 2008; Ibrahim et al., 2011). Nas regiões onde ocorrem, as plantas de soja ficam pequenas e amareladas, as folhas apresentam manchas cloróticas e necrose entre as nervuras, além de ser observado um intenso abortamento das vagens e um amadurecimento prematuro das plantas infectadas. Estas espécies de nematoides formadores de galhas são frequentemente encontradas no norte do Rio Grande do Sul, sudoeste e norte do Paraná, sul e norte de São Paulo e sul do Triângulo Mineiro. Na região Central do Brasil, o problema é crescente, com severos danos em lavouras no Mato Grosso do Sul e Goiás. Os danos causados por *Meloidogyne javanica* e *Meloidogyne incognita* são menos severos quando associados com rotação da cultura de soja com amendoim, algodão, sorgo resistente (AG 2005-E, AG



2501-C), mamona ou milho resistente. Desta forma, a utilização de cultivares de soja resistentes aos nematoides de galha é o meio de controle mais eficiente e mais adequado para o agricultor, associados com rotações de cultura não hospedeiras e adubação verde (Davis et al., 2004).

Outro nematoide muito importante em infecções causadas em soja seria o *Heterodera glycines* (nematoide de cisto), uma das principais pragas da cultura pelos prejuízos que pode causar e pela facilidade de disseminação. Ele penetra nas raízes da planta de soja e dificulta a absorção de água e nutrientes, condicionando porte e número de vagens reduzidos, clorose e baixa produtividade. O sistema radicular fica reduzido e infestado por minúsculas fêmeas do nematoide com formato de limão ligeiramente alongado. Inicialmente de coloração branca, a fêmea, posteriormente, adquire a coloração amarela. Após ser fertilizada pelo macho, cada fêmea produz de 100 a 250 ovos, armazenando a maior parte deles em seu corpo. Quando a fêmea morre, seu corpo se transforma em uma estrutura dura denominada cisto, de coloração marrom escura, cheia de ovos, leve, altamente resistentes à deterioração e à dessecação, que se desprende da raiz e fica no solo. Apesar da disponibilidade de centenas de cultivares de soja resistentes à *Heterodera glycines*, associadas com a prática de rotações de cultura com plantas não hospedeiras e ação de produtos químicos (fenamifos e carbofuram), o nematoide continua a ser um fator limitante na produção deste grão. Estudos moleculares realizados com este patógeno ajudaram a identificar possíveis genes relacionados com sua virulência e com o estabelecimento de sua infecção, o que pode prover novos alvos para obtenção de mais cultivares de soja com mecanismos diferenciados de resistência ao nematoide (Niblack et al., 2006).

Nos últimos anos, outra praga emergiu com grande importância no cultivo de diversas *commodities* com relevância econômica no mundo. O inseto *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae) é uma das mais importantes pestes em diversas partes do mundo, pois suas lagartas podem se alimentar tanto dos órgãos vegetativos como reprodutivos de várias espécies de plantas com importância econômica. Estima-se que a perda mundial causada por lagartas de *Helicoverpa armigera*, nas diferentes culturas em que ataca, chega anualmente a 5,0 bilhões de dólares (Avila et al., 2013). Sharma e colaboradores (2008) relataram que a perda anual causada por este inseto supera 2,0 bilhões de dólares apenas na região dos trópicos semiáridos da Europa, e que o

custo anual da aplicação de inseticidas nas lavouras, para o controle dessa praga, é de 500,0 milhões de dólares (Sharma et al., 2008). No Brasil, os maiores danos econômicos causados por lagartas de *Helicoverpa armigera* são observados, até então, nas culturas de algodão, milho, soja, sorgo e tomate. Este inseto pode se alimentar de folhas e hastes dessas plantas, mas possuem preferência pelas estruturas reprodutivas, como botões florais, frutos, maçãs, espigas e inflorescências, causando deformações ou necrose nestas estruturas, ou até mesmo a sua queda. Essa inerente capacidade do *Helicoverpa armigera* causar danos nas partes reprodutivas das culturas, em associação à sua habilidade de atacar um grande número de hospedeiros, são fatores que elevam a importância econômica da praga (Cunningham et al., 1999). Como exemplo pode-se observar as estatísticas da safra brasileira de 2012/2013, onde foram constatados ataques de lagartas de *Helicoverpa armigera* em cultivos de soja nos Estados do Maranhão e Piauí. Em Mato Grosso, essa praga foi também observada atacando lavouras de algodão, soja e milho, enquanto que em Goiás os danos foram mais acentuados nas lavouras de tomate e soja (Czepak et al., 2013). Quanto ao manejo, o aumento da resistência do inseto à inseticidas destaca a necessidade do desenvolvimento de alternativas independentes de produtos químicos para o controle efetivo desta praga, destacando-se principalmente a obtenção de plantas com genótipos resistentes, associadas com a presença de inimigos naturais do inseto, como por exemplo algumas espécies de nematoides (Romeis & Shanower, 1996; Arthurs et al., 2004).

#### **1.6.4. Tolerância à Seca.**

A seca é a modalidade de estresse abiótico mais estudada e com maior relevância no cultivo de soja, uma vez que foi responsável pela perda de 12,2 % na produção do grão no início de 2014 em todo o Brasil. Estudos em campo e em casas de vegetação mostram que a seca leva à uma significativa redução na produção de sementes (24,0 - 50,0 %) nos mais distintos locais e condições ambientais (Frederick et al., 2001; Sadeghipour & Abbasi, 2012). Desta forma, para facilitar o seu cultivo em regiões de clima árido e semiárido é importante classificar as cultivares de soja de acordo com sua tolerância à seca. Para isso, vários parâmetros são adotados, podendo variar para os diferentes fatores ambientais e temporais. O primeiro deles é a definição de tolerância à seca, termo

vagamente usado para incluir todos os mecanismos que permitem a melhor sobrevivência da soja sob déficit hídrico.

Os parâmetros conseguintes estão relacionados com mudanças fisiológicas e morfológicas nas cultivares de soja em resposta a esse tipo de estresse, presumivelmente devido às diferenças em seus genótipos. Dentre eles, destacam-se o tamanho e quantidade das raízes, que durante a seca apresentam-se em baixa densidade na superfície desidratada e altamente densa em regiões mais profundas, onde o conteúdo hídrico é maior (Liu et al., 2005; Benjamin & Nielsen, 2006); o aumento relação entre o crescimento das raízes e dos brotos, o que indica uma redução no desenvolvimento da parte aérea da planta mas não das raízes (Wu & Cosgrove, 2000); aumento da área foliar, associado à redução da condutância estomática em amplo espectro e diminuição da taxa fotossintética em menor amplitude, o que diminui diretamente a perda de água, minimizando a redução da fotossíntese (Benjamin & Nielsen, 2006; Manavalan et al., 2009; Stolf-Moreira et al., 2010); e manutenção da turgidez, através do ajuste osmótico pelo acúmulo de soluto no meio intracelular (Sloane et al., 1990).

Adaptações bioquímicas também ocorrem em cultivares de soja sob déficit hídrico, destacando-se a eliminação de espécies reativas de oxigênio (ROS ou EROS). Nestas condições de estresse, ocorre um acúmulo de ROS, que é eliminado pelo aumento da expressão dos genes de enzimas como *superóxido desmutases*, *catalases* e *glutathione peroxidases* em cultivares de soja mais tolerantes (Masoumi et al., 2010).

Além disso, condições ambientais adversas podem propiciar o enovelamento incorreto de proteínas, que por sua vez se acumulam no retículo endoplasmático (ER). Cultivares de soja mais tolerantes à déficit hídrico induzem respostas que culminam na ativação destas proteínas localizadas no ER, principalmente por intermédio de chaperonas moleculares (Irsigler et al., 2007; Valente et al., 2009; Liu & Howell, 2010).

Em nível molecular, as cultivares de soja mais tolerantes à seca percebem o déficit hídrico por osmosensores (que não foram completamente caracterizados), que transmitem o sinal por rotas de sinalização que ativam a transcrição de genes importantes neste tipo de resposta (Ahuja et al., 2010). A tolerância não envolve somente a atividade de receptores, proteínas cinases, fatores de transcrição e efetores, mas também a produção de metabólitos e mensageiros para a transdução de sinal, sendo de natureza multigênicas, envolvendo mecanismos e redes moleculares complexas, que possuem ampla sobreposição com

respostas à estresse osmótico (Ahuja et al., 2010). É neste nível que a maior parte dos esforços na obtenção de cultivares de soja tolerantes à seca atuam, principalmente com o auxílio da tecnologia do DNA recombinante e fundamentos de engenharia genética.

#### **1.6.5. Projeto GENOSOJA.**

Com a finalidade de melhor compreender o desenvolvimento da soja em condições normais e de estresse, assim como melhor caracterizar as cultivares adaptadas em território brasileiro, foi criado em 2007 (e concluído em 2010) o Consórcio Brasileiro do Genoma da Soja (GENOSOJA), formado pela associação de nove instituições brasileiras de diferentes regiões com órgãos brasileiros de apoio à pesquisa e fomento (Ministério da Ciência e Tecnologia - MCT; e Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq). A proposta teve como objetivo estudar o genoma da soja a partir de sua organização em nível estrutural, buscando caracterizar e sequenciar importantes regiões genômicas e seus produtos, assim como contribuir com a identificação de genes expressos em resposta a diferentes estresses bióticos e abióticos, que afetam a cultura no hemisfério Sul, com auxílio de métodos transcricionais e proteômicos. A rede GENOSOJA não tinha como objetivo somente determinar se a expressão de um gene seria induzida ou reprimida, mas também definir os níveis em que seriam expressos, a interação com outros genes, além de sua localização física e produtos, permitindo assim a identificação de genes e rotas metabólicas essenciais para o desenvolvimento e para o estudo de tolerância da soja em situações adversas (Benko-Iseppon et al., 2012).

Desta forma, o projeto foi composto de diversos componentes, dentre os quais se destacam: (i) *genômica estrutural*, que incluiu pesquisas relacionadas com arquitetura física do genoma da soja (cultivar Conquista), incluindo análises de regiões promotoras e sequenciamento de regiões gênicas e identificação de polimorfismos de base única (SNPs); (ii) *transcriptômica*, compreendendo pesquisas que abrangeram a determinação de vários perfis de expressão de genes expressos em estresses bióticos (ferrugem asiática - *Phakopsora pachyrhizi*; vírus da necrose da haste da soja ou *cowpea mild mottle virus* - CPMMV; e os nematoides *Meloidogyne javanica* e *Pratylenchus brachyurus*) e abiótico (déficit hídrico), utilizando inúmeras estratégias (bibliotecas subtrativas de cDNA,

SuperSAGE, bibliotecas de miRNAs e sequenciamento de cDNAs); (iii) *proteômica*, que visou estudar o perfil de expressão e interações entre proteínas de soja nas mesmas condições estabelecidas nas análises de transcriptômica com a finalidade de garantir a integração dos bancos de dados gerados; e (iv) *ensaios de expressão* (transgênese), que visou a validação dos resultados de genômica e transcriptômica e identificação de regiões regulatórias (elementos *cis*-atuantes) em promotores gênicos (Benko-Iseppon et al., 2012).

Como resultado, este projeto identificou milhares de genes diferencialmente expressos, que foram posteriormente validados por reação em cadeia da polimerase em tempo real (qPCR). Muito deles não eram caracterizados anteriormente, sem nenhuma função atribuída, o que conferiu relevância para estudos futuros de funcionalidade e de importância para o melhoramento de plantas (Benko-Iseppon et al., 2012).

## 2. JUSTIFICATIVA.

Tendo em vista a importância da soja como agronegócio no mundo, é necessário o melhor entendimento de sua fisiologia durante seu desenvolvimento e em respostas à adversidades ambientais. Desta forma, a caracterização e compreensão do papel dos fitormônios nestes eventos seria um ponto chave neste processo de compreensão, visando posterior melhoramento desta *commodity* à estas situações de estresse. De acordo com a literatura, pouco se sabe sobre as rotas de biossíntese e sinalização mediada por etileno em soja, principalmente sobre o ponto de vista molecular. Em associação, visto que a transdução de sinal que é desencadeada por este fitormônio culmina na regulação de inúmeros genes, a caracterização de moléculas que possam vir coordenar estes eventos no núcleo da célula vegetal também se faz necessária. Em plantas modelo, esta função é atribuída aos fatores de transcrição, sendo neste caso membros da superfamília AP2/ERF.

### 3. OBJETIVOS.

#### 3.1. OBJETIVOS GERAIS.

O trabalho visou caracterizar *in silico* os principais componentes moleculares (genes e respectivas proteínas por eles codificadas) que podem participar das rotas de biossíntese e sinalização mediada por etileno em soja, além de avaliar a sua importância em respostas à seca nesta mesma espécie vegetal.

#### 3.2. OBJETIVOS ESPECÍFICOS.

Este trabalho visou inicialmente identificar e caracterizar *in silico* os membros relacionados com a biossíntese e com a rota metabólica sinalizada por etileno, juntamente com os representantes da superfamília AP2/ERF de fatores de transcrição presentes no genoma da soja, tendo como base conhecimentos prévios importados de plantas modelo, como *Arabidopsis thaliana* e *Oryza sativa*.

Em seguida, os resultados obtidos foram correlacionados com os dados de transcriptoma depositados no banco de dados públicos e disponibilizados pelo Projeto GENOSOJA, tendo como objetivo avaliar o nível de expressão dos genes identificados em cultivares de soja (suscetível e tolerante à seca) em condições de déficit hídrico.

Após a validação dos dados obtidos *in silico*, com auxílio de metodologias de análise *in vitro* e *in vivo*, foi possível inferir o papel deste fitormônio gasoso e da rota metabólica por ele sinalizada, além do perfil de expressão dos fatores de transcrição da superfamília AP2/ERF em cultivares de soja crescidas em sistema hidropônico e submetidas à seca.

## 4. MATERIAL E MÉTODOS.

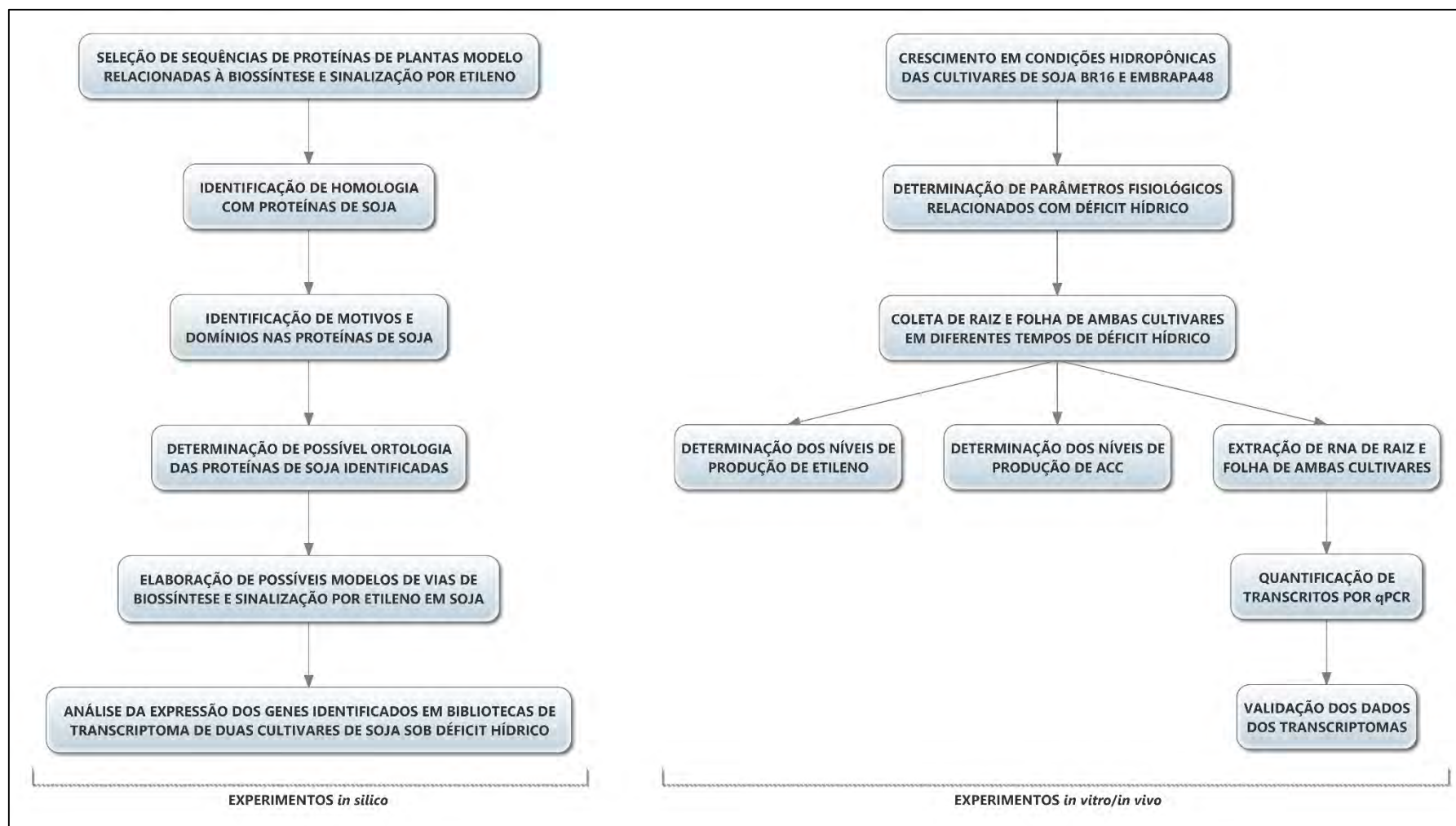
O estudo foi realizado em duas etapas, conforme ilustrado na Figura 2: a primeira *in silico*, onde foram identificados e caracterizados os genes/proteínas de soja relacionados com a biossíntese e sinalização por etileno, tendo como base o que já foi descrito para *Arabidopsis thaliana*. Em seguida, o perfil de expressão dos transcritos destes genes foi avaliado com base em dados de transcriptoma de soja sob déficit hídrico (seca). Em um segundo momento, experimentos *in vitro/in vivo* foram utilizados como ferramentas para validação dos dados obtidos *in silico*.

### 4.1. EXPERIMENTOS IN SILICO.

#### 4.1.1. Identificação dos Genes Alvo.

Inicialmente, foram selecionados no banco de dados TAIR (*The Arabidopsis Information Resource*; <http://www.arabidopsis.org/>) (Swarbreck et al., 2008) os genes referentes à biossíntese e sinalização por etileno em *Arabidopsis thaliana*, com exceção aos que se referem a fatores de transcrição da superfamília AP2/ERF, que foram selecionados a partir de dados publicados em artigo científico (Nakano et al., 2006). Em seguida, com auxílio da ferramenta *Basic Alignment Search Tool* (BLAST) (Altschul et al., 1990), foram identificadas as sequências de resíduos de aminoácidos homólogas que apresentaram  $score \geq 200$  bits em varreduras de bancos de dados *online* específicos de soja - *Glycine max* [L.] Merrill (*GENOSOJA*; <http://bioinfo03.ibi.unicamp.br/soja/>) e arroz - *Oryza sativa* Nipponbare (*Rice Genome Annotation Project*; <http://rice.plantbiology.msu.edu/>), além do portal *Phytozome* versão 9.1 (*The JGI Comparative Plant Genomics Portal*; <http://www.phytozome.net/>) (Goodstein et al., 2012; Nascimento et al., 2012; Kawahara et al., 2013). Os genes referentes aos fatores de transcrição responsivos a etileno (ERF) de arroz também foram selecionados a partir de dados de Nakano e colaboradores (2006).





**Figura 2. Organograma de Trabalho.** O fluxograma apresentado esquematiza os principais experimentos realizados neste trabalho. Cronologicamente, os experimentos *in silico* antecederam os *in vitro* e *in vivo*.

Posteriormente, os genes de soja foram classificados de acordo com a sua ontologia através de análise com *software Blast2GO (Gene Ontology)* (Conesa et al., 2005), onde foram categorizados quanto à: (i) *componente celular*, onde foram sugeridos locais de atuação de cada um deles, em nível de subestruturas celulares e complexos macromoleculares; (ii) *função molecular*, onde foram descritas as atividades catalíticas ou ligação em nível molecular de cada um; e (iii) *processos biológicos*, onde foram descritas rotas bioquímicas coordenadas por um ou mais conjuntos de funções moleculares. Para tal fim, as sequências de nucleotídeos de cada gene de soja foram processadas com auxílio da ferramenta *BlastX* (que realiza buscas nos bancos de dados de proteínas, utilizando sequências de nucleotídeos traduzidas nas seis fases de leitura possíveis), tendo como banco de dados as sequências de proteínas de *A. thaliana* e sendo selecionadas somente as correspondências com  $e\text{-value} \leq e^{-10}$ . Após a devida anotação, a funcionalidade das sequências foi analisada com auxílio da ferramenta *online InterproScan versão 5.0* (Jones et al., 2014) e, finalmente, a ontologia determinada pelo *software online GO-Slim* (Ashburner et al., 2000).

#### **4.1.2. Determinação de Domínios Proteicos.**

A partir da sequência de aminoácidos das possíveis proteínas selecionadas nos três organismos estudados foram identificados domínios proteicos comuns e característicos, depositados em bancos de dados *online*. Para este fim, foram utilizadas ferramentas de bioinformática presentes no site *PFAM (Protein Family; http://pfam.xfam.org/)* (Punta et al., 2012), utilizando como parâmetros de seleção aqueles definidos como significantes pelos elaboradores da ferramenta ( $e\text{-value} < 1,0$ ).

#### **4.1.3 Alinhamentos, Ideogramas, Dendogramas e Mapas Metabólicos.**

O ideograma que representa a localização dos genes de soja analisados nos 20 cromossomos da *commodity* foi construído em escala proporcional ao tamanho do cromossomo (1,000 cm corresponde à 5,00 megabases), onde foram consideradas a

localização de cada gene e a fita de DNA em que estavam presentes (senso ou anti-senso). A posição dos centrômeros, assim como o tamanho de cada cromossomo foram obtidos a partir do genoma de referência (Schmutz et al., 2010).

Os alinhamentos e dendogramas foram obtidos com auxílio dos programas *BIOEDIT versão 7.0.9.0* (Hall, 1999) (matriz BLOSUM64) e *MEGA versão 5* (Tamura et al., 2011) respectivamente, sendo *Neighbor-joining* o método de análise estatística deste último, com 10.000 (dez mil) repetições do teste de *Bootstrap*. Os modelos propostos para biossíntese e sinalização por etileno em soja foram obtidos com auxílio dos bancos de dados *SoyCyC versão 3.0* (*Soybean Metabolic Pathway Database; Soybase; <http://www.soybase.org:8082/>*) e *KEGG* (*Kyoto Encyclopedia of Genes and Genomes; <http://www.genome.jp/kegg/>*) (Selkov et al., 1998; Kanehisa & Goto, 2000).

#### **4.1.4. Determinação de Possíveis Sítios de Fosforilação.**

Os possíveis sítios de fosforilação por MAP cinases e proteínas cinases cálcio dependentes, presentes nas enzimas 1-aminociclopropano-1-ácido carboxílico sintases (ACSs) tipo I e do tipo II respectivamente, foram determinados pelo programa de bioinformática online *NetPhos versão 2.0* (<http://www.cbs.dtu.dk/services/NetPhos/>) (Blom et al., 1999). Foram analisadas a presença de possíveis sítios de fosforilação em resíduos de tirosina, serina e treonina no C-terminal das enzimas em questão.

#### **4.1.5. Determinação de Motivos Grupo-Específicos Comuns.**

A determinação de motivos característicos de cada grupo de fatores de transcrição responsáveis a etileno (ERFs) da família ERF de soja foi obtida com o auxílio da ferramenta online *MEME* (*Multiple Em for Motif Elucidation; <http://meme.nbcr.net/meme/cgi-bin/meme.cgi/>*) (Bailey et al., 2009). Foram selecionados motivos com até 60 resíduos de aminoácidos de extensão que estivessem presente em pelo menos duas sequências proteicas distintas que apresentassem entre si no máximo 50,0 % de similaridade. O domínio AP2 não está incluso nesta análise.

Após a identificação dos motivos característicos de cada grupo da família ERF de soja, foi determinada a presença de cada um deles em fatores de transcrição semelhantes em outras espécies de plantas, utilizando sua sequência consenso (expressão regular) e a ferramenta de bioinformática *online Motif Search - GenomeNet* (<http://www.genome.jp/tools/motif/>) (Kanehisa, 1997). A função de cada motivo foi inferida por homologia com sequências proteicas presentes em outras proteínas. Tal análise foi realizada com auxílio da ferramenta *online* de bioinformática de análise de motivos *BLOCK* (<http://blocks.fhrc.org/>), onde são analisados múltiplos alinhamentos de segmentos altamente conservados de proteínas, com ausência de *gaps* (deleções ou “espaços”), utilizando banco de dados e ferramentas de busca específicos, denominado de *LAMA (Local Alignment of Multiple Alignments)*; [http://blocks.fhrc.org/blocks-bin/LAMA\\_search.sh](http://blocks.fhrc.org/blocks-bin/LAMA_search.sh)) (Henikoff & Henikoff, 1994; Henikoff et al., 1995; Pietrokovski, 1996; Henikoff et al., 2000). A análise de homologia de motivos foi realizada utilizando os parâmetros padrões das ferramentas utilizadas.

#### **4.1.6. Modelagem Molecular dos Domínios AP2.**

Os domínios AP2 da família ERF de soja tiveram sua estrutura terciária gerada através de metodologias *in silico* de modelagem molecular por homologia estrutural, utilizado como modelo a estrutura cristalográfica deste domínio presente no fator de transcrição responsivo à etileno 1 (ERF1) de *A. thaliana*, depositado em banco de dados de proteínas (PDB - código 1GCC) (Allen et al., 1998). Os modelos virtuais foram construídos inicialmente a partir do alinhamento da sequência de aminoácidos de cada candidato com o modelo de *A. thaliana*, utilizando a ferramenta de bioinformática *online Swiss Model* (<http://swissmodel.expasy.org/>) (Biasini et al., 2014). Posteriormente, o refinamento da estrutura foi obtido com auxílio do *software Modeller versão 9.14* (Eswar et al., 2001), gerando 200 possíveis modelos para cada sequência analisada, dentre os quais foram escolhidos os que apresentaram menor valor *Z* (*Z-score*).

#### **4.1.7. Experimentos de Melhor “Hit” Bidirecional (BBH).**

Utilizando a ferramenta BLAST, cada sequência proteica identificada no banco de dados de soja (*GENOSOJA*), denominadas de *A*, foi comparada separadamente com o banco de dados de *Arabidopsis thaliana* (*TAIR*) e de arroz (*Rice Genome Annotation Project*), obtendo-se assim as proteínas homólogas *B* e *C* respectivamente. Foram determinados eventos BBH (*Best Bidirectional Hit*) positivos aqueles em que quando *B* e/ou *C* foram comparados com o banco de dados de soja, a melhor similaridade foi *A*. A duplicação gênica foi considerada com a finalidade de evitar falsos negativos (Dalquen & Dessimoz, 2013).

#### **4.1.8. Identificação de Elementos *cis*-Atuantes em Possíveis Promotores.**

Foi analisada a presença de elementos *cis*-atuantes em possíveis regiões promotoras [2.000 pares de bases à montante da fase aberta de leitura (*open read frame* - ORF)] de cada gene de soja selecionado para este estudo. Esta análise foi realizada com auxílio da ferramenta de bioinformática *MatInspector versão 8.0* (*Genomatix*<sup>®</sup>), utilizando “*Plants*” como *matrix group*, “0,85” como o valor para a similaridade das principais bases que compõe cada elemento *cis*-atuante (*core similarity*), e “*Optimized +I*” como valor para a similaridade da matriz (*matrix similarity*) (Cartharius et al., 2005).

#### **4.1.9. Análise de Expressão em Transcriptomas de Soja.**

As bibliotecas subtrativas das cultivares de soja BR16 e EMBRAPA48 sob déficit hídrico foram construídas por Rodrigues e colaboradores (Rodrigues et al., 2012). As plantas foram crescidas em hidroponia e expostas a déficit hídrico de curta duração (o crescimento e a extração de RNA de raiz e folha das plantas serão descritos respectivamente na seção 4.2.1 e 4.2.5 deste estudo). As amostras de cada replicata biológica foram extraídas separadamente e quantidades equimolares de RNA total foram agrupadas de acordo com o tempo de déficit hídrico (25 à 50, 75 à 100 e 125 à 150

minutos). Em seguida, a partir do RNA total agrupado, foi purificada a fração poliA<sup>+</sup> (mRNA) com o kit *FastTrack MAG mRNA Isolation* (Invitrogen, Inc.) para a produção subsequente do cDNA por transcrição reversa (Morrissey & Collins, 1989; Stone et al., 1996). As bibliotecas de raiz e folha das plantas BR16 e EMBRAPA48 foram construídas com o kit *PCR Select cDNA Subtraction* (Clontech, Inc.) (Diatchenko et al., 1996). A primeira fita de cDNA foi sintetizada a partir de 2,0 µg de amostras de mRNA poliA<sup>+</sup> agrupados (*tester* e *driver*). Após a síntese da segunda fita, os cDNAs (*tester* e *driver*) foram digeridos com a enzima de restrição *Rsa I* com a finalidade de produzir fragmentos com extremidades abruptas (*blunt end*). Uma vez digeridos, as amostras *tester* foram divididas em dois grupos, sendo que cada um foi ligado a dois diferentes adaptadores (I e 2R respectivamente). Como o objetivo era detectar os transcritos expressos em planta sob déficit hídrico, as amostras controle (com zeros minutos de tratamento) foram usadas como *driver* no passo de hibridização. O cDNA *driver* (sem adaptadores) foi adicionado às amostras *tester* (I e 2R) e submetidos à primeira etapa de hibridização, que eliminou sequências comuns em ambas as amostras ou presentes no controle (*driver*). Para a segunda hibridização, as amostras *tester* I e 2R, obtidas na primeira hibridização, foram misturadas, seguindo com a adição do cDNA *driver* desnaturado, com a finalidade de enriquecer as sequências expressas nas amostras *tester*. A amplificação por PCR (Reação em Cadeia da Polimerase) foi realizada com cDNAs dupla fita com diferentes adaptadores como modelo (*template*). As sequências foram exponencialmente amplificadas com oligonucleotídeos (*primers*) específicos para a região dos adaptadores. A segunda PCR foi realizada com *primers nested* (próximos aos adaptadores) com a finalidade de enriquecer as sequências diferencialmente expressas.

Os produtos de PCR foram sequenciados utilizando Sequenciamento por Tecnologia de Síntese Illumina/Solexa (Shen et al., 2005). O sequenciamento das seis bibliotecas da cultivar BR16 (três de raiz e três de folha) geraram sequências únicas de 36 pb (pares de base), com auxílio do equipamento *Genome Analyzer IIe*. Já as bibliotecas da cultivar EMBRAPA48 foram sequenciadas com o equipamento *Genome Analyzer IIz* gerando sequências únicas de 76 pb. Após o *base-calling*, as sequências obtidas foram mapeadas tendo como base o genoma de referência (Schmutz et al., 2010) com auxílio do programa *SOAP* (*Short Oligonucleotide Alignment Program*) (Li et al., 2008). As

sequências não mapeadas foram reunidas em *contigs* e analisadas seguindo o padrão do programa *Edena* (Hernandez et al., 2008).

Com o sequenciamento das bibliotecas construídas foram gerados um total de aproximadamente 42,3 milhões de sequências únicas (27,8% referentes à BR16 e 72,2% à EMBRAPA48), representando um total de 26.627 genes diferencialmente expressos, sendo 8.208 na cultivar BR16 (1.515 em raiz e 6.693 em folha) e 18.419 na cultivar EMBRAPA48 (12.035 em raiz e 6.384 em folha). Os dados referentes a este projeto estão depositados no banco de dados *GENOSOJA* (<http://bioinfo03.ibi.unicamp.br/soja/>) (Nascimento et al., 2012).

A partir destes dados, a expressão de cada gene identificado como sendo membro das rotas metabólicas que envolvem etileno (biossíntese e sinalização) foi avaliada. Os níveis de expressão são apresentados em gráficos indicando os valores normalizados de *FPKM* (fragmentos por quilobase de transcrito por milhão de fragmentos mapeados) para cada gene diferencialmente expresso nas seis bibliotecas geradas para cada cultivar.

## **4.2. EXPERIMENTOS IN VITRO/IN VIVO.**

### **4.2.1. Crescimento das Plantas.**

Sementes das cultivares de soja (*Glycine max* [L.] Merrill) BR16 e EMBRAPA48, classificadas respectivamente como sensível e tolerante a déficit hídrico (Tabela 1) (Casagrande et al., 2001; Texeira et al., 2008), foram germinadas em papel de filtro *Germitest* (Cial, Inc.) por sete dias em câmara de crescimento à temperatura constante de  $25,0 \pm 1,0$  °C e 100,0 % de umidade relativa (Figura 3-A). As mudas foram transferidas para caixas de acrílico com volume de 36,0 L contendo 50,0 % (v/v) de solução nutritiva (solução de Hoagland - Tabela 2) (Hoagland & Arnon, 1938), que foi constantemente aerada e repostada semanalmente (Figura 3-B). Estas plantas foram transferidas para casa de vegetação sob um fotoperíodo natural com ciclos de aproximadamente 12/12 h claro/escuro, temperatura de  $30,0 \pm 5,0$  °C e umidade relativa do ar de  $60,0 \pm 10,0$  %, onde cresceram até atingirem estágio V5 (Figuras 3-C até 3-J) (Fehr et al., 1971). O plano

**Tabela 1.** Características Botânicas e Agrônômicas das Cultivares de Soja BR16 e EMBRAPA48

Classificação	Característica Geral	Característica das Cultivares	
		BR16	EMBRAPA48
Genealógica	–	D69-B10-M58 X DAVIS	DAVIS X PARANÁ X (IAS-4 X BR-5)
Botânica	Cor do Hipocótilo	Verde	Verde
	Cor da For	Branca	Branca
	Cor da Pubescência	Cinza	Cinza
	Cor da Vagem	Marrom Clara	Marrom Clara
	Cor do Tegumento	Amarelo Semi Brilhante	Amarelo Brilhante
	Cor do Hilo	Marrom Claro	Marrom Claro
	Grupo de Maturação	Semi Precoce	Precoce
	Altura da Planta	68,0 cm	84,0 cm
	Peso de 100 Grãos	16,6 g	15,4
	Teor de Óleo	22,6 %	21,2 %
	Teor de Proteína	39,0 %	38,2 %
	Reação à Peroxidase	Negativa	Positiva
Estresse Biótico	Resistente	Cancro da Haste	
		Mancha Olho de Rã	–
		Pústula Bacteriana	
	Suscetível	Mosaico Comum da Soja	
		<i>Meloidogyne incognita</i>	–
		<i>Meloidogyne javanica</i>	
Estresse Abiótico	Tolerância à Seca	Sensível	Tolerante





**Figura 3. Crescimento de Soja em Condições Hidropônicas.** As figuras ilustram em ordem cronológica todo o processo de crescimento em condições hidropônicas das cultivares de soja BR16 e EMBRAPA48. **A** - germinação em papel *Germitest* (dia zero); **B** - plantas selecionadas com 7 dias; **C** e **D** - exemplares vegetais (BR16 e EMBRAPA48, respectivamente) transferidos para o sistema hidropônico; **E** - raiz com 10 dias; **F** e **G** - porção aérea (folhas e caule) com 10 dias (BR16 e EMBRAPA48, respectivamente); **H** e **I** - plantas com 20 dias (BR16 e EMBRAPA48, respectivamente); **J** - cultivares com 25 dias (tempo zero de déficit hídrico); **K** até **P** - 25, 50, 75, 100, 125 e 150 minutos de déficit hídrico, respectivamente.

experimental consistiu em um bloco fatorial randomizado 2 X 7 com três repetições. Os fatores foram as duas cultivares (BR16 e EMBRAPA48) e os sete tempos de déficit hídrico (0, 25, 50, 75, 100, 125 e 150 minutos - Figuras 3-J até 3-P). O estresse consistiu na retirada das plantas da solução hidropônica deixando-as nas caixas vazias por até 150 minutos, nas condições ambientais da casa de vegetação. Para cada tempo de déficit hídrico, foram coletadas amostras de raiz e folha de três plantas de cada cultivar. Todos os experimentos realizados estudo foram desenvolvidos com as plantas crescidas nestas condições.

**Tabela 2.** Reagentes e Concentrações para Preparo da Solução Nutritiva

Solução	Nome	Fórmula	P.M. (g/mol) <sup>1</sup>	I (g) <sup>2</sup>	II (mM) <sup>3</sup>	III (µM) <sup>4</sup>	IV (mL) <sup>5</sup>
A	Nitrato de Potássio	KNO <sub>3</sub>	101,10	101,10	1.000,00	6.000,00	6,00
B	Nitrato de Cálcio Tetra Hidratado	Ca(NO <sub>3</sub> ) <sub>2</sub> .4H <sub>2</sub> O	236,20	236,20	1.000,00	4.000,00	4,00
C	Fosfato de Amônio Monobásico	NH <sub>4</sub> H <sub>2</sub> PO <sub>4</sub>	115,00	115,00	1.000,00	2.000,00	2,00
D	Sulfato de Magnésio Hepta Hidratado	MgSO <sub>4</sub> .7H <sub>2</sub> O	246,50	246,50	1.000,00	2.000,00	2,00
E	Ácido Bórico	H <sub>3</sub> BO <sub>3</sub>	61,80	2,70	46,30	46,30	
	Cloreto de Manganês Tetra Hidratado	MnCl <sub>2</sub> .4H <sub>2</sub> O	197,90	0,72	3,70	3,70	
	Molibdênio Trióxido/Anidrido Molibídico	MoO <sub>3</sub>	143,90	0,02	0,10	0,10	1,00
	Sulfato de Cobre II Penta Hidratado	CuSO <sub>4</sub> .5H <sub>2</sub> O	249,80	0,08	0,30	0,30	
	Sulfato de Zinco Hepta Hidratado	ZnSO <sub>4</sub> .7H <sub>2</sub> O	287,50	0,22	0,80	0,80	
F	Ferro (III) - Ácido Etilenodiamino Tetra Acético	C <sub>10</sub> H <sub>13</sub> FeN <sub>2</sub> O <sub>8</sub>	345,06	4,47	12,95	12,95	1,00

<sup>1</sup> Peso molecular em g/mol.<sup>2</sup> Peso (em gramas) necessário para preparo de 1000 mL de solução estoque.<sup>3</sup> Concentração final (em mM) do reagente na solução estoque.<sup>4</sup> Concentração final (em µM) do reagente na solução nutritiva.<sup>5</sup> Volume (em mL) necessário da solução estoque para 1000 mL de solução nutritiva. Após adição das seis soluções estoque, ajustar para pH 6,6 e completar com água para 1000 mL.

#### **4.2.2. Determinação de Parâmetros Fisiológicos.**

Os parâmetros taxa fotossintética ( $A$ ), radiação fotossinteticamente ativa ( $PAR$ ), concentração interna de  $CO_2$  ( $C_i$ ), condutância estomática ( $g_s$ ) e taxa de transpiração ( $E$ ), foram avaliados com auxílio do equipamento *LI-6400 Portable Photosynthesis System* (LiCor, Inc.). As medições foram realizadas em triplicata, a partir do tempo zero de déficit hídrico nas folhas médias do trifólio mais jovem, totalmente expandido, sob uma densidade de fluxo de fótons de  $1.300 \mu\text{mol m}^{-2} \text{s}^{-1}$ . A variação de temperatura ( $\Delta T$ ) é aqui definida como sendo a diferença entre a temperatura do ar ( $T_{ar}$ ) e a temperatura da folha ( $T_{folha}$ ), ambas avaliadas também com o equipamento LI-6400. A eficiência do uso da água ( $WUE$ ) foi determinada pela razão entre as taxas fotossintética ( $A$ ) e de transpiração ( $E$ ). Os dados foram submetidos à coleção de modelos estatísticos *ANOVA (Analysis of Variance)* (Eisenhart, 1947), com auxílio dos programas *SAS* e *SANEST (Statistical Analysis System versão 8.0)* (Zonta, 1984; Schlotzhauer & Littell, 1997), onde os tratamentos foram comparados pelo teste de Turkey ( $p \leq 0,05$ ).

#### **4.2.3. Determinação da Produção de Etileno.**

As amostras para a determinação da produção de etileno foram coletadas em frascos de vidro com volume de 50,0 mL vedados com tampa de silicone. Para cada tempo de déficit hídrico (0, 25, 50, 75, 100, 125 e 150 minutos), foram coletadas, em triplicata (plantas distintas), 0,5 g de raiz e folha das duas cultivares. Devido ao fato da concentração de etileno ser indetectável na atmosfera do frasco imediatamente, as amostras foram mantidas vedadas à temperatura ambiente por um período de 24 h com a finalidade de tornar mensurável o gás na atmosfera do frasco. Para a análise de etileno, foi empregado o método descrito por Mainardi e colaboradores (2006) utilizando um cromatógrafo a gás da *Hewlett-Packard modelo GC-6890* equipado com detecção por ionização de chama (*Flame Ionization Detector - FID*). A coluna utilizada foi a *HP-Plot Q* (30,0 m, D.I. 0,53 mm). O volume de injeção foi de 1,0 mL. As condições de injeção foram: pressão de 20,0 psi por 2 minutos, fluxo de ventilação de 20,0 mL/min após 30 segundos de injeção e a temperatura do injetor foi mantida à 200,0 °C. As demais condições cromatográficas empregadas

foram: corrida isotérmica a 30,0 °C empregando hélio como gás carregador em fluxo constante de 1,0 mL/min; temperatura do detector em 250,0 °C, fluxo de ar e hidrogênio no detector em 400,0 mL/min e 40,0 mL/min, respectivamente. A estimativa da quantidade de etileno produzida foi feita em relação à injeção de um padrão de 0,1 µL/L de etileno em ar sintético da *Air Liquid*. A produção de etileno é dada em nL (grama de massa fresca)<sup>-1</sup> h<sup>-1</sup> (Mainardi et al., 2006).

#### **4.2.4. Determinação de ACC Livre.**

As determinações de 1-aminociclopropano-1-ácido carboxílico (ACC) livre foram feitas com base no método descrito por Liu e colaboradores (2012). A coleta das amostras foi semelhante à determinação da produção de etileno, com exceção de que foi coletado 1,0 g de amostra em tubos de polipropileno de 50,0 mL (Falcon, Inc.), sendo posteriormente congelada em nitrogênio líquido. Aproximadamente 0,5 g do material foi triturado em gral de porcelana na presença de N<sub>2</sub> líquido. O pó resultante foi transferido para um tubo de centrífuga contendo 5,0 mL de metanol a 60,0 % (v/v) em água ultrapura. O material foi mantido sob agitação por 1 hora à temperatura ambiente e, em seguida, centrifugado a 14.000 x g por 10 minutos à 25 °C. O sobrenadante foi transferido para outro tubo e seco completamente em corrente de nitrogênio. O resíduo foi retomado em 200,0 µL de água ultrapura e transferido para tubo de microcentrífuga de 1,5 mL (Eppendorf, Inc.), ao qual foi adicionado 300,0 µL de tampão borato 200 mM, pH 8,0 e 360,0 µL de fluorescamina 1,0 mM dissolvida em acetona. A mistura foi agitada vigorosamente e mantida a 25,0 °C por 10 minutos, sendo em seguida filtrada através de membrana de 0,45 µm de poro para um frasco de vidro de 2,0 mL. Uma alíquota de 20,0 µL da mistura filtrada foi injetada em um cromatógrafo à líquido acoplado a um detector de fluorescência (*Agilent 1100*). A amostra foi eluída através de uma coluna *Luna C18* (5,0 µm, 300,0 x 4,0 mm, *Supelco, Sigma-Aldrich, USA*) e o efluente monitorado em comprimento de onda de excitação de 378 nm e de emissão de 475 nm. Os resultados foram calculados por padronização externa utilizando curva-padrão de ACC (*Sigma-Aldrich, USA*) no intervalo de 0,1 a 10,0 µg (Liu et al., 2012). A determinação de ACC livre é dada em ng (grama de massa fresca)<sup>-1</sup>.

#### **4.2.5. Extração e Quantificação de RNA Total.**

Utilizando o protocolo do reagente Trizol (Invitrogen, Inc.), o RNA total de raiz e folha das cultivares de soja BR16 e EMBRAPA48 foi extraído em triplicata (com plantas distintas), em diferentes tempos de déficit hídrico (0, 25, 50, 75, 100, 125 e 150 minutos) (Chomczynski & Sacchi, 1987). Ao final de cada tempo, foi triturado aproximadamente 0,5 g de cada tecido em gral de porcelana na presença de N<sub>2</sub> líquido. O pó resultante foi transferido para tubo de microcentrífuga com capacidade para 2,0 mL. Em seguida, cada amostra foi homogeneizada vigorosamente com 1,0 mL de Trizol à temperatura ambiente por 15 segundos. A solução foi posteriormente centrifugada a 12.000 x g por 15 minutos a 4,0 °C. O sobrenadante resultante foi transferido para outro tubo, onde foi adicionado 200,0 µL de clorofórmio (Merck, Inc.). A solução foi agitada vigorosamente por 15 segundos, incubada à temperatura ambiente por 5 minutos (com inversões) e centrifugada a 15.000 x g por 15 minutos a 4,0 °C. A fase aquosa (superior), que correspondeu a cerca de 60,0 % do volume total, foi transferida para um novo tubo, onde o RNA total foi precipitado com adição de 0,5 mL de isopropanol (Merck, Inc.). Após incubação a 4,0 °C por 10 minutos, a solução foi centrifugada a 12.000 x g por 15 minutos a 4,0 °C. O sobrenadante foi descartado e o precipitado de RNA lavado com 1,0 mL de etanol 75,0 % (v/v), seguindo com uma centrifugação a 12.000 x g por 15 minutos a 4,0 °C. O sobrenadante foi mais uma vez descartado e o precipitado seco à temperatura ambiente, com os tubos invertidos, por 5 minutos. O RNA total foi ressuspensão com 50,0 µL de água ultrapura e quantificado em espectrofotômetro (*Qubit* - Invitrogen, Inc.), onde foi analisada também a pureza de cada amostra ( $A_{260nm}/A_{280nm} \geq 1,8$ ;  $A_{320nm} \geq 0,1A_{260nm}$ ). Foi realizada eletroforese em gel de agarose 1,0 % (p/v) para análise da integridade dos RNAs obtidos.

#### **4.2.6. Tratamento das Amostras de RNA Total com DNase I.**

Com a finalidade de excluir qualquer DNA contaminante, as amostras de RNA total foram tratadas com DNase I (Invitrogen, Inc.) (Sutton et al., 1997). Para um volume final de 9,0 µL, foram tratados 4,0 µg de RNA com 0,2 µL de DNase I 10 U/µL em 0,9 µL de Tampão 10 X [200,0 mM Tris-HCl (pH 8,4), 20,0 mM MgCl<sub>2</sub>, 500,0 mM KCl] e água

ultrapura para completar o volume. As amostras foram incubadas em Termomix (Eppendorf, Inc.) a 37,0 °C por 20 minutos, seguidos de 5 minutos a 65,0 °C para inativação da enzima. Para posterior teste de eficácia do tratamento, cada amostra tratada foi molde para realização de uma PCR convencional com oligonucleotídeos que reconhecem especificamente íntrons, seguida de visualização em gel de agarose 1,0 % (p/v). Desta forma, as amostras selecionadas foram somente as PCR negativas.

#### **4.2.7. Síntese de DNA Complementar (cDNA).**

As amostras de RNA total previamente tratadas com DNase I foram utilizadas como molde para a síntese de cDNA. A enzima escolhida para tal procedimento foi a *Moloney Murine Leukemia Virus Reverse Transcriptase* (M-MLV RT) (Kotewicz et al., 1985). Inicialmente, a mistura contendo o RNA previamente tratado com DNase I (4,0 µg em 9,0 µL) e 1,0 µL do oligonucleotídeo NV(dt)<sub>30</sub> (*N*: qualquer desoxirribonucleotídeo e *V*: dGTP, dATP e dCTP) na concentração de 500,0 µg/mL, foi submetida a um processo de desnaturação (incubação a 70,0 °C por 3 minutos seguida de outra a 0,0 °C por mais 3 minutos). Em seguida, a estes 10,0 µL de solução foram adicionados: 5,0 µL do 5 X *First Strand Buffer* [250,0 mM Tris-HCl (pH 8.3 a temperatura ambiente), 375,0 mM KCl, 15 mM MgCl<sub>2</sub>]; 2,0 µL de 0,1 M ditioneitol (DTT); 0,5 µL de *RNAse Out* 40 U/µL (Inibidor Recombinante de RNase - Invitrogen, Inc.); 1,0 µL de 10,0 mM desoxirribonucleotídeos (dNTPs); 0,5 µL de água ultrapura. Esta mistura foi pré-aquecida em Termomix à 42,0 °C por 2 minutos, momento este em que foi adicionado 1,0 µL de M-MLV RT 200 U/µL (Invitrogen, Inc.). Por fim, a reação de síntese da primeira fita de cDNA, com volume final de 20,0 µL, foi incubada a 42,0 °C por 90 minutos, seguida de inativação da enzima por 15 minutos a 70,0 °C. O cDNA gerado foi usado como substrato em experimentos posteriores de determinação de níveis de expressão de transcritos gênicos.

#### 4.2.8. PCR Quantitativa em Tempo Real (qPCR).

Com a finalidade de validar dados de transcriptoma, foram realizados experimentos de PCR em Tempo Real (qPCR) com oito genes identificados como participantes das rotas metabólicas de biossíntese e sinalização por etileno em soja: três referentes à biossíntese (MAT, ACS e ACO); e cinco referentes à sinalização (dois de transdução de sinal - ETR e CTR; e três pertencentes à superfamília de fatores de transcrição AP2/ERF - ERF069, ERF223 e RAV). A seleção destes genes foi feita levando-se em consideração principalmente níveis de expressão diferencial no transcriptoma e relevância nas rotas estudadas.

Os oligonucleotídeos foram desenhados com auxílio do programa *Primer 3 Plus* (<http://primer3plus.com/>) (Untergasser et al., 2007) usando o critério de geração de produtos de PCR (*amplicons*) de 120 a 200 pb com  $T_m = 60,0 \pm 2,0$  °C. Foram utilizados quatro genes como normalizadores de expressão ou endógenos (de referência): dois para as amostras de raiz (*ACT11* e *UBC2*) e dois para folha (*CYP2* e *ELF1A*), sendo que a escolha dos mesmos foi determinada por dados da literatura (Kulcheski et al., 2010; Miranda et al., 2013) e análise de estabilidade nas amostras. Análises de curvas de fusão (*melting*) e de géis de eletroforese dos *amplicons* confirmaram que os oligonucleotídeos amplificaram somente produtos únicos e do tamanho esperado (as informações sobre os oligonucleotídeos encontram-se na Tabela 3).

Os experimentos foram realizados em triplicatas técnicas para cada uma das três réplicas biológicas, totalizando assim nove repetições experimentais para cada amostra de cDNA de raiz e folha das duas cultivares analisadas (BR16 e EMBRAPA48) nos sete diferentes tempos de déficit hídrico (0, 25, 50, 75, 100, 125 e 150 minutos).

As PCRs foram realizadas em placas óticas de 96 poços no equipamento *ABI Real Time PCR System 7500 Fast* (Applied Biosystem, Inc.) com o método de quantificação de  $C_t$  (ciclo limiar - *cycle threshold*) comparativa ( $\Delta\Delta C_t$ ). O corante utilizado foi *SYBR Green*, que possui a propriedade de se intercalar em moléculas de DNA fita dupla (Papin et al., 2004). Para cada reação de 10,0  $\mu$ L, foram utilizados 2,0  $\mu$ L de cDNA fita simples com aproximadamente 50,0 ou 100 ng (diluições de 1:80 e 1:40 respectivamente), 0,5  $\mu$ L de cada oligonucleotídeo (*forward* e reverso) na concentração inicial de 4,0 ou 10,0  $\mu$ M, 2,5  $\mu$ L de *SYBR Green Rox Plus* (LGC Biotecnologia, Inc.) e 4,5  $\mu$ L de água ultrapura.

**Tabela 3.** Relação de Genes e Oligonucleotídeos para PCR em Tempo Real

Símbolo	Nome Genérico	Locus ID	Sequência do Oligonucleotídeo <i>Forward</i> (5'-3')	Sequência do Oligonucleotídeo Reverso (5'-3')	T.A. (pb) <sup>1</sup>	[ ] nM <sup>2</sup>	Eficiência (%)
ACO	GmACO#012	Glyma14g05350	CCA ATG CGC CAT TCC ATT GTT G	TGA GGC TAC GGA CAT TCT GGT C	126	500	82,5±0,8
ACS	GmACS#006	Glyma05g36250	CTC TTA ACC TTC ATT CTT GCT AAC C	TTG CTT CTG CTT CTT TGT ATG C	187	200	86,3±0,3
CTR	GmCTR#003	Glyma10g07610	ATG ATC CCA ATT TCC TCG ATC C	GGA ATC CAT CCC GTG AAT TAG G	188	200	81,4±0,2
ERF069	GmERF#069	Glyma05g32040	TCT GGT TTG AAC AGA GAG AAG G	AGA AGA AGG TGG CAT AGA ACC	198	500	87,8±0,5
ERF223	GmERF#223	Glyma16g32330	TCT CCT CTT CCT TGG ACT CG	CAC CTC ACT TAC CCA CTT ATC C	191	500	83,5±0,4
ETR	GmETR#007	Glyma19g40090	ATG GAT GCC TTC AAG AAG TGG	GCA CAT ATC TTC CCA CAA GAG G	188	500	85,8±0,3
MAT	GmMAT#008	Glyma15g21890	TGA TGA GAA GAC CAT TTT CCA TTT G	CAC AAT GTA GGC ACC ACT CCT A	187	500	81,7±0,3
RAV	GmRAV#001	Glyma01g22260	GTG AGC AGC TAT TCC AGA AGG	CTT TCC CTC CAA CGT CTT CG	180	500	80,9±0,4
–	ACT11 <sup>3</sup>	Glyma18g52780	CGG TGG TTC TAT CTT GGC ATC	GTC TTT CGC TTC AAT AAC CCT A	142	500	84,6±0,5
–	CYP2 <sup>4</sup>	Glyma12g02790	CGG GAC CAG TGT GCT TCT TCA	CCC CTC CAC TAC AAA GGC TCG	154	500	83,3±0,3
–	ELF1A <sup>4</sup>	Glyma05g24110	GAC CTT CTT CGT TTC TCG CA	CGA ACC TCT CAA TCA CAC GC	161	500	84,4±0,3
–	UBC2 <sup>3</sup>	Glyma18g16160	TCC CCT CAC ACC CTT CCT C	CCA TCC CAA GGG GTG TCAT	185	500	83,6±0,3

<sup>1</sup> Tamanho do *amplicon* (pares de base).<sup>2</sup> Concentração final dos oligonucleotídeos utilizados nos experimentos.<sup>3</sup> Genes endógenos utilizados para normalização de amostras foliares.<sup>4</sup> Genes endógenos utilizados para normalização de amostras radiculares.



As diluições do cDNA assim como a concentração final dos oligonucleotídeos para cada gene alvo e endógeno encontram-se na Tabela 3. Já no equipamento, as reações foram incubadas a 95,0 °C por 15 minutos para ativação da enzima, seguida por 40 ciclos de 30 segundos a 95,0 °C (desnaturação), 30 segundos a 60,0 °C (anelamento) e 30 segundos a 72,0 °C (extensão). Os demais parâmetros de reação foram os padrões estabelecidos pelo fabricante. A variação dos valores normalizados ( $\Delta R_n$  - *raw fluorescence data*) gerados nos experimentos foram processados pelo programa *online Real Time PCR Miner* (<http://www.miner.ewindup.info/>) (Zhao & Fernald, 2005), fornecendo parâmetros importantes para posterior análise dos dados (eficiência dos oligonucleotídeos e  $C_t$ , apresentados nas Tabelas 3). O programa *Miner* foi usado para determinar os pontos de início e o fim da fase inicial das PCRs a partir do  $\Delta R_n$ . Ele também estimou a eficiência de amplificação dos oligonucleotídeos através de um algoritmo de regressão linear sem a necessidade de uma curva padrão. Além disso, os valores de ciclo limiar ( $C_t$ ) foram convertidos pelo programa *qBASE versão 1.3.5* (Hellemans et al., 2007) em quantificações relativas normalizadas (NRQ). Os valores foram apresentados em gráficos, todos em função do menor valor de expressão relativa, cujo valor é determinado como 1,0 (um). As análises estatísticas foram feitas com auxílio do programa *REST 2009 (Relative Expression Software Tool - Qiagen, Inc.)* (Pfaffl et al., 2002) de duas formas: a primeira onde foram comparados os valores relativos de expressão dos genes de interesse entre as duas cultivares de soja (BR16 e EMBRAPA48) em um mesmo tecido (raiz ou folha) sob as mesmas condições de estresse; e a segunda onde os valores de expressão dos genes alvo em um mesmo tecido da planta não estressada (tempo zero) foram confrontados com os da estressada (demais tempos de déficit hídrico) de uma mesma cultivar.

## 5. RESULTADOS.

### 5.1. IDENTIFICAÇÃO IN SILICO E ONTOLOGIA DOS GENES.

Foram analisados um total de 957 genes, sendo 455 deles relacionados com as plantas modelo (221 referentes à *A. thaliana* e 234 à *O. sativa*) e 502 com *Glycine max*. Estes genes foram divididos em três grupos principais: *biossíntese de etileno*, *transdução de sinal mediada por este fitormônio* e *membros da superfamília de fatores de transcrição AP2/ERF* (famílias AP2, ERF e RAV) (Tabela 4). Os genes de biossíntese de etileno em *A. thaliana*, *O. sativa* e *Glycine max* correspondem respectivamente à 19,91 %, 16,24 % e 21,51 % em relação ao total de genes analisados para cada espécie vegetal. Já os genes relativos à transdução de sinal mediada por este fitormônio correspondem a 13,57 %, 14,53 % e 13,55 % do total de genes analisados em *A. thaliana*, *O. sativa* e *Glycine max* respectivamente. Por fim, os membros da superfamília de fatores de transcrição AP2/ERF representam 66,52 % dos genes analisados de *A. thaliana*, 69,23 % de *O. sativa* e 64,94 % de soja. Desta forma, observa-se que a variação relativa da quantidade de genes nas classes analisadas pouco varia entre as espécies em questão.

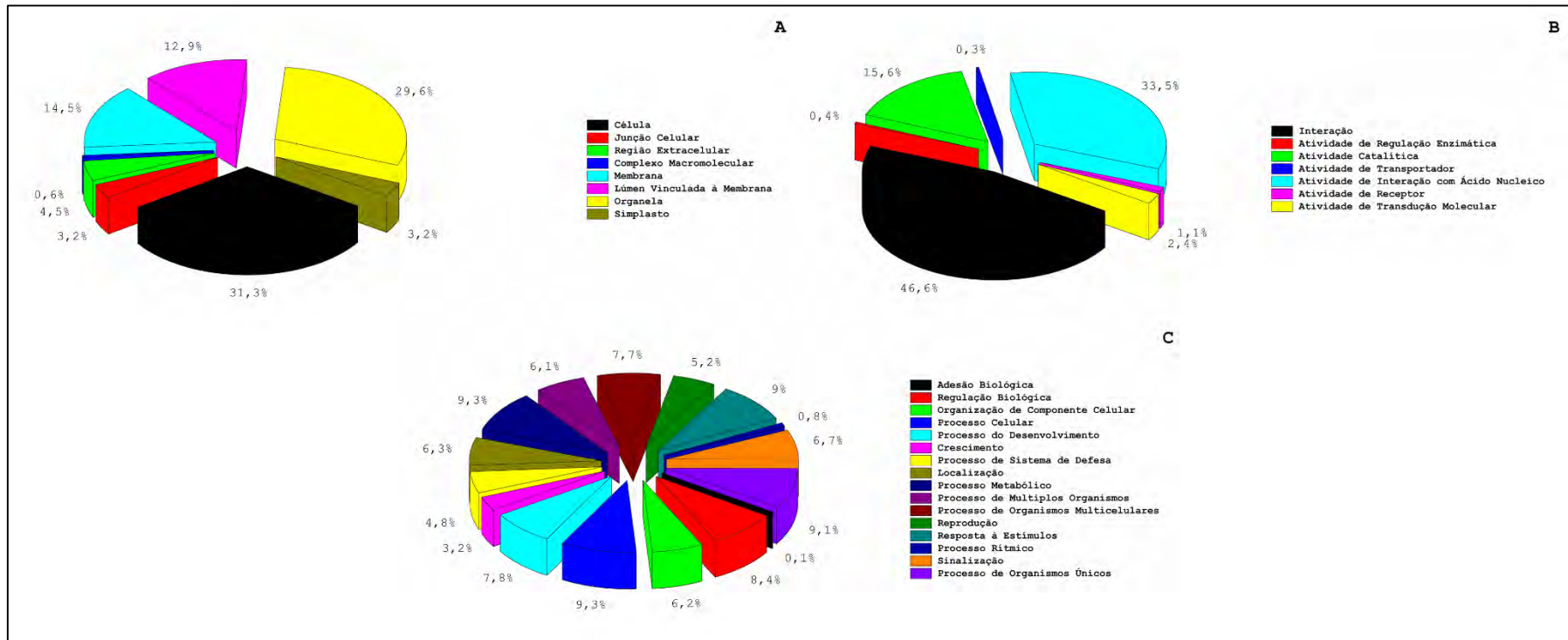
**Tabela 4.** Sumário dos Genes Analisados em Diferentes Plantas

Grupo	Número de Genes		
	<i>Arabidopsis thaliana</i>	<i>Glycine max</i>	<i>Oryza sativa</i>
Biossíntese	44	108	38
Transdução de Sinal	30	68	34
Superfamília AP2/ERF	147	326	162
<b>Total</b>	<b>221</b>	<b>502</b>	<b>234</b>

A organização dos 502 genes analisados no genoma da soja foi representada na Figura S1, sendo que a maioria deles encontra-se geralmente nas extremidades de um dos 20 cromossomos com alta taxa de duplicação. Com relação à localização na fita de DNA, 49,60 % deles se localizam na fita positiva (senso) e 50,40 % na fita negativa (anti-senso). Quanto ao número de genes por cromossomo, em valores absolutos, o cromossomo 13 é o

que apresenta o maior número de genes identificados (41 genes), juntamente com o maior número de representantes da superfamília de fatores de transcrição AP2/ERF (28 genes). Já os cromossomos 18 e 3 apresentam, respectivamente, o maior número de genes relacionados com a biossíntese de etileno (10 genes) e com a transdução de sinal mediada por este fitormônio (7 genes). Além disso, os cromossomos 15 e 16 não possuem representantes do grupo relativo à transdução de sinal. Após normalização do número de genes com o tamanho de cada cromossomo (em megabases -  $10^6$  pares de bases), observou-se que a maior densidade de genes relativos à biossíntese de etileno encontra-se nos cromossomos 13 e 18, assim como os cromossomos 3 e 13 apresentam respectivamente a maior densidade de genes relativa à transdução de sinal e fatores de transcrição da superfamília AP2/ERF.

Com relação à ontologia dos 502 genes de soja, após a anotação funcional, os resultados foram agrupados em três categorias: *componentes celulares*, *função molecular* e *processos biológicos* (Figura 4). Em relação aos componentes celulares, foram encontrados 83 *GO terms* em 11 níveis, sendo que para o nível 2 (Figura 4-A), 31,3 % foram encontrados na célula, 29,6 % na organela, 14,5% na membrana, 12,9 % no lúmen vinculado à membrana, 4,5 % na região extracelular, 3,2 % com junção celular e simplasto e 0,6 % com complexo macromolecular. Já para a função molecular, foram encontrados 199 *GO terms* em 7 níveis, sendo que para o nível 2 (Figura 4-B), 46,6 % estão relacionados com interação em geral, 33,5 % com interação com ácidos nucleicos, 15,6 % com atividade catalítica, 2,4 % com atividade de transdução molecular, 1,1 % com atividade de receptor, 0,4 % com atividade de regulação enzimática e 0,3 % relacionados com transporte. Por fim, para processos biológicos, foram encontrados 1.319 *GO terms* em 14 níveis. No nível 2 (Figura 4-C), a maior parte dos genes está relacionada com processos metabólicos e celulares (9,3 % cada), seguidos por processo de organismos únicos (9,1 %), resposta à estímulos (9,0 %), regulação biológica (8,4 %), processos de desenvolvimento e organismos multicelulares (7,8 % e 7,7 %, respectivamente), sinalização (6,7 %), localização (6,3 %), organização de componente celular (6,2 %), processo de múltiplos organismos (6,1 %), reprodução (5,2 %), processo de sistema de defesa (4,8 %), crescimento (3,2 %), processo rítmico (0,8 %) e adesão biológica (0,1 %). Algumas



**Figura 4. Classificação Ontológica das Sequências de Soja Analisadas.** Os três gráficos apresentam os subgrupos ontológicos (nível 2) em que as 502 sequências de soja selecionadas estão inseridas. *Legendas:* A - componente celular; B - função molecular; C - processo biológico

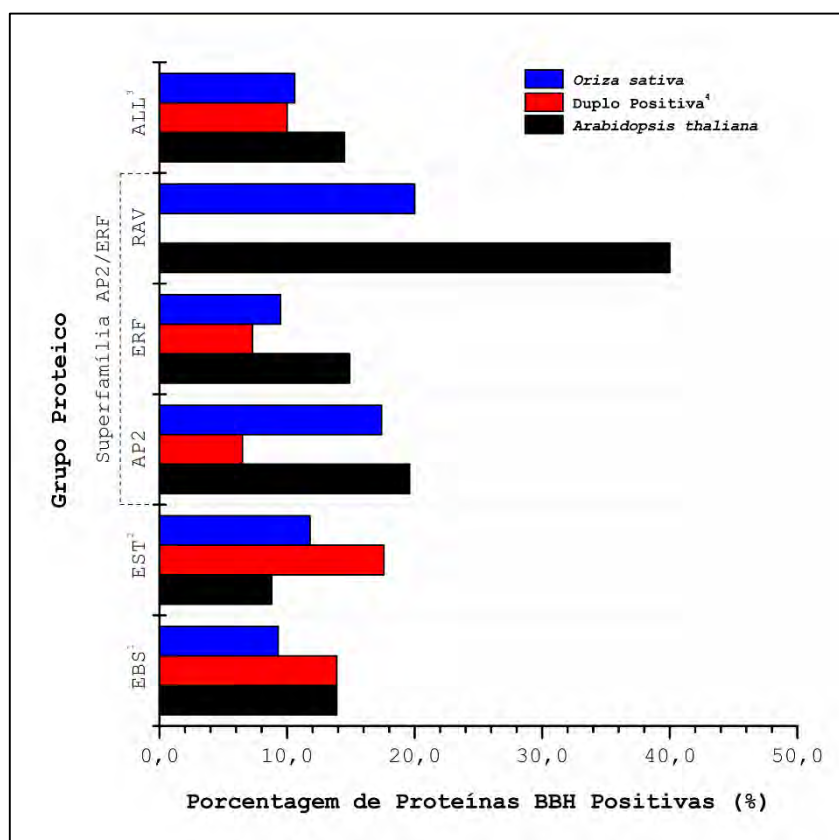
sequências foram classificadas em mais de uma categoria, sendo assim, a soma das sequências nas categorias excedeu o número total de 502.

Em seguida, foram realizadas análises das proteínas codificadas pelos genes de *Arabidopsis thaliana* (Tabelas S1 à S5), *Oryza sativa* (Tabelas S6 à S10) e *Glycine max* (S11 à S15), onde foram identificadas as principais características de cada sequência de resíduos de aminoácidos predita (tamanho e presença/localização de domínios proteicos conservados). Em relação à soja, também foram identificadas as sequências homólogas com *Arabidopsis thaliana* e *Oryza sativa*, o que propiciou a inferência de possíveis funções. Nas três espécies vegetais analisadas, foram identificados segundo o banco de dados *PFAM* um total de 36 domínios proteicos conservados, sendo 15 deles em proteínas do grupo relacionado com a biossíntese de etileno, 19 em representantes da transdução de sinal mediada por este fitormônio e dois domínios em proteínas da superfamília AP2/ERF.

## **5.2. EXPERIMENTO DE BBH.**

Com a finalidade de identificar possíveis proteínas de soja ortólogas à proteínas de *A. thaliana* e *O. sativa*, foi realizado o experimento de *melhor “hit” bidirecional (Best Bidirectional Hit)* ou *BBH*, comparando os bancos de dados das três espécies vegetais. Dentre as proteínas codificadas pelos 502 genes de soja analisados, foram identificados 176 eventos BBH positivos com *A. thaliana* e/ou *O. sativa*, sendo 23,73 % relacionados com a biossíntese de etileno, 14,77 % com a transdução de sinal mediada por este fitormônio e 62,5 % com membros da superfamília de fatores de transcrição AP2/ERF (11,36 % com a família AP2; 49,43 % com a família ERF; e 1,71 % com a família RAV) (Tabelas S16 à S20). Com relação ao total de proteínas de soja identificadas, 13,90 % das relacionadas com a biossíntese do fitormônio gasoso são BBH positivas com proteínas de *A. thaliana*, 9,30 % com *O. sativa* e 13,90 % com as duas espécies vegetais. As proteínas de soja relacionadas com a transdução de sinal mediada por etileno apresentam 8,80 % de seu total BBH positivas com *A. thaliana*, 11,80 % com *O. sativa* e 17,6 % com as duas espécies vegetais. Por fim, os membros da superfamília AP2/ERF de fatores de transcrição de soja apresentam 15,95 % de seu total BBH positivas com *A. thaliana* (19,60 % dos membros da família AP2; 14,90 % da família ERF; e 40,00 % dos membros da família

RAV), 10,74 % com proteínas de *O. sativa* (17,40 % dos membros da família AP2; 9,50 % da família ERF; e 20,00 % dos membros da família RAV), e 7,06 % com as duas espécies vegetais (6,50 % dos membros da família AP2; 7,30 % da família ERF; e nenhum membro da família RAV) (Figura 5). Desta forma, com exceção do grupo relativo referente à proteínas relacionadas com a transdução de sinal mediada por etileno, todos os grupos proteicos de soja analisados apresentaram um maior número de possíveis ortólogos com *A. thaliana*, quando comparados à *O. sativa*, indicando uma maior proximidade filogenética das proteínas de soja analisadas, e, possivelmente das referidas rotas metabólicas, com a planta modelo *A. thaliana*.



**Figura 5. Experimento de Melhor “Hit” Bidirecional (BBH).** O gráfico exibe a porcentagem de proteínas de soja BBH positivas presentes em cada grupo proteico analisado. <sup>1</sup> EBS e <sup>2</sup> EST - proteínas relacionadas à biossíntese e transdução de sinal mediada por etileno, respectivamente; <sup>3</sup> ALL - percentual geral (considerando todas as proteínas de soja analisadas). <sup>4</sup> Duplo Positiva - proteínas de soja BBH positivas com *Arabidopsis thaliana* e *Oryza sativa* simultaneamente. As proteínas identificadas neste experimento *in silico* encontram-se listadas nas Tabelas S18 - S22.

### **5.3. PREDIÇÃO DE MODELOS *IN SILICO* EM SOJA.**

#### **5.3.1. Biossíntese de Etileno.**

Os tecidos vegetais produzem etileno a partir de metionina por uma série de reações enzimáticas, onde a S-adenosil-metionina (AdoMet ou SAM), fornecida pelo ciclo de Yang, é convertida à ACC (1-aminociclopropano-1-ácido carboxílico) por ação da enzima ACC sintase (ACS), que por sua vez é convertido à etileno por ação da enzima ACC oxidase (ACO) (Yang & Hoffman, 1984). Desta forma, com base no modelo proposto para a biossíntese de etileno em *A. thaliana*, os 108 genes de soja relacionados com esta rota metabólica, foram divididos em três subgrupos: ciclo de Yang, biossíntese de etileno propriamente dita e degradação/conjugação de ACC (Figura 6). A partir desta subdivisão, 21,30 % dos genes deste grupo estão associados com o ciclo de Yang, 44,44 % com a biossíntese de etileno propriamente dita (genes *MAT*, *ACS* e *ACO*) e 34,26 % com a degradação/conjugação de ACC [conjugação com N-malonil (genes *ACT*),  $\gamma$ -glutamil (genes *GGT*) e degradação (genes *ACD*)] (Tabela S11).

Pommerrenig e colaboradores (2011), apresentaram um possível modelo para as reações de reciclagem da metionina através do ciclo de Yang em *Plantago* e *A. thaliana* (Pommerrenig et al., 2011). Com base nesse neste trabalho, foi proposto para soja um modelo *in silico* para esta rota, onde foram identificados homólogos para todas as enzimas que a compõe: *MTN* (5-metil-tio-adenosina nucleosidase), *MTK* (5-metil-tio-ribose cinase), *MTI* (5-metil-tio-ribose-1-fosfato isomerase), *DEP* (complexo desidratase-enolase-fosfatase), *ARD* (aciredutona dioxigenase), e *AAT* (aminoácido transferase) ou *ASP* (aspartato aminotransferase) (Figura 6).

Com relação à biossíntese de etileno propriamente dita, foram identificadas 11 homólogos em soja relacionados com a primeira enzima desta rota em *A. thaliana*, enzima esta denominada de *MAT* (metionina-adenosil transferase) ou *SAMS* (S-adenosil-metionina sintetase) e é responsável pela produção de AdoMet tanto para biossíntese de etileno quanto para outras rotas metabólicas, como biossíntese de ligninas e poliaminas (Yang & Hoffman, 1984; Amthor, 2003).

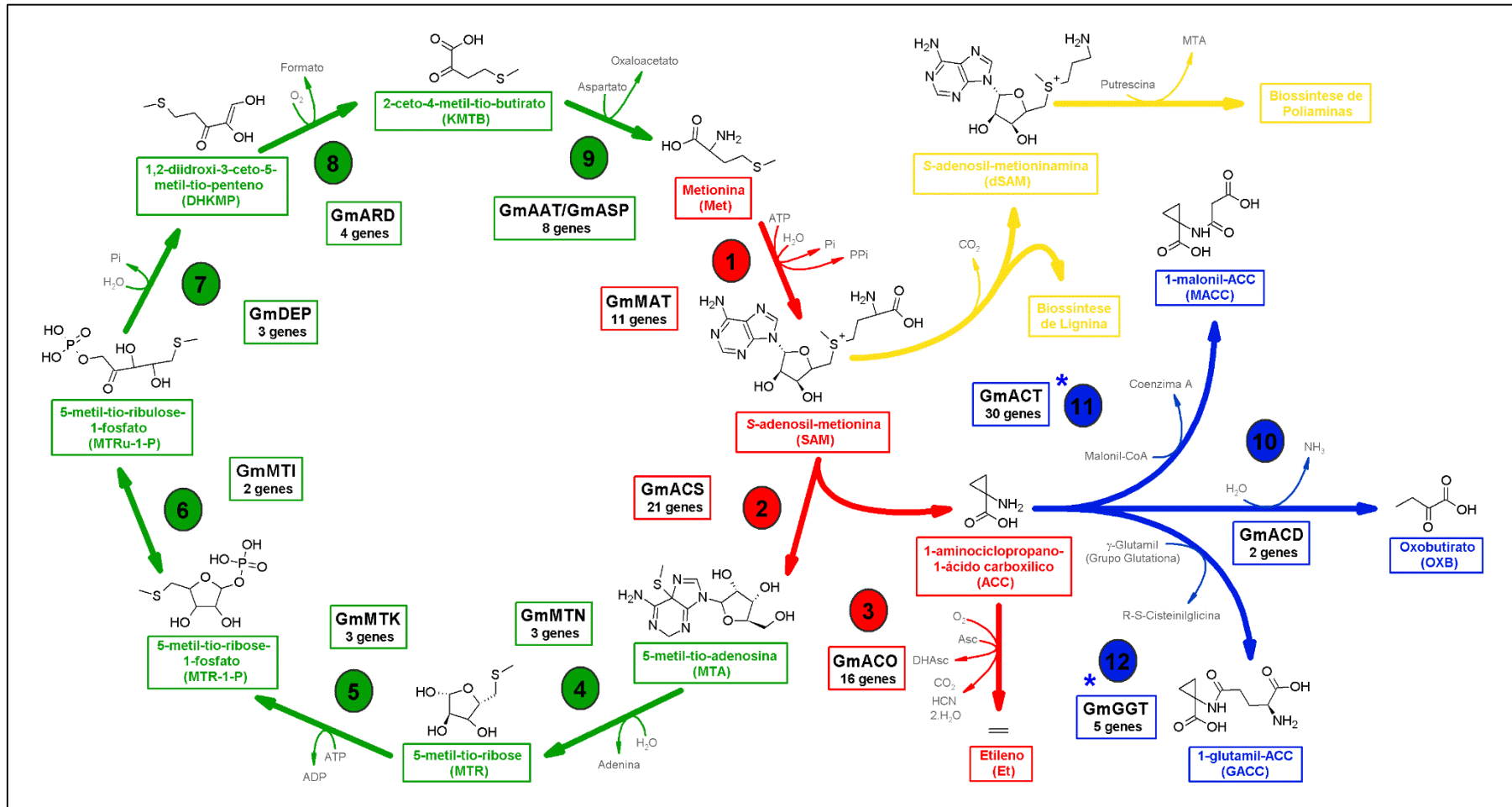


Figura 6. Modelo Proposto para Biossíntese de Etileno em Soja.



**Figura 6 (cont.). Modelo Proposto para Biossíntese de Etileno em Soja.** Experimentos *in silico* identificaram 108 proteínas que podem estar envolvidas direta ou indiretamente com a biossíntese de etileno em soja. No modelo pode-se observar: **verde** - ciclo de Yang; **vermelho** - biossíntese de etileno; **azul** - degradação e conjugação de ACC (1-aminociclopropano-1-ácido carboxílico) com outros metabólitos (grupamentos malonil e glutamyl); **amarelo** - biossíntese de poliamina/lignina (exemplo de desvio da produção de S-adenosil-metionina para outras rotas metabólicas). Enzimas: **1** - *MAT* (metionina-adenosil transferase) ou *SAMS* (S-adenosil-metionina sintetase); **2** - *ACS* (1-aminociclopropano-1-ácido carboxílico sintase); **3** - *ACO* (1-aminociclopropano-1-ácido carboxílico oxidase); **4** - *MTN* (5-metil-tio-adenosina nucleosidase); **5** - *MTK* (5-metil-tio-ribose cinase); **6** - *MTI* (5-metil-tio-ribose-1-fosfato isomerase); **7** - *DEP* (complexo desidratase-enolase-fosfatase); **8** - *ARD* (aciredutona dioxigenase); **9** - *AAT* (aminoácido transferase) ou *ASP* (aspartato aminotransferase); **10** - *ACD* (1-aminociclopropano-1-ácido carboxílico deaminase); **11** - *ACT* (aciltransferase; N-malonil transferase); **12** - *GGT* ( $\gamma$ -glutamyl transpeptidase). Outras siglas: *Asc* - ascorbato; *DHAsc* - diidroxiascorbato; *HCN* - cianeto de hidrogênio. Os **asteriscos azuis (\*)** presentes nos números **11** e **12** indicam enzimas que podem ser candidatas a desempenhar a função descrita no modelo, mas suas funções descritas *in vitro* e *in vivo* não estão associadas primariamente com esta rota metabólica. Cada enzima é representada por um nome genérico (ver Tabela S11).

Dentre as três famílias de enzimas que participam da biossíntese propriamente dita de etileno, a mais estudada é a responsável pela produção do precursor imediato do fitormônio, a *1-aminociclopropano-1-ácido carboxílico sintase* (ACS), que converte S-adenosil-metionina em 1-aminociclopropano-1-ácido carboxílico (ACC). Pelo fato desta enzima pertencer à uma família multigênicas e pela presença de sítios de fosforilação na porção C-terminal de alguns de seus membros, Liu e colaboradores (2004) propuseram para *A. thaliana* uma classificação baseada nestas possíveis modificações pós traducionais detectadas nesta porção. São classificadas como *ACS do tipo I*, as que possuem em seu C-terminal sítios de fosforilação para as enzimas cinases dependentes de cálcio (CPK ou CDPK) e para proteínas MAP cinases, semelhantes à AtCDPK2 e AtMPK6 de *A. thaliana*, respectivamente; *ACSs do tipo II* as que apresentam na mesma região somente sítios de fosforilação somente para CPKs; e *ACSs do tipo III* as que não possuem sítio de fosforilação em seu C-terminal. (Liu & Zhang, 2004). Posteriormente, Tucker e colaboradores (2010) propuseram tal classificação para as 21 sequências que apresentam similaridade com ACSs identificadas em soja, mostrando uma possível relação filogenética das mesmas com as de *A. thaliana*, além de relacionar sua expressão com a infecção ocasionada pelo nematoide *Heterodera glycines* (Tucker et al., 2010). O presente trabalho apresenta uma divisão da sequência de resíduos de aminoácidos das possíveis ACSs identificadas em soja, correlacionando-as filogeneticamente não somente com *A. thaliana*, mas também com seus homólogos em *O. sativa*, além de inferir *in silico* os possíveis sítios de fosforilação pelas respectivas cinases (Figura 7).

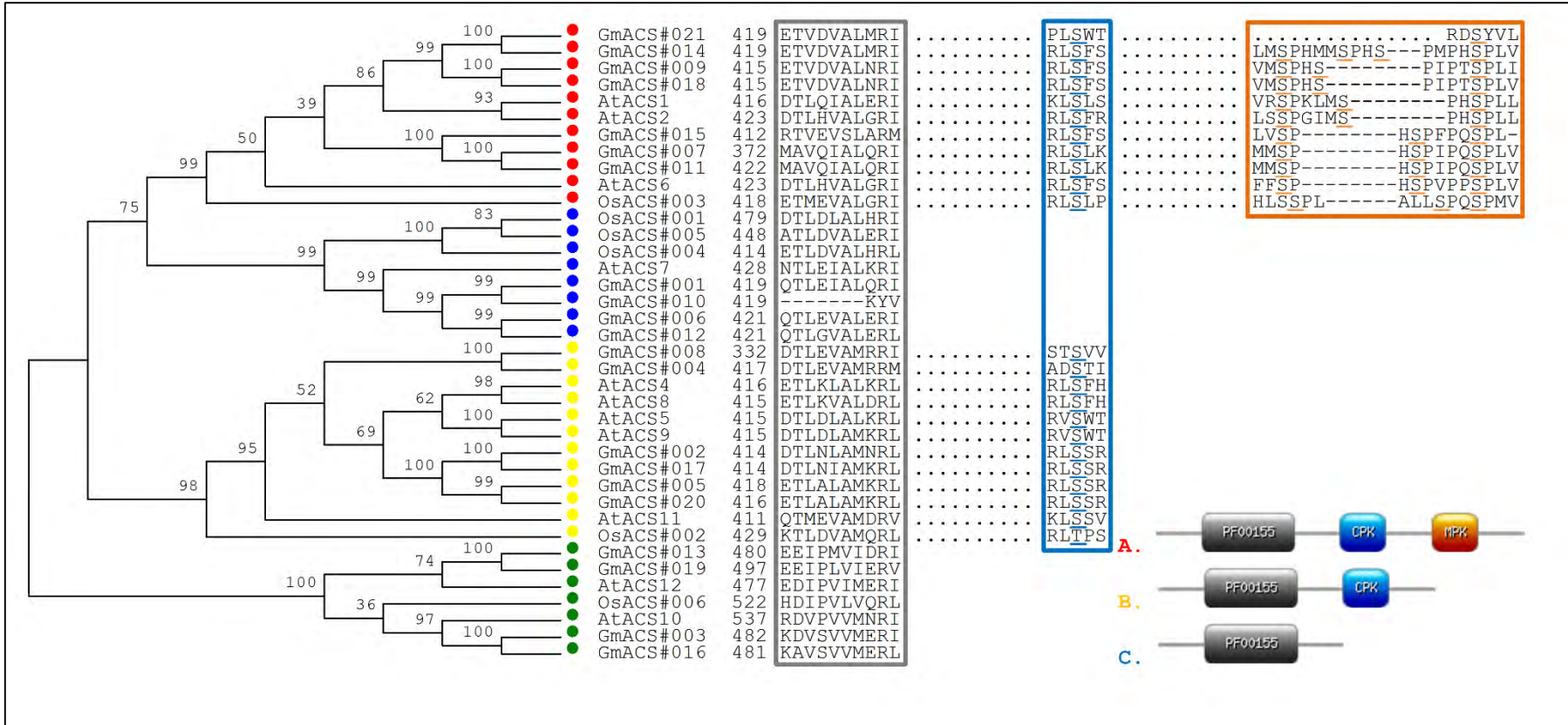


Figura 7. Classificação das ACSs de *Arabidopsis thaliana*, *Glycine max* e *Oryza sativa*.

**Figura 7. Classificação das ACSs de *Arabidopsis thaliana*, *Glycine max* e *Oryza sativa*.** A figura apresenta a clusterização existente entre as sequências de aminoácidos de ACSs (1-aminociclopropano-1-ácido carboxílico sintase) de soja com *Arabidopsis thaliana* e *Oryza sativa*. Esta relação possibilita a classificação de acordo com a presença/ausência de possíveis sítios de fosforilação por proteínas cinase cálcio dependente (CDPK ou CPK) e/ou proteínas MPK6 (proteína cinase ativada por mitógeno 6 - MAPK6) na extremidade C-terminal destas proteínas. Desta forma, estas sequências de aminoácidos podem ser divididas em três classes: *tipo I* (**círculos vermelhos; modelo A**) - presença de sítios de fosforilação por CPK e MPK6; *tipo II* (**círculos amarelos; modelo B**) - presença de sítios de fosforilação somente para CPK; e, *tipo III* (**círculos azuis; modelo C**) - ausência de sítios de fosforilação para ambas as cinases (Liu & Zhang, 2004; Tucker et al., 2010). As sequências representadas pelos **círculos verdes** são classificadas como semelhantes à ACSs, uma vez que AtACS10 e AtACS12 provavelmente não possuem atividade ACS e podem também ser classificadas como possíveis aminoácido transferase (AAT) (Yamagami et al., 2003). AtACS1 não possui atividade ACS por possuir deleções em seu centro catalítico, mas AtACS2 possui (Yamagami et al., 2003). O **retângulo cinza** destaca a porção C-terminal do centro catalítico (posição à esquerda), o **retângulo azul** o sítio de fosforilação por CPK e o **retângulo laranja** por MPK6. Os **resíduos de aminoácidos sublinhados** são os com maior probabilidade de serem fosforilados em cada sequência. Cada proteína é representada por um nome genérico (ver Tabelas S1, S6 e S11).

Observa-se uma divisão das sequências de resíduos de aminoácidos semelhante a apresentada por Tucker, com sequências das três espécies distribuídas uniformemente entre os três tipos de ACSs em questão, o que indica um certo nível de conservação evolutiva desta enzima nas espécies vegetais em questão. Além disso, algumas sequências (GmACS#003, GmACS#013, GmACS#016 e GmACS#019), mesmo apresentando alta similaridade com sequências de ACS, encontraram-se separadas filogeneticamente das demais, onde o principal ponto de diferença foi detectado no domínio catalítico. Desta forma, essas sequências foram classificadas como sendo semelhantes à ACS (*ACS-like*), pertencentes à família das enzimas classificadas como *aminoácido transferases* (AATs), família em que as ACSs estão inseridas.

Após à síntese de ACC, o mesmo é convertido à etileno pela ação da enzima ACC oxidase (ACO). No genoma de soja, foram identificados 16 genes referentes à esta enzima, dentre os quais 6 codificam proteínas ortólogas com *A. thaliana* e /ou *O. sativa* (GmACO#004, GmACO#006, GmACO#007, GmACO#008, GmACO#009 e GmACO#014) (Tabela S16).

Além de ser substrato para a biossíntese de etileno, o ACC pode ser utilizado em reações de conjugação com grupamentos malonil e glutamyl, gerando respectivamente malonil-ACC (MACC) e glutamyl-ACC (GACC) (Kende, 1993; Martin et al., 1995). A enzima responsável pela formação de MACC é a chamada *1-aminociclopropano-1-ácido carboxílico-N-malonil transferase*, identificada em extratos proteicos vegetais, mas que

não teve seu respectivo gene ainda caracterizado em plantas modelo. Desta forma, foram selecionados trinta possíveis genes candidatos à esta função em soja, tendo utilizado como modelo seis *aciltransferases* (ACT) de *A. thaliana* que foram descritas com atividades de *malonil transferases* (Tabelas S1 e S11). Deste total, 5 destas enzimas foram consideradas BBH positivos com *A. thaliana* e/ou *O. sativa* (GmACT#003, GmACT#006, GmACT#017, GmACT#020 e GmACT#023) (Tabela S16). Vale a pena ressaltar que mesmo que a maioria das enzimas com atividade de malonil transferase possuam papel na biossíntese de ácidos graxos, pode-se imaginar que algum de seus representantes possa vir a ter atividade de N-malonização o que, após este processo de seleção *in silico*, precisa ser profundamente analisado em experimentos tanto *in vitro* quanto *in vivo*. Com relação à formação de GACC, foram identificadas *in silico* cinco enzimas *γ-glutamil transpeptidase* (GGT) em soja, capazes de participar desta reação de conjugação, sendo somente duas (GmGGT#001 e GmGGT#003) BBH positivas com *Arabidopsis thaliana* e *Oryza sativa* (Tabela S16).

Por fim, o ACC pode também ser substrato da enzima *ACC deaminase* (ACD), que converte irreversivelmente o precursor de etileno em oxobutirato (Figura 7) (McDonnell et al., 2009). Em soja, foram identificados dois genes que codificam enzimas homólogas à ACD de *A. thaliana* (GmACD#001 e GmACD#002), sendo somente uma considerada BBH positiva (GmACD#001) (Tabela S16).

### **5.3.2. Transdução de Sinal Mediada por Etileno.**

Semelhante à outros sistemas de transdução de sinal, o etileno em plantas é reconhecido por receptores específicos que desencadeiam mecanismos moleculares envolvendo cascatas com proteínas cinases e regulatórias que culminam em respostas específicas, como a regulação transcricional de genes alvo. Em *A. thaliana*, o etileno é percebido por uma família de cinco receptores que regulam negativamente as rotas de sinalização mediada por este fitormônio (Hua & Meyerowitz, 1998; Wang et al., 2003; Qu & Schaller, 2004). Na ausência de etileno, os receptores ativam CTR1, uma cinase serina/treonina que suprime a resposta por este gás (Kieber et al., 1993; Clark et al., 1998). O alvo direto de fosforilação de CTR1 é EIN2, uma proteína com similaridade à

transportadores Nramp, que é mantida em um estado inativo quando fosforilado por CTR1 (Alonso et al., 1999; Ju & Chang, 2012; Qiao et al., 2012). Na presença de etileno, os receptores inativam CTR1, atenuando assim a supressão em elementos de sinalização à jusante. Como resultado, EIN2 sofre proteólise em seu domínio C-terminal, que após ser liberado migra até o núcleo por mecanismos ainda não conhecidos (Ju & Chang, 2012; Qiao et al., 2012; Wen et al., 2012), ativando direta ou indiretamente os fatores de transcrição EIN3/EIL, que por sua vez iniciam a resposta transcricional mediada por etileno (Chao et al., 1997; Solano et al., 1998; Alonso et al., 2003).

No presente trabalho foram identificados 68 genes relacionados com a transdução de sinal mediada por etileno em soja, sendo que 38,24 % das proteínas por eles codificadas são consideradas ortólogas com *A. thaliana* e/ou *O. sativa* por experimento de BBH, representando a maioria dos componentes analisados da rota (Tabela S17). Do total de genes relativos à transdução de sinal mediada por etileno em soja, 32,38 % correspondem à receptores específicos (ETR) e à proteínas importantes para sua atividade (RAN e RTE), 7,35 % com proteínas CTR, 4,41 % com proteínas EIN2, 19,11 % com proteínas cinases (CPK, MKK e MPK), 24,99 % com proteínas *F-box*, importantes em processos proteolíticos via proteassoma 26S (ETO/EOL e EBF), 7,35 % com os fatores de transcrição EIN3/EIL e 4,41 % com a exoribonuclease EIN5, importante na degradação dos transcritos de EIN3/EIL (Figura 8).

Dentre os cinco receptores para etileno descritos em *A. thaliana* (ETR1, ETR2, ERS1, ERS2, EIN4), somente quatro foram identificados em soja, sendo eles homólogos à ETR1 e ETR2 (pertencentes à subfamília I - GmETR#001, GmETR#003, GmETR#006 e GmETR#007), e homólogos à ERS1 e EIN4 (subfamília II - GmETR#002, GmETR#004, GmETR#005, GmETR#008, GmETR#009, GmETR#010 e GmETR#011) (Chang et al., 1993; Hua et al., 1995; Hua et al., 1998; Sakai et al., 1998). Semelhante à mesma planta modelo, os receptores identificados em soja são compostos por quatro principais domínios: (i) *domínio regulador de resposta de receptor* (PF00072); (ii) *domínio histidina cinase A* (fosfoceptor) (PF00512); (iii) *domínio GAF* (PF01590); e (iv) *histidina cinase, DNA girase B e semelhante à ATPase* (PF02518). Foi observado que a combinação destes domínios caracteriza as famílias de receptores identificadas, sendo que os homólogos à ETR1 apresentam os quatro domínios em sua estrutura, os homólogos à

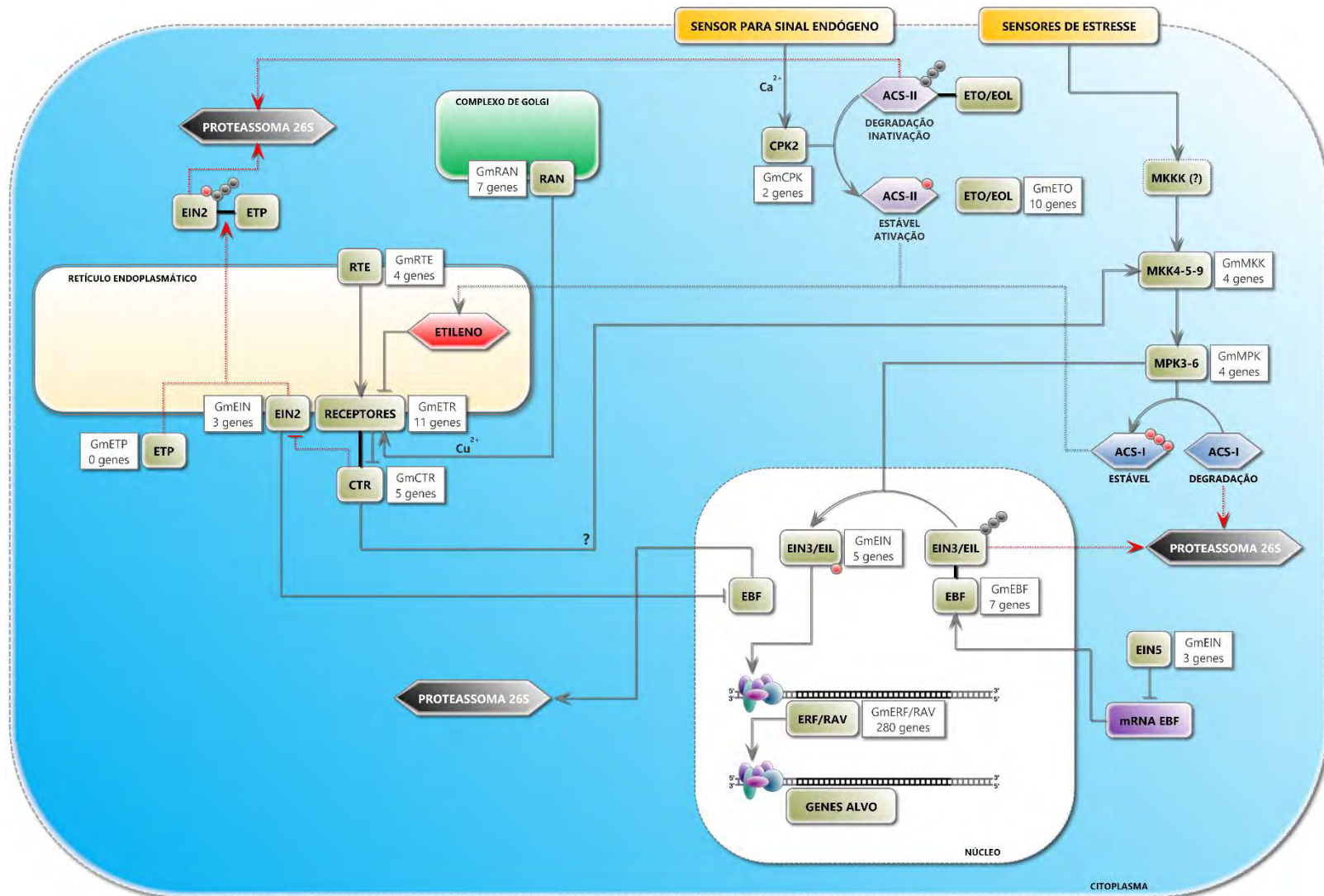


Figura 8. Modelo Proposto para Transdução de Sinal Mediada por Etileno em Soja.

**Figura 8 (cont.). Modelo Proposto para Transdução de Sinal Mediada por Etileno em Soja.**

Experimentos *in silico* identificaram 68 proteínas envolvidas com a transdução de sinal mediada por etileno em soja. Os **retângulos marrons** exibem proteínas da rota identificadas em *Arabidopsis thaliana* e nos **retângulos brancos** encontram-se os genes de soja que codificam proteínas homólogas a esta planta modelo; os **retângulos laranjas** ilustram sensores de membrana que respondem a estresse biótico e abiótico e receptores para sinais endógenos (i.e. outros fitormônios); o **retângulo roxo** representam mRNAs referentes à proteínas ETP; o **retângulo com o contorno pontilhado** (acompanhado de ponto de interrogação) representa proteína na rota que ainda não foi identificada em nenhuma das plantas estudadas; os **hexágonos azuis e roxos** representam as enzimas ACSs do tipo I e do tipo II respectivamente; os **círculos pretos e vermelhos** correspondem à ubiquitina e grupamentos fosfato respectivamente; as **setas cinza** correspondem às rotas que ocorrem na presença de etileno e/ou a sinais ambientais de estresse biótico/abiótico; as **setas pontilhadas em vermelho e cinza** representam respectivamente as vias que intercorrem na ausência do fitormônio e rotas que culminam na biossíntese de etileno; as **linhas pretas** indicam interação entre proteínas. Os compartimentos celulares representados são: **retículo endoplasmático (bege)**, **complexo de Golgi (verde)**, **núcleo (branco)** e **citoplasma (azul)**. Siglas: **ACS**: 1-aminociclopropano-1-ácido carboxílico sintase; **CPK** (ou **CDPK**): proteína cinase cálcio dependente; **CTR**: proteína constitutiva de resposta tripla; **EBF**: proteína *F-box* de ligação à EIN3; **EIL**: proteína semelhante à EIN; **EIN**: proteína insensível a etileno; **EOL**: proteína semelhante à ETO; **ERF**: fator de transcrição responsivo à etileno; **ETP**: proteína-alvo de EIN2; **ETO**: proteína relacionada à superprodução de etileno; **MKKK** (ou **MAPKKK**): proteína MAP cinase cinase; **MKK** (ou **MAPKK**): proteína MAP cinase; **MPK** (ou **MAPK**): proteína cinase ativada por mitógeno; **RAN**: proteína responsiva à agonista; **RAV**: proteína relacionada à ABI3/VP1; **RTE**: proteína que reverte a sensibilidade à etileno. A rota intracelular da transdução de sinal é iniciada pela interação do etileno com receptores de membrana (codificados por genes do tipo ETR) e, por meio da modulação da atividade de CTR, regula a atividade e expressão de diversos genes, como EIN3. Os receptores juntamente com CTR (proteína cinase semelhante à RAF - MKKK) funcionam como reguladores negativos da via e, na ausência de etileno, reprimem componentes positivos a jusante da transdução de sinal. A ligação do hormônio bloqueia os receptores em uma conformação inativa, reduzindo a repressão dos reguladores positivos da rota metabólica (Bleecker, 1999). Na ausência de etileno, CTR fosforila o domínio C-terminal de EIN2, propiciando sua interação com as proteínas *F-box* ETP (não identificadas em soja) e sua subsequente degradação via proteassoma 26S (Qiao et al., 2009). Na ausência da fosforilação do C-terminal de EIN2 (presença do hormônio), este domínio é clivado e se desloca para o núcleo, onde estimula a atividade de EIN3/EIL pela repressão de EBF (estimula a degradação desta proteína *F-box* por mecanismos ainda desconhecidos), que por sua vez induzem a transcrição de genes alvo, por intermédio de alguns membros da superfamília AP2/ERF de fatores de transcrição (Ju et al., 2012). Além da interação com o C-terminal de EIN2, a atividade de EIN3/EIL pode ser influenciada pela cascata de fosforilação MKK4-5-9 → MPK3-6, independente de CTR e EIN2. Na presença de um sinal, os fatores de transcrição EIN3/EIL podem ser fosforilados por MPK3-6, deixando assim de interagir com a proteína *F-box* EBF, o que impede a degradação destes fatores de transcrição via proteassoma 26S. Desta forma, estes fatores se acumulam no núcleo, interagem com promotores de genes alvo e desencadeiam diversas respostas relacionadas ao etileno (Stepanova & Alonso 2009). Outro regulador positivo da via é EIN5, uma exoribonuclease 5'-3' que promove o decaimento do mRNA dos genes de EBF, aumentando os níveis das proteínas EIN3/EIL no núcleo (Olmedo et al., 2006). Além disso, a biossíntese de etileno também é regulada. Possíveis receptores para sinais endógenos (i.e. outros fitormônios) podem induzir o acúmulo de metabólitos secundários (i.e. cálcio) no meio intracelular, ativando assim proteínas cinases (i.e. CPK2), o que culmina na estabilização de ACSs do tipo II, importante enzima da biossíntese de etileno. Desta forma, as ACSs do tipo II (em *A. thaliana* AtACS5 e AtACS9) são fosforiladas por CPK2, o que impede a interação destas enzimas com ETO/EOL e a subsequente degradação pelo proteassoma 26S, o que desta forma induz o aumento na produção de etileno (Ecker, 2004). Além disso, diversos tipos de estresse (bióticos e abióticos) induzem a ativação de módulos de MAPKs (em *A. thaliana* MKK4-5-9 e MPK3-6). As cinases MPK3 e MPK6 são capazes de fosforilar o C-terminal de ACSs do tipo I (em *A. thaliana* AtACS2 e AtACS6), fato este que as estabiliza e as preserva da degradação via proteassoma 26S, aumentando desta forma a produção de etileno e induzindo outras vias de sinalização dependentes deste fitormônio (Hahn et al., 2009). Não existe consenso com relação à participação direta de CTR na rota envolvendo MPK3-6 (Vandenbussche et al., 2012). A atividade dos receptores está relacionada a duas proteínas: RAN, proteína transportadora de íons cobre, um cofator importante na atividade dos receptores (Binder et al., 2010); e RTE, proteína com mecanismo de ação desconhecido, mas importante na transição de um dos receptores (ETR1) do seu estado inativo para o nativo (Dong et al., 2008;



Stepanova & Alonso 2009). Cada proteína é representada por um nome genérico: *EIN2*: GmEIN#002, GmEIN#004 e GmEIN#007; *EIN3*: GmEIN#001, GmEIN#005, GmEIN#006, GmEIN#008 e GmEIN#010; *EIN5*: GmEIN#003, GmEIN#009 e GmEIN#011; *MKK4*: GmMKK#001 e GmMKK#003; *MKK5*: sem representantes identificados em soja; *MKK9*: GmMKK#002 e GmMKK#004; *MPK3*: GmMPK#003 e GmMPK#004; *MPK6*: GmMPK#001 e GmMPK#002; *Receptores*: *EIN4*: GmETR#002, GmETR#004, GmETR#008 e GmETR#011; *ERS1*: GmETR#001 e GmETR#007; *ERS2*: sem representantes identificados em soja; *ETR1*: GmETR#003 e GmETR#006; *ETR2*: GmETR#005, GmETR#009 e GmETR#010. Para mais informações ver Tabelas S2, S7 e S12. A classificação das enzimas ACSs encontra-se na Figura 7 (ver também Tabelas S1, S6 e S11).

ETR2 e EIN4 possuem somente os domínios (i), (ii) e (iii) e, por fim, os homólogos à ERS1, apresentam os domínios (ii), (iii) e (iv) (Tabela S12).

Associadas à atividade dos receptores ETR, estão as proteínas RAN (*responsivo à agonista*) e RTE (*reversion to ethylene sensitivity 1*), que em *A. thaliana* atuam respectivamente como transportador de cobre, cofator determinante na atividade dos receptores de etileno (Binder et al., 2010), e na regulação das funções do receptor de etileno ETR1 (Resnick et al., 2006). Em soja, foram identificados sete genes homólogos do transportador RAN e quatro homólogos de RTE com a planta modelo (Tabela S12).

Em *A. thaliana*, a proteína CTR1 é considerada principal MKKK (MAP cinase cinase) envolvida na sinalização mediada por etileno (Ju et al., 2012). Em soja, foram identificados cinco genes que codificam proteínas homólogas à CTR1 desta planta modelo, apresentando alto grau de identidade e, conseqüentemente, domínios proteicos conservados (Tabela S12).

Uma das rotas vegetais reguladas por CTR em respostas à etileno seria a representada em *A. thaliana* pela proteína EIN2 (*ethylene insensitive 2*), um regulador positivo das rotas de sinalização mediadas por etileno, atuando à jusante de CTR (Roman et al., 1995; Alonso et al., 1999). Em soja, foram identificados três possíveis homólogos com esta proteína presente nesta planta modelo (GmEIN#002, GmEIN#004 e GmEIN#007) (Tabela S12).

Um ponto importante a se ressaltar é a não identificação de homólogos da proteína *F-box ETP (EIN2 targeting proteins)* em soja. Em *A. thaliana*, esta proteína possui um importante papel na sinalização mediada por etileno ao direcionar a proteína EIN2 para degradação via proteassoma 26S na ausência deste fitormônio (Chen et al., 2011; Ju et al., 2012). A não identificação desta proteína em soja é um indicativo de que outra proteína da mesma família pode desempenhar o mesmo papel que ETP desempenha em *A. thaliana* ou



a existência de outro mecanismo de regulação da atividade de EIN2 ainda desconhecido nesta espécie vegetal.

Em *A. thaliana*, na presença de etileno, a rota de transdução de sinal mediada por etileno induz a atividade dos fatores de transcrição EIN3/EIL (*ethylene insensitive 3; EIN3-like*). Em soja, foram identificados cinco genes que codificam proteínas homólogas à EIN3/EIL (GmEIN#001, GmEIN#005, GmEIN#006, GmEIN#008 e GmEIN#010), inclusive com conservação de domínios específicos (Tabela S12).

A atividade de EIN3/EIL pode ser regulada em *A. thaliana* pelas proteínas *F-box* EBF (*EIN3 binding F-box*) e pelas exoribonuclease EIN5. Na ausência de etileno, as proteínas EBF interagem diretamente com EIN3/EIL e acredita-se que os direcionem para degradação proteolítica via proteassoma 26S. Já a exoribonuclease EIN5, na presença de etileno, seria a responsável pela degradação do mRNA de EBFs, impedindo assim a degradação de EIN3/EIL (Guo & Ecker, 2003; Olmedo et al., 2006; Stepanova & Alonso, 2009). Em soja, foram identificados cinco genes homólogos à EBFs de *A. thaliana* e três homólogos à EIN5 da mesma planta modelo (GmEIN#003, GmEIN#009 e GmEIN#011) (Tabela S12).

Como visto anteriormente, a biossíntese de etileno também é passível de regulação. Foi demonstrado que membros dos dois principais grupos das enzimas ACS de biossíntese de etileno (tipo I e tipo II), representadas em *A. thaliana* por ACS5/ACS9 e ACS6/ACS2, são reguladas por diferentes proteínas cinases. Estudos nesta mesma planta modelo em situações de estresse mostram que a estabilidade das ACSs do tipo I é regulada pelas proteínas cinases MPK3 e MPK6 a partir da cascata de fosforilação MKK4/MKK5/MKK9 → MPK3/MPK6 (Ecker, 2004; Liu & Zhang, 2004). Vale a pena ressaltar que além desta função, alguns estudos com *A. thaliana* associam esta rota de fosforilação com o acúmulo de EIN3/EINL no núcleo (Ecker, 2004; Vandenbussche et al., 2012). Em soja, foram identificados oito homólogos à estas cinases de *A. thaliana*, sendo quatro referentes à MKK4/MKK9 e quatro à MPK3/MPK6 (Tabela S12). Dentre as MKKs, duas são homólogas à MKK4 (GmMKK#001 e GmMKK#003) e duas à MKK9 (GmMKK#002 e GmMKK#004). De forma semelhante, dois homólogos à MPK3 (GmMPK#003 e GmMPK#004) e dois homólogos à MPK6 (GmMPK#001 e GmMPK#002) foram identificados nesta *commodity*.

As enzimas ACS do tipo II possuem um mecanismo similar de regulação, uma vez que, semelhante às ACSs do tipo I são continuamente alvos de degradação na ausência de estímulos, mediada pela pelas proteínas *F-box* ETO (*ethylene overproducer*) e EOL (*ETO-like*). Na ausência de um sinal endógeno, as proteínas ETO/EOL interagem com a exterminada C-terminal das ACSs do tipo II, o que torna estas enzimas alvos de degradação via proteassoma 26S. Existem evidências em *A. thaliana* de que na presença de um sinal, a porção C-terminal desta classe de ACSs seja fosforilada pela *proteína cinase dependente de cálcio 2* (CPK2 ou CDPK2), evento esse que estabiliza estas enzimas inibindo sua degradação (Ecker, 2004; Sebastia et al., 2004; Wang et al., 2004). Em soja, foram identificados dez homólogos à ETO/EOL de *A. thaliana* e cinco à CPK2 da mesma planta modelo (Tabela S12).

#### **5.4. FATORES DE TRANSCRIÇÃO DA SUPERFAMÍLIA AP2/ERF.**

A superfamília de fatores de transcrição AP2/ERF é conhecida por apresentar em sua estrutura o domínio AP2 (*APETALA 2*) de interação com o DNA, e está envolvida na regulação de diversas rotas metabólicas relacionadas com o desenvolvimento vegetal e com respostas aos mais diferentes tipos de estresses (Licausi et al., 2013). No presente trabalho foram identificados um total de 326 proteínas em soja que possuem este domínio em sua estrutura, sendo divididas em três famílias de acordo com a filogenia obtida pela sequência de resíduos de aminoácidos: *AP2*, que apresentam um ou dois domínios AP2; *ERF*, que possuem somente um domínio AP2, filogeneticamente distinto dos apresentados pela família AP2; e *RAV*, que possuem um domínio AP2 e um domínio B3 (também de interação com o DNA) (Tabela 5). Além disso, existe uma sequência adicional pertencente à superfamília AP2/ERF, denominada de *Soloísta*, conservada em todos os genomas sequenciados e que possui uma estrutura que diverge dos outros fatores de transcrição que possuem este domínio (Nakano et al., 2006; Licausi et al., 2013). Foram identificados em soja duas sequências (Glyma08g46250 e Glyma18g33460) que poderiam ser homólogas a este fator de transcrição de *A. thaliana*, mas pelo fato de apresentarem baixa similaridade com seu possível ortólogo presente na planta modelo, e o alinhamento ser restrito a uma pequena região da proteína, estas sequências não foram inseridas neste estudo.

**Tabela 5.** Sumário da Superfamília AP2/ERF de Soja

Classificação	Grupo	N <sup>o1</sup>
Família AP2		<b>46</b>
	Domínio AP2/ERF Duplo	35
	Domínio AP2/ERF Simples	11
Família ERF		<b>275</b>
	Grupos I - IV	99
	Grupos V - X	175
	Grupo VI-L	1
Família RAV		<b>5</b>
	<b>Total</b>	<b>326</b>

<sup>1</sup> Número de genes. Total para cada família está em negrito.

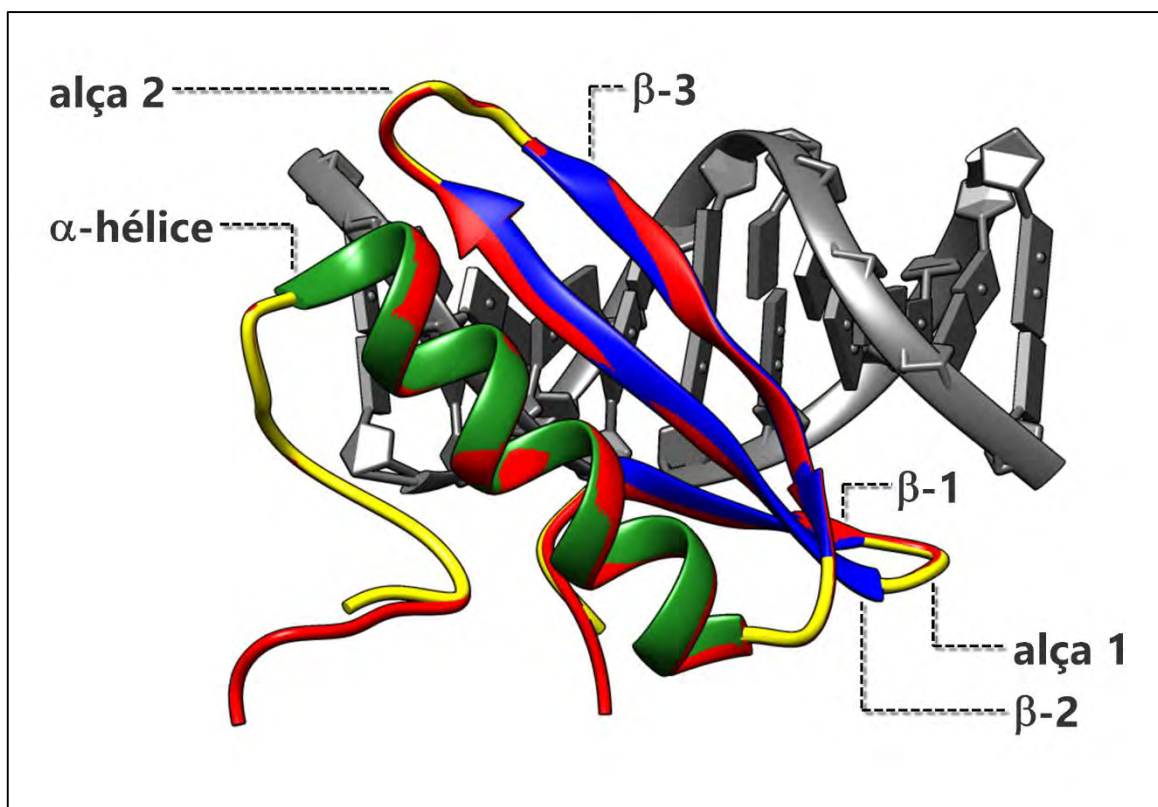
Os membros da família ERF foram agrupados de acordo com a similaridade filogenética de seus domínios AP2 (Figura S2), e posteriormente classificados de acordo com as classificações propostas por Sakuma et al. (2002) e Nakano et al. (2006) (Figura S3). De acordo com a classificação de Sakuma, dos 275 fatores de transcrição da família ERF de soja, 36,00 % correspondem à homólogos do grupo *A* (subfamília CBF/DREB - *A1* até *A6*) e 64,00 % do grupo *B* (subfamília ERF - *B1* até *B6*). Com relação à classificação proposta por Nakano, 5,09 % dos fatores de transcrição da família ERF de soja pertencem ao grupo *I*; 7,27 % ao grupo *II*; 17,45 % ao grupo *III*; 6,18 % ao grupo *IV*; 12,36 % ao grupo *V*; 5,82 % ao grupo *VI*; 0,36 % ao grupo *VI-L*; 3,64 % ao grupo *VII*; 12,00 % ao grupo *VIII*; 23,27 % ao grupo *IX*; e 6,56 % ao grupo *X*. Vale a pena ressaltar que os grupos de *I* à *IV* da classificação de Nakano correspondem ao grupo *A* da classificação de Sakuma, assim como os grupos de *V* à *X*, incluindo *VI-L*, correspondem ao grupo *B* (Figura S3).

Uma vez realizada a caracterização e categorização *in silico* da família ERF de soja, foram identificados motivos proteicos presentes em cada grupo da classificação de Nakano, com a finalidade de identificar possíveis semelhanças entre os membros de cada grupo. Sendo assim, foram selecionados 63 motivos (Tabelas S21 até S83) distribuídos entre os grupos, onde 22,22 % deles estão presentes em fatores de transcrição com domínio AP2/ERF de outras espécies vegetais (Tabela S84), dentre os quais se destaca

principalmente o motivo EAR (*ethylene responsive element binding factor associated with amphiphilic repression*), identificado nos grupos II e VIII, e conhecido por sua participação em mecanismos de repressão transcricional em plantas (Tabelas S29 e S64) (Kagale & Rozwadowski, 2011). Além disso, 23,81 % do total de motivos selecionados possui grande homologia com sequências provenientes de bactérias, leveduras e vírus, enquanto que 17,46 % deles possui homologia com motivos presentes em proteínas da maquinaria eucariótica de replicação, transcrição e tradução (Tabela S84).

A estrutura terciária do domínio AP2 é formada pela interface de uma  $\alpha$ -hélice e três folhas  $\beta$  antiparalelas, sendo que a interação com o DNA ocorre diretamente com as folhas  $\beta$  (Ohme-Takagi & Shinshi, 1995). Com base na estrutura cristalográfica deste domínio depositada em banco de dados, foram construídos modelos moleculares *in silico* dos domínios AP2 da família ERF identificados em soja, apresentados de acordo com a classificação proposta por Nakano e colaboradores (Figura S4). Observou-se que mesmo com a proximidade filogenética entre as sequências de resíduos de aminoácidos deste domínio em fatores classificados em um mesmo grupo, a estrutura terciária dos mesmos apresentava diferenças, localizadas principalmente nas alças e nas folhas  $\beta$ , o que pode acarretar em pequenas variações na interação com a sequência específica de DNA por eles reconhecida.

Em paralelo, foram selecionados em soja possíveis candidatos à ERF1, baseando-se na identidade com a sequência primária de seu homólogo em *A. thaliana* (At3g23240) e na possibilidade de serem regulados por EIN3/EIL, isto é, apresentarem elementos *cis*-atuantes responsivos à este fator de transcrição em seus possíveis promotores. Desta forma, foram selecionados oito possíveis candidatos (GmERF#127, GmERF#129, GmERF#166, GmERF#169, GmERF#171, GmERF#255, GmERF#256 e GmERF#259), que tiveram a estrutura de seus domínios AP2 modelada *in silico*, possibilitando assim avaliar e confirmar sua conservação (Figura 9). Vale a pena ressaltar que pelo experimento de BBH, somente GmERF#134 foi considerado como um possível órtologo do referido ERF1 de *A. thaliana* (Tabela S19).



**Figura 9. Estrutura Terciária do Domínio AP2 de Possíveis Candidatos à Fatores Responsíveis à Etileno 1 (ERF1) de Soja.** A figura ilustra as estruturas terciárias dos domínios AP2 de oito possíveis candidatos à ERF1 de soja, obtidas por modelagem molecular e evidenciando suas principais estruturas secundárias: uma  $\alpha$ -hélice, três folhas  $\beta$  ( $\beta$ -1,  $\beta$ -2 e  $\beta$ -3) e duas alças (alça 1 e 2). As sequências de resíduos de aminoácidos destes candidatos de soja possuem homologia com a proteína ERF1 de *Arabidopsis thaliana* (At3g23240) e possuem elementos *cis*-atuantes responsivos à EIN3/EINL em seus possíveis promotores: GmERF#127, GmERF#129, GmERF#166, GmERF#169, GmERF#171, GmERF#255, GmERF#256 e GmERF#259. Cada candidato à ERF1 apresenta a estrutura de seu domínio AP2 sobreposta à estrutura cristalográfica proposta para este domínio presente na proteína ERF1 de *A. thaliana*, obtida em banco de dados de proteínas (PDB). As estruturas secundárias em **verde**, **azul** e **amarelo** correspondem respectivamente à  $\alpha$ -hélices, folhas  $\beta$  e alças dos domínios AP2 de soja. Já as estruturas secundárias em **vermelho** correspondem ao modelo de *A. thaliana*. A estrutura em **cinza** representa a sequência de DNA reconhecida pelo domínio.

### 5.5. ANÁLISE E VALIDAÇÃO DE TRANSCRIPTOMAS DE SOJA.

Após a identificação e caracterização *in silico* dos genes e proteínas por eles codificadas relacionadas com a biossíntese e sinalização mediada por etileno em soja, foram realizadas análises da expressão diferencial dos mesmos em transcriptomas de duas cultivares da *commodity* cultivadas em sistema hidropônico, sob diferentes tempos de estresse hídrico, sendo uma delas suscetível e outra tolerante a seca (BR16 e EMBRAPA48, respectivamente). Estes transcriptomas, disponibilizados pelo Projeto

GENOSOJA, foram construídos utilizando a metodologia denominada de biblioteca subtrativa, que detecta os transcritos diferencialmente expressos em situações de interesse, que no caso é a de déficit hídrico (seca). Desta forma, 23,90 % do total de genes identificados apresentaram expressão diferencial em pelo menos uma das situações analisadas, dentre os quais 25,83 % estão relacionados com a biossíntese de etileno, enquanto que 35,00 % estão relacionados com a transdução mediada por este fitormônio e 39,17 % com membros da superfamília de fatores de transcrição AP2/ERF (Figuras S5 – S7). Além disso, 19,17 % dos genes diferencialmente expressos foram identificados exclusivamente no banco de dados da cultivar de soja suscetível à seca (BR16), 54,17 % são exclusivos da cultivar tolerante (EMBRAPA48), e 26,66 % estão presentes no banco de dados de ambas as cultivares. Tal diferença pode ser explicada tanto pelo fato da variação genética entre as duas cultivares propiciar essa variação relativa no número de genes diferencialmente expressos entre ambas, ou também pelo fato do número de sequências únicas obtidas nos transcriptomas em questão ser relativamente discrepante quando comparam-se os bancos de dados das duas cultivares (42,3 milhões de sequências únicas geradas, sendo que 27,8% são referentes à BR16 e 72,2% à EMBRAPA48) (Rodrigues et al., 2012).

Quanto à expressão tecido-específica, observa-se que 36,67 % dos genes diferencialmente expressos identificados foram detectados exclusivamente em raiz (dentro os quais 5,88 % foram identificados exclusivamente no banco de dados da cultivar BR16, 90,59 % são exclusivos da cultivar EMBRAPA48 e 3,53 % estão presentes no banco de dados de ambas as cultivares), 29,17 % foram detectados exclusivamente em folha (dentro os quais 47,37 % foram identificados exclusivamente no banco de dados da cultivar BR16, 32,89 % são exclusivos da cultivar EMBRAPA48 e 19,74 % estão presentes no banco de dados de ambas as cultivares) e 34,16 % dos genes diferencialmente expressos tanto em raiz quanto em folha. Com base nestes dados, associados com análise de valores de expressão normalizados e plotados em gráficos de dispersão (Figura S8), infere-se que a maior parte dos genes diferencialmente expressos em raiz é observada na cultivar tolerante à seca, observação essa que é atenuada quando o tecido analisado é o foliar, onde a expressão diferencial entre as cultivares é mais balanceada, ora tendendo para cultivar suscetível, ora para a tolerante, dependendo também do grupo de genes e do tempo de estresse analisado.

Com relação à expressão diferencial nos grupos de genes analisados, observou-se que 28,70 % dos genes de biossíntese de etileno foram detectados em transcriptomas de soja sob déficit hídrico, tanto em raiz quanto em folha de cultivares suscetível e tolerante à seca (Figura S5). Dentre eles, 16,13 % tiveram sua expressão detectada somente na cultivar suscetível à seca (BR16), majoritariamente em folhas. Por outro lado, cerca de 48,39 % dos genes diferencialmente expressos neste grupo foram detectados exclusivamente na cultivar tolerante (EMBRAPA48), principalmente no tecido radicular (Figura S5).

Já 61,76 % dos representantes do grupo relacionado com a transdução de sinal mediada por este fitormônio apresentaram expressão diferencial em transcriptomas de soja sob déficit hídrico, em ambos os tecidos de ambas as cultivares analisadas (Figura S6). Dentre eles, 33,33 % tiveram sua expressão detectada somente na cultivar suscetível à seca (BR16), com a expressão diferencial majoritariamente em folhas, semelhante à mesma análise realizada com genes de biossíntese. Por outro lado, 42,86 % dos genes diferencialmente expressos neste grupo foram detectados exclusivamente na cultivar tolerante (EMBRAPA48), principalmente em raízes (Figura S6).

Com a relação aos fatores de transcrição da superfamília AP2/ERF em transcriptomas de soja em condições de déficit hídrico, observou-se que cerca de 14,40 % dos genes relativos a este grupo apresentaram-se como diferencialmente expressos, pertencentes exclusivamente às famílias ERF e RAV. Desse total, 36,18 % tiveram sua expressão diferencial detectada exclusivamente em raiz, 31,91 % exclusivamente em folhas e 31,91 % nos dois tecidos simultaneamente. Além disso, 8,51 % destes genes foram diferencialmente expressos na cultivar suscetível à seca (BR16) enquanto que 68,09 % foram diferencialmente expressos exclusivamente na cultivar tolerante à déficit hídrico (EMBRAPA48) e 23,40 % em ambas as cultivares. Observa-se a predominância numérica de genes diferencialmente expressos tanto no tecido radicular quanto no foliar da cultivar EMBRAPA48 (Figura S7).

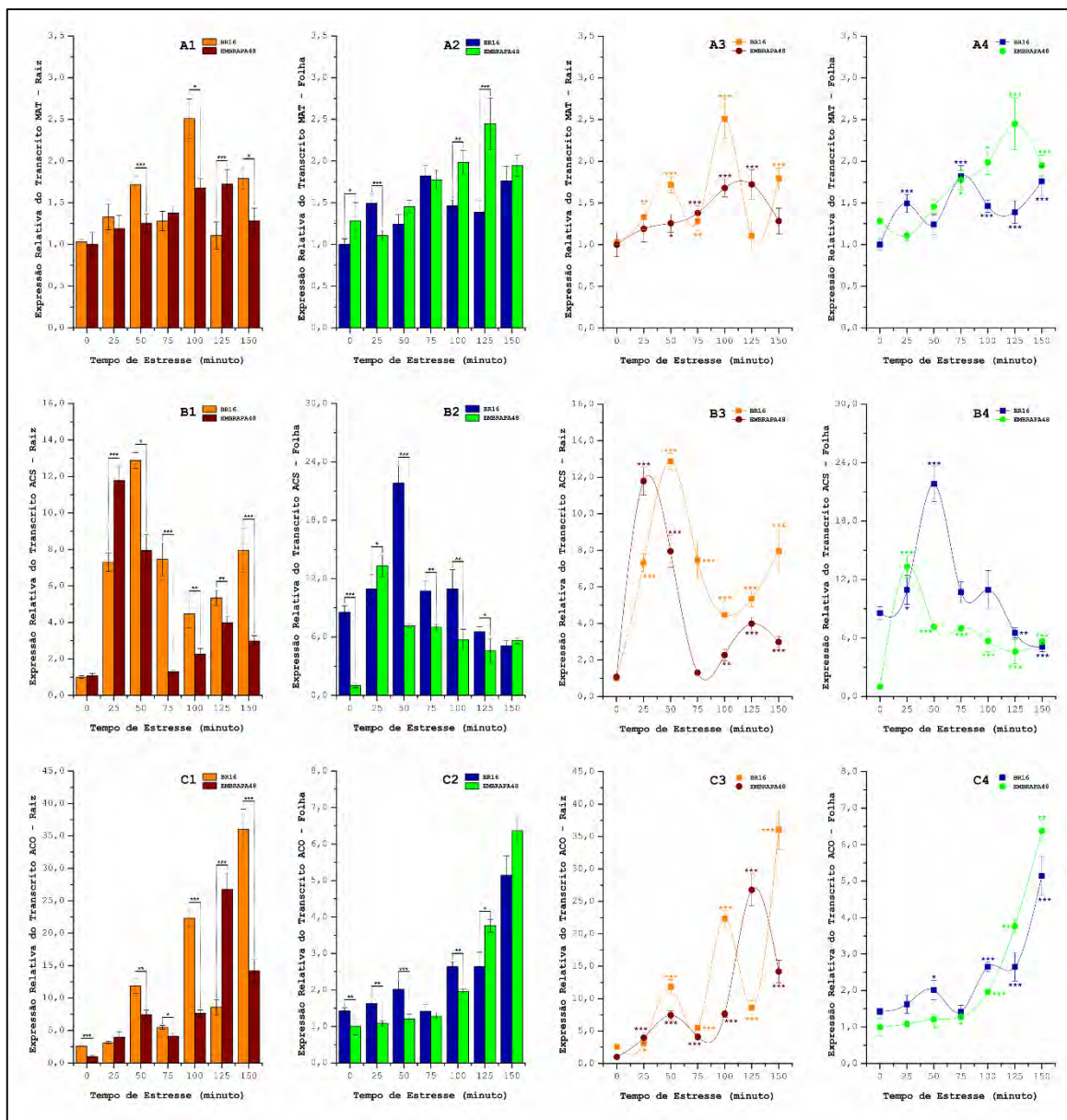
Ainda com relação aos membros da superfamília AP2/ERF diferencialmente expressos, observou-se que 17,02 % deles pertence ao grupo *I* (GmERF#051, GmERF#068, GmERF#079, GmERF#105, GmERF#161, GmERF#182, GmERF#198 e GmERF#241); 8,51 % ao grupo *II* (GmERF#012, GmERF#047, GmERF#075 e GmERF#145); 12,77 % ao grupo *III* (GmERF#010, GmERF#121, GmERF#141, GmERF#173, GmERF#223 e GmERF#267); 10,64 % ao grupo *VII* (GmERF#015,

GmERF#044, GmERF#125, GmERF#210 e GmERF#214); 8,51 % ao grupo *VIII* (GmERF#085, GmERF#160, GmERF#179 e GmERF#205); 23,40 % ao grupo *IX* (GmERF#038, GmERF#095, GmERF#122, GmERF#128, GmERF#129, GmERF#133, GmERF#171, GmERF#232, GmERF#264, GmERF#269 e GmERF#272); 17,02 % ao grupo *X* (GmERF#020, GmERF#021, GmERF#57, GmERF#069, GmERF#099, GmERF#106, GmERF#219 e GmERF#222); e 2,13 % à família RAV (GmRAV#001) (Figura S7). Dentre estes genes, alguns destacam-se quanto à função predita *in silico*, sendo os principais: **ABR1** - repressor *ABA 1* (GmERF#069), **CBF2** - repetição *C/fator de interação com DRE 2* (GmERF#173), **DREB1D** - proteína com elemento de interação responsivo à desidratação *1D* (GmERF#223), **ERF1** (GmERF#129 e GmERF#171), **ERF1A** (GmERF#232), **RAP2.12** (GmERF#044 e GmERF#214), **RAP2.3** (GmERF#015, GmERF#125 e GmERF#210), **RAP2.4/WIND1** - *desdiferenciação induzida por ferimento 1* (GmERF#051, GmERF#079, GmERF#161 e GmERF#198), **RAP2.5** (GmERF#160), **RAP2.6** (GmERF#057), **ORA47** (GmERF#047 e GmERF#075), **RRTF** - *fator de transcrição de resposta redox* (GmERF#099), **TINY2** - *proteína com elemento de interação responsivo à desidratação 2* (GmERF#121) e **TDR1** - *regulador transcricional de resposta à defesa 1* (GmERF#128) (Tabela S14; Figura S7).

Com a tentativa de validar os dados obtidos *in silico* com dados moleculares, foram mensurados por PCR em Tempo Real (qPCR) os níveis de alguns transcritos diferencialmente expressos dos três grupos de genes soja (biossíntese de etileno, transdução de sinal por ele mediada e fatores de transcrição da superfamília AP2/ERF), tanto em raiz quanto em folha das cultivares BR16 e EMBRAPA48. As plantas foram cultivadas nas mesmas condições em que foram obtidos os transcriptomas analisados e os valores médios do ciclo limiar ( $C_t$ ) obtidos encontram-se nas Tabelas S85 e S86.

Quando se confrontam plantas não estressadas com estressadas, as análises de qPCR para os genes MAT, ACS e ACO (Figura 10-A, 10-B e 10-C, respectivamente) comprovaram o caráter diferencial que os mesmos apresentaram em seus dados originados em transcriptomas, mas devido à limitações da técnica na construção de bibliotecas subtrativas, associadas com variações experimentais, observou-se modificações nos perfis de expressão dos transcritos em questão. Dentre estas modificações, destaca-se principalmente a expressão diferencial do gene ACS, que segundo os dados de qPCR foi detectado como diferencialmente expresso nos dois tecidos de ambas as cultivares, mas



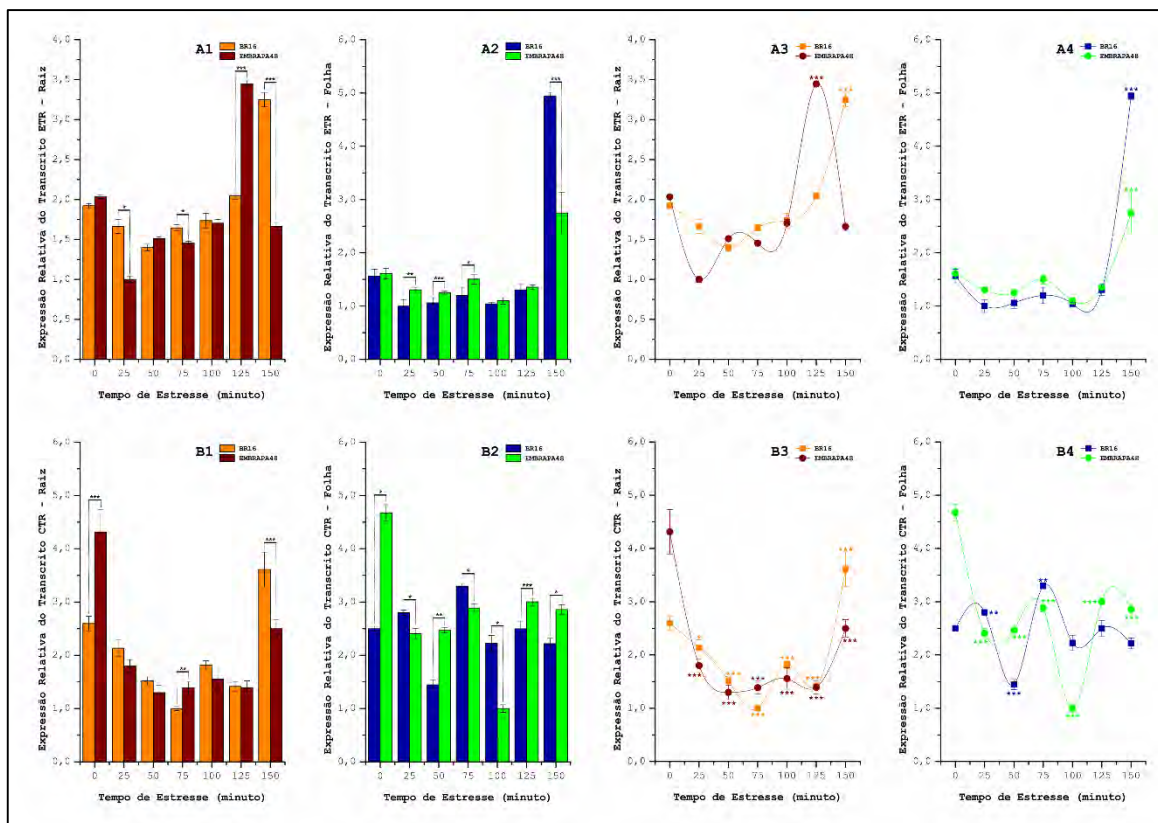


**Figura 10. Níveis de Expressão dos Genes de Biossíntese de Etileno em Cultivares de Soja Sob Déficit Hídrico.** Os gráficos ilustram os níveis de expressão, obtidos por qPCR, de três genes de soja que estão relacionados com a biossíntese de etileno (MAT, ACS e ACO). O experimento comparou os níveis de expressão destes genes em raiz e folha das cultivares de soja BR16 e EMBRAPA48 em diferentes tempos de déficit hídrico. Os códigos A1 - A4, B1 - B4 e C1 - C4 nomeiam os gráficos referentes à expressão dos genes MAT, ACS e ACO respectivamente. Nos gráficos de barra A1/A2, B1/B2 e C1/C2, a análise estatística foi obtida comparando-se os dados entre tecidos semelhantes das duas cultivares em um mesmo tempo de déficit hídrico. Já nas curvas de tendência A3/A4, B3/B4 e C3/C4, a estatística aplicada foi por comparação dos dados da planta não estressada (tempo zero) com a estressada (demais tempos de déficit hídrico). Sendo assim, a diferença entre os dados representados nos gráficos 1 e 2 com o 3 e 4 (tanto em A quanto em B e C) é somente a análise estatística. Os asteriscos representam o nível de significância estatística: (\*)  $p \leq 0,05$ ; (\*\*)  $0,01 \leq p < 0,05$ ; (\*\*\*)  $0,001 \leq p < 0,01$ . Cada ponto representa o valor médio ( $\pm$  desvio padrão) de três repetições técnicas (em uma mesma amostra) em três amostras biológicas (plantas diferentes), totalizando assim nove replicatas. A ausência da representação gráfica do desvio padrão em alguns pontos ocorre pelo fato de seus valores absolutos serem inferiores à escala. Após serem normalizados com base em genes endógenos, os valores apresentados nos gráficos são relativos ao de menor expressão, cujo valor é determinado como 1,0 (um). Informações sobre os genes alvo e endógenos são apresentadas na Tabela 3.

havia sido inicialmente detectado apenas em transcriptomas de raízes da cultivar EMBRAPA48.

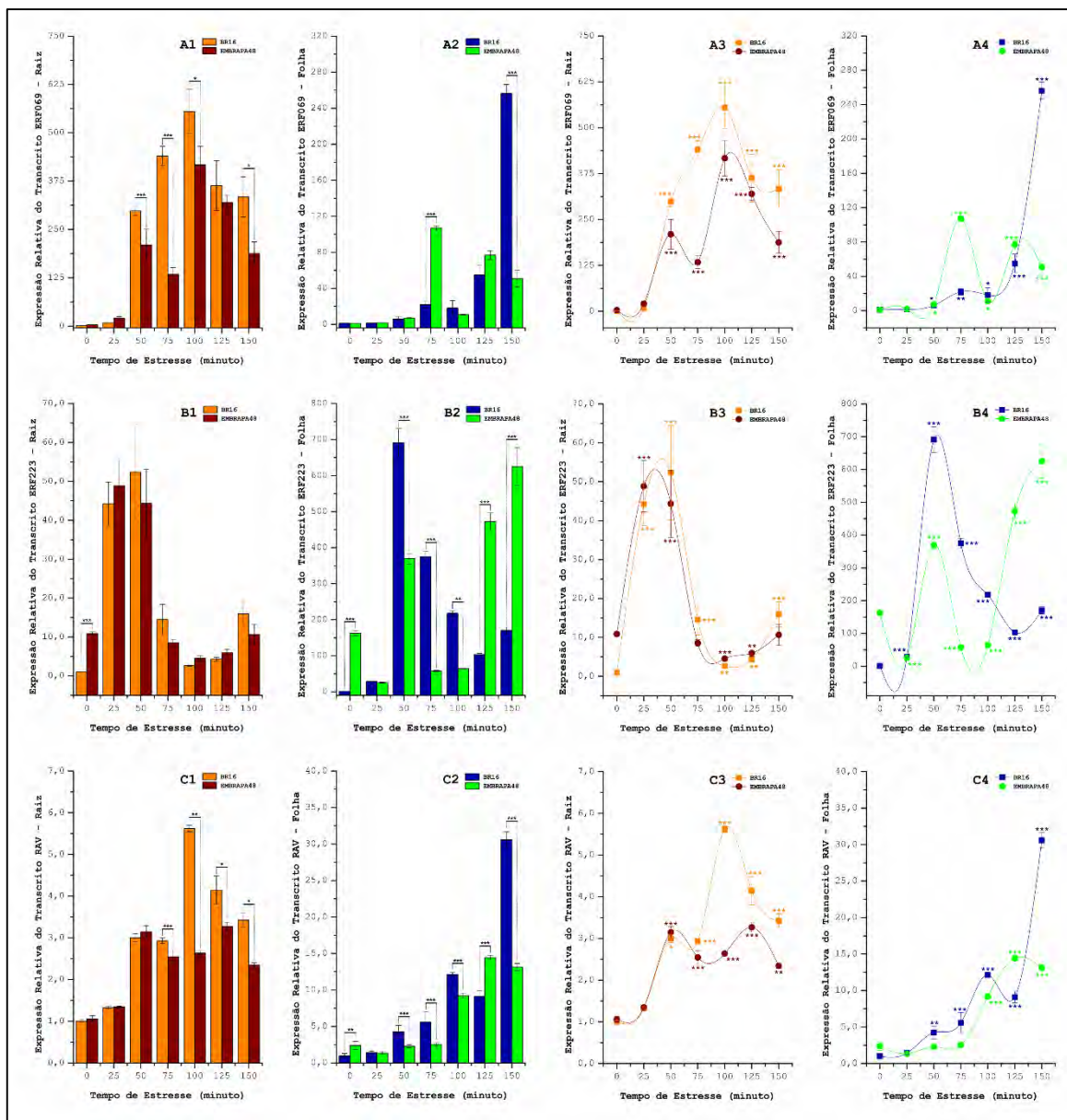
Análises similares foram realizadas com genes relativos aos receptores de etileno (ETR) e à proteína cinase CTR (Figura 11-A e 11-B, respectivamente), ambos pertencentes ao grupo relativo à transdução de sinal mediada por etileno em soja. Poucas diferenças foram observadas no padrão de expressão dos transcritos destes genes entre as cultivares. Nas raízes de ambas as cultivares, observa-se uma redução no nível de transcritos de ETR quando se comparam plantas estressadas e não estressadas (tempo zero), onde os picos de expressão são alcançados somente nos tempos finais de estresse (125 minutos na cultivar EMBRAPA48 e 150 minutos na BR16). Ainda no mesmo tecido, os transcritos do gene CTR apresentam redução gradativa em seus níveis de expressão, tendo um aumento significativo somente com 150 minutos de déficit hídrico em ambas as cultivares. Em folhas, quando se comparam as situações de estresse e de não estresse observa-se uma leve redução dos níveis dos transcritos de ETR nos primeiros 125 minutos de seca e um pico de elevação dos mesmos no final da análise (150 minutos). Já com relação ao transcrito CTR observa-se que a expressão do mesmo na cultivar tolerante à seca é maior na situação não estressada (tempo zero).

Também foram selecionados três candidatos dentre os fatores de transcrição da superfamília AP2/ERF diferencialmente expressos para validação por qPCR, nomeados como ERF069, ERF223 e RAV (Figura 12-A, 12-B e 12-C, respectivamente), detectados respectivamente no banco de dados de raiz da cultivar EMBRAPA 48 (125 - 150 minutos de estresse), de folhas da cultivar EMBRAPA48 (25 - 50 minutos de seca) e de raiz da cultivar BR16 (75 - 100 minutos de déficit hídrico). Quando se comparam as análises de qPCR de plantas estressadas com não estressadas (tempo zero), os dados presentes no banco de dados do Projeto GENOSOJA são confirmados, mas a expressão diferencial é observada nas duas cultivares, que expressam os três genes nos dois tecidos analisados (raiz e folha), mesmo que em níveis diferentes. Ainda segundo à análise de qPCR, a indução de ERF223 e RAV nestas condições seria mais evidente em folhas, principalmente na cultivar BR16, e de ERF069 nas raízes da mesma cultivar. Tanto ERF069 em raiz quanto ERF223 em folha podem ser induzidos mais de 600 vezes, quando comparados com as plantas não estressadas nas mesmas condições experimentais (Figura 12-A e 12-B). Vale a pena ressaltar que os níveis de transcritos do gene ERF223 na cultivar



**Figura 11. Níveis de Expressão dos Genes Relacionados à Transdução de Sinal Mediada por Etileno em Cultivares de Soja Sob Déficit Hídrico.** Os gráficos ilustram os níveis de expressão, obtidos por qPCR, de dois genes de soja que estão relacionados com a biossíntese de etileno (ETR e CTR). O experimento comparou os níveis de expressão destes genes em raiz e folha das cultivares de soja BR16 e EMBRAPA48 em diferentes tempos de déficit hídrico. Os códigos **A1 - A4** e **B1 - B4** nomeiam os gráficos referentes à expressão dos genes ETR e CTR respectivamente. Nos gráficos de barra **A1/A2** e **B1/B2**, a análise estatística foi obtida comparando-se os dados entre tecidos semelhantes das duas cultivares em um mesmo tempo de déficit hídrico. Já nas curvas de tendência **A3/A4** e **B3/B4**, a estatística aplicada foi por comparação dos dados da planta não estressada (tempo zero) com a estressada (demais tempos de déficit hídrico). Sendo assim, a diferença entre os dados representados nos gráficos **1** e **2** com o **3** e **4** (tanto em **A** quanto em **B**) é somente a análise estatística. Os asteriscos representam o nível de significância estatística: (\*)  $p \leq 0,05$ ; (\*\*)  $0,01 \leq p < 0,05$ ; (\*\*\*)  $0,001 \leq p < 0,01$ . Cada ponto representa o valor médio ( $\pm$  desvio padrão) de três repetições técnicas (em uma mesma amostra) em três amostras biológicas (plantas diferentes), totalizando assim nove replicatas. A ausência da representação gráfica do desvio padrão em alguns pontos ocorre pelo fato de seus valores absolutos serem inferiores à escala. Após serem normalizados com base em genes endógenos, os valores apresentados nos gráficos são relativos ao de menor expressão, cujo valor é determinado como 1,0 (um). Informações sobre os genes alvo e endógenos são apresentadas na Tabela 3.





**Figura 12. Níveis de Expressão dos Genes da Superfamília de Fatores de Transcrição AP2/ERF em Cultivares de Soja Sob Déficit Hídrico.** Os gráficos ilustram os níveis de expressão, obtidos por qPCR, de três genes de soja que estão relacionados com superfamília de fatores de transcrição AP2/ERF (ERF069, ERF223 e RAV). O experimento comparou os níveis de expressão destes genes em raiz e folha das cultivares de soja BR16 e EMBRAPA48 em diferentes tempos de déficit hídrico. Os códigos A1 - A4, B1 - B4 e C1 - C4 nomeiam os gráficos referentes à expressão dos genes ERF069, ERF223 e RAV respectivamente. Nos gráficos de barra A1/A2, B1/B2 e C1/C2, a análise estatística foi obtida comparando-se os dados entre tecidos semelhantes das duas cultivares em um mesmo tempo de déficit hídrico. Já nas curvas de tendência A3/A4, B3/B4 e C3/C4, a estatística aplicada foi por comparação dos dados da planta não estressada (tempo zero) com a estressada (demais tempos de déficit hídrico). Sendo assim, a diferença entre os dados representados nos gráficos 1 e 2 com o 3 e 4 (tanto em A quanto em B e C) é somente a análise estatística. Os asteriscos representam o nível de significância estatística: (\*)  $p \leq 0,05$ ; (\*\*)  $0,01 \leq p < 0,05$ ; (\*\*\*)  $0,001 \leq p < 0,01$ . Cada ponto representa o valor médio ( $\pm$  desvio padrão) de três repetições técnicas (em uma mesma amostra) em três amostras biológicas (plantas diferentes), totalizando assim nove replicatas. A ausência de representação gráfica do desvio padrão em alguns pontos ocorre pelo fato de seus valores absolutos serem inferiores à escala. Após serem normalizados com base em genes endógenos, os valores apresentados nos gráficos são relativos ao de menor expressão, cujo valor é determinado como 1,0 (um). Informações sobre os genes alvo e endógenos são apresentadas na Tabela 3.

EMBRAPA48 não estressada (tempo zero) são mais elevados que os da cultivar BR16, tanto em raízes quanto em folhas (Figura 12-B). Sabendo-se que ERF223 é homólogo à uma proteína de resposta à seca de *A. thaliana* (DREB1D), pode-se inferir que esta seja uma possível adaptação positiva para sua maior tolerância à seca, uma vez que este fato pode acarreta em respostas mais rápidas à estresses.

## 5.6. ANÁLISE DE POSSÍVEIS PROMOTORES GÊNICOS DE SOJA.

Com a finalidade de melhor compreender os mecanismos de regulação transcricional dos 502 genes de soja analisados, foi realizada uma varredura *in silico* nas possíveis regiões promotoras dos mesmos. Desta forma, foram identificadas 44.890 correspondências referentes à 29 famílias de elementos *cis*-atuantes descritos na literatura, importantes na regulação transcricional de diversas espécies vegetais. Desse total, 20,82 % destas correspondências estão presentes em possíveis promotores de genes de soja relacionados à biossíntese de etileno, 11,23 % em promotores de genes relacionados com a transdução de sinal mediada por este fitormônio e 67,95 % em promotores de membros da superfamília de fatores de transcrição AP2/ERF (Tabela S87)<sup>2</sup>.

Dentre as possíveis regiões promotoras de soja analisadas todas apresentam elementos da família *PTBP* (*TATA-box* de plantas) e/ou *CAAT* (elementos *CCAAT-box*), o que sugere a real possibilidade de que as sequências analisadas possam vir atuar como promotores gênicos *in vivo*, uma vez que ambas as famílias podem participar de promotores gênicos basais (Figura 13).

Além das famílias de elementos *cis*-atuantes *PTBP* e *CAAT*, as que apresentaram elementos com maior representatividade nos possíveis promotores de soja foram as relacionadas com fatores de transcrição MYB, MYC e NAC (Tabela S87; Figura 13), além das que respondem à fatores abióticos como luz e à calor (famílias *LREM* e *HEAT*, respectivamente). Vale a pena ressaltar que no presente trabalho não foram identificados possíveis promotores de soja com elementos *cis*-atuantes referentes à família *RAV3*, o que pode ser explicado por variações na sequência reconhecida pelo domínio B3 presente nos representantes da família *RAV* em soja, ou que as possíveis regiões promotoras analisadas

---

<sup>2</sup> Devido ao tamanho, a Tabela S87 encontra-se somente no material em anexo no CD.

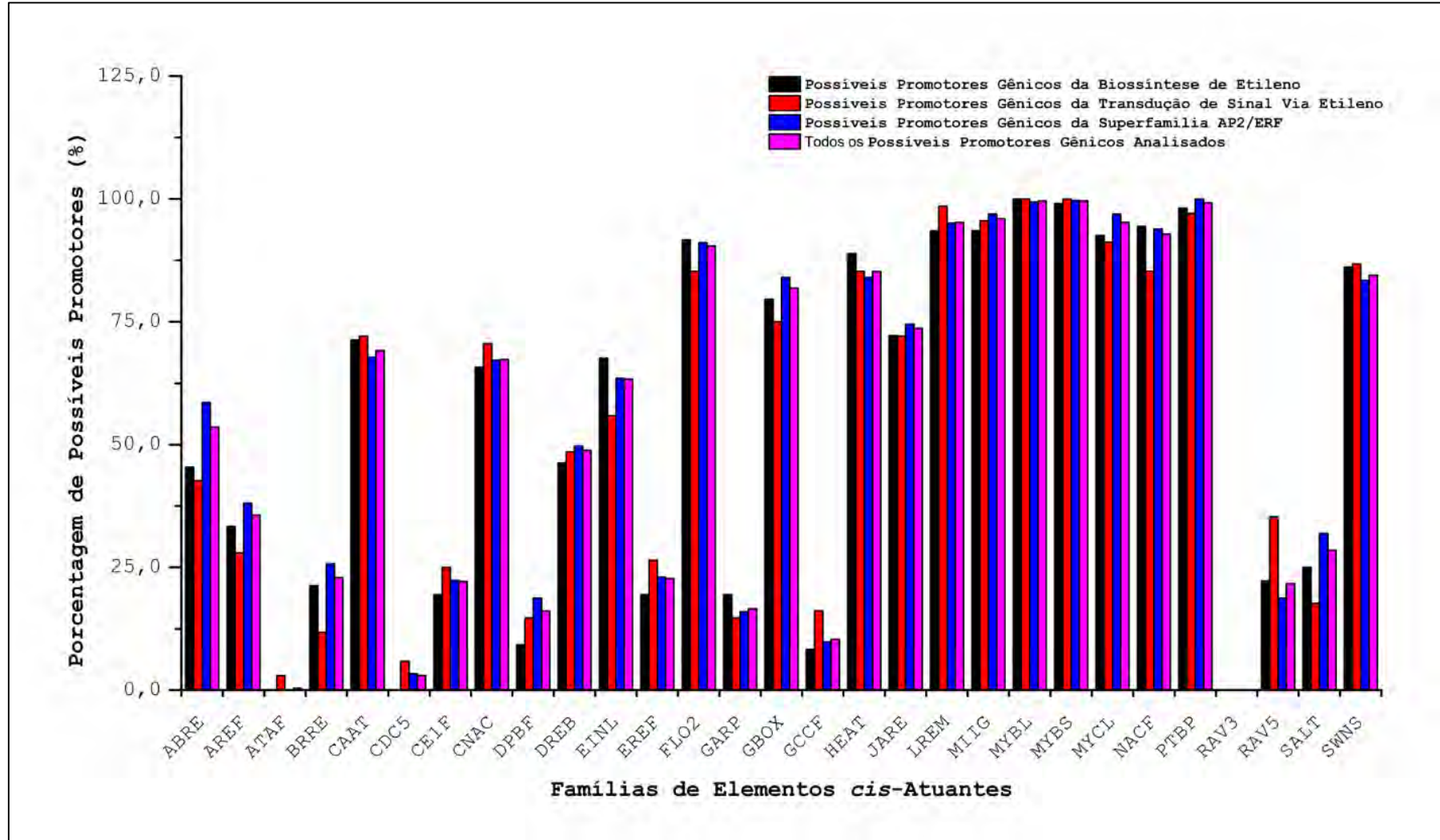


Figura 13. Distribuição de Elementos *cis*-Atuantes em Possíveis Regiões Promotoras de Genes de Soja.

**Figura 13 (cont.). Distribuição de Elementos *cis*-Atuantes em Possíveis Regiões Promotoras de Genes de Soja.** O gráfico apresenta a distribuição de elementos *cis*-atuantes em regiões promotoras dos genes de soja relacionados com a biossíntese de etileno, com a transdução de sinal mediada por este fitormônio gasoso e com fatores de transcrição da superfamília AP2/ERF. As famílias de elementos *cis*-atuantes identificadas foram: *ABRE* (elementos de resposta à ABA); *AREF* (elementos de resposta à auxina); *ATAF* (elementos de interação com proteínas contendo domínio NAC semelhante à *ATAF*); *BRRE* (elementos de resposta à brassinoesteróides); *CAAT* (elementos *CCAAT-box*); *CDC5* (elementos homólogos ao *CDC5* de *Arabidopsis thaliana*); *CEIF* (elementos de acoplamento 1); *CNAC* (elementos de interação com fatores NAC regulados por cálcio); *DPBF* (elementos de interação em promotores *Dc3*); *DREB* (proteínas com elementos de interação responsivos à desidratação); *EINL* (elementos de interação com fatores semelhantes à *EIN3*); *EREF* (elementos responsivos à etileno); *FLO2* (elementos de interação com proteína homeótica floral *APETALA 2*); *GARP* (elementos de interação com DNA relacionados à MYB - Golden2, ARR, Psr); *GBOX* (elementos *G-box/C-box* de interação com bZIP de plantas); *GCCF* (elementos *GCC-box*); *HEAT* (elementos de choque térmico); *JARE* (elementos de resposta à jasmonato); *LREM* (elementos responsivos à luz, não modulados por diferentes tipos de luz); *MIIG* (sítios de interação com MYB IIG); *MYBL* (elementos de interação com proteínas semelhantes à MYB); *MYBS* (MYB com elementos únicos de interação com DNA); *MYCL* (elementos de interação com proteínas semelhantes à MYC); *NACF* (elementos de interação com fatores NAC específicos de plantas); *PTBP* (*TATA-box* de plantas); *RAV3* (porção 3' do sítio de interação com *RAV1*); *RAV5* (porção 5' do sítio de interação com *RAV1*); *SALT* (elementos responsivos à seca e variação de salinidade); *SWNS* (elementos de resposta à NACS - parede secundária).

possam ser reguladas preferencialmente pela interação do domínio AP2 presentes nestes fatores com elementos *cis*-atuantes da família *RAV5*, amplamente identificados nas possíveis regiões promotoras analisadas (Wittkopp & Kalay, 2012).

Dentre as famílias de elementos *cis*-atuantes identificadas, as que podem estar relacionadas diretamente com as rotas metabólicas sinalizadas por etileno seriam as famílias *EINL* (elementos de interação com fatores semelhantes à *EIN3*) e a família *GCCF* (elementos *GCC-box*), presentes respectivamente em 63,35 % e 10,36 % do total de possíveis promotores de soja analisados no presente trabalho. Podem-se destacar também dentre os elementos envolvidos com o desenvolvimento e respostas à estresses os pertencentes as família *DREB* (proteínas com elementos de interação responsivos à desidratação) e *EREF* (elementos responsivos à etileno), presentes respectivamente em 48,80 % e 22,71 % dos possíveis promotores de soja analisados (Figura 13).

Levando-se em consideração somente os possíveis promotores de soja relacionados com a biossíntese de etileno, observa-se que 67,59 % deles possuem elementos da família *EINL* em sua estrutura enquanto que 8,33 % deles possuem elementos da família *GCCF*, sendo que ambas as famílias de elementos *cis*-atuantes estão associadas com a regulação mediada diretamente por etileno (Figura 13). Além disso, observa-se também a presença de famílias de elementos que respondem à outros fitormônios, sendo que dentre elas destacam-se a que responde diretamente ao ácido jasmônico (família *JARE*), que está

presente em mais de 70,00 % dos possíveis promotores de soja deste grupo, seguida pelas famílias *ABRE* e *CEIF*, que podem responder diretamente à ABA, presentes em respectivamente 45,37 % e 19,44 % do mesmo grupo de possíveis promotores. Além disso, 30,00 % deles possuem elementos de resposta à auxinas e 21,33 % à brassinoesteróides. Por fim, observa-se a presença de elementos *cis*-atuantes da família *DREB* e *EREF* em respectivamente 46,30 % e 19,44 % dos possíveis promotores de soja do referido grupo.

Considerando-se agora o grupo relativo à transdução de sinal mediada por etileno, observa-se a presença de elementos da família *EINL* em 55,88 % dos possíveis promotores de soja analisados, enquanto que 16,86 % deles possuem elementos da família *GCCF*. Além disso, observa-se também a presença de famílias de elementos que respondem à outros fitormônios, sendo que dentre elas destacam-se a responsiva ao ácido jasmônico (família *JARE*), que semelhante à análise feita com os possíveis promotores de biossíntese, é a com o maior número de correspondências identificadas, estando presente em cerca de 70,00 % dos possíveis promotores deste grupo, seguida pelas famílias *ABRE* e *CEIF*, que podem responder diretamente ao ácido abscísico (ABA), presentes em respectivamente 42,65 % e 25,00 % do mesmo grupo de possíveis promotores de soja. Além disso, cerca de 28,00 % deles possuem elementos de resposta à auxinas e somente 11,76 % à brassinoesteróides. Por fim, observa-se a presença de elementos *cis*-atuantes da família *DREB* e *EREF* em respectivamente 48,53 % e 26,47 % dos possíveis promotores de soja do referido grupo (Figura 13).

Por fim, o grupo de possíveis promotores que regulam os genes que codificam fatores de transcrição da superfamília AP2/ERF apresenta elementos das famílias *EINL* e *GCCF* em respectivamente 63,50 % e 9,82 % de seus representantes, enquanto que cerca de 49,60 % deles apresenta elementos *cis*-atuantes da família *DREB* e 23,01 % da família *EREF*. Elementos que estão associados com a regulação transcricional mediada por ácido jasmônico (74,54 %), seguidos dos responsivos à ABA (família *ABRE* - 58,59 %), auxinas (38,04 %), brassinoesteróides (25,77 %) e por elementos da família *CEIF* (também responsiva à ABA - 22,39 %) também foram identificados em possíveis promotores de genes referentes à estes fatores de transcrição em soja (Figura 13). Além disso, observou-se que dentre os possíveis promotores dos genes deste grupo, detectados como diferencialmente expressos nos transcriptomas de soja analisados, cerca de 53,00 % apresentaram elementos da família *EINL* e 17,02 % da família *GCCF*, enquanto que cerca

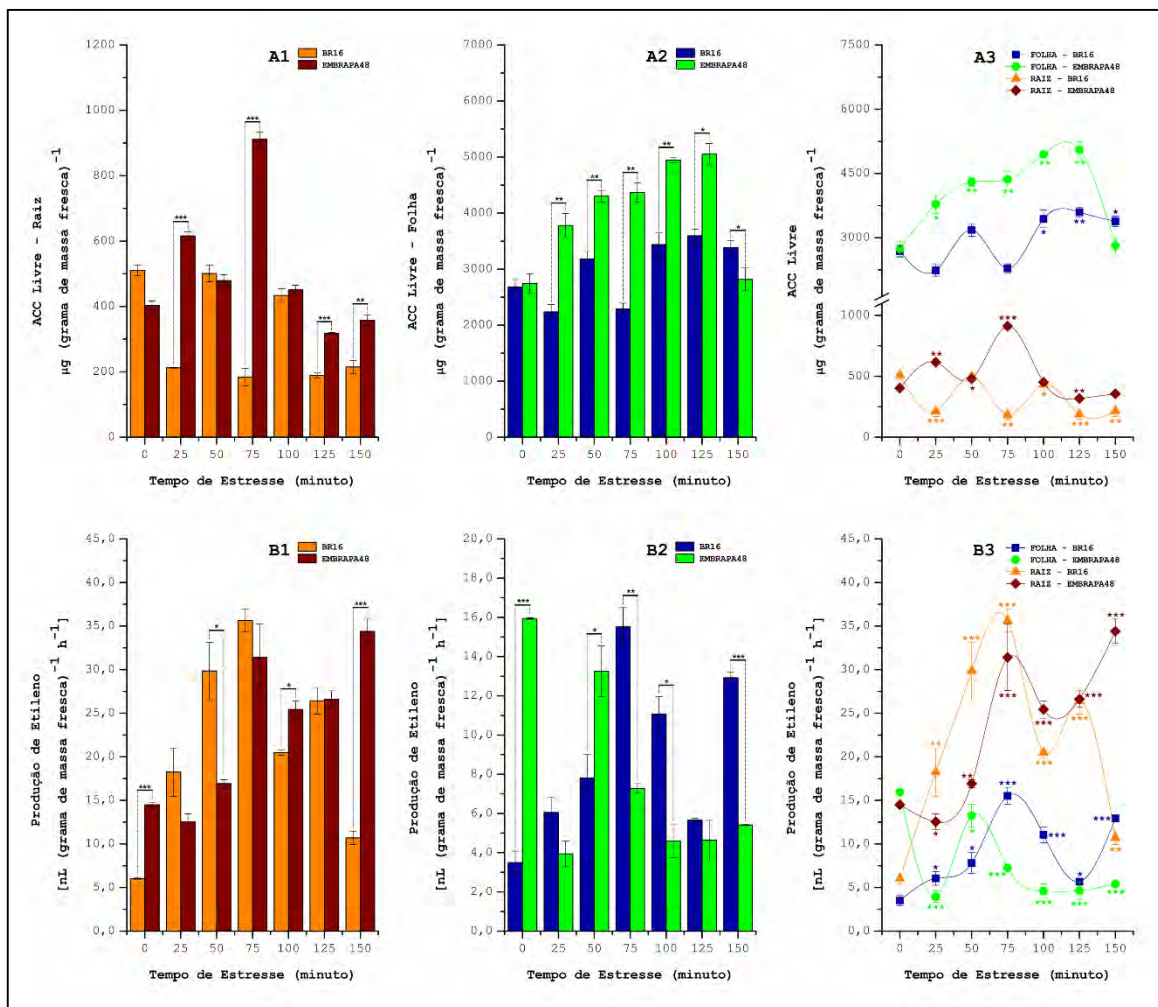


de 61,70 % e 23,40 % deles apresentam elementos *cis*-atuantes das famílias *ABRE* e *CEIF* respectivamente, 82,98 % da família *JARE* (ácido jasmônico), 40,43 % da família *AREF* (auxinas), 29,79 % da família *BRRE* (brassinosteróides), 44,68 % da família *DREB* e 25,53 % da família *EREF*.

### **5.7. NÍVEIS DE PRODUÇÃO DE ETILENO E DE ACC LIVRE EM SOJA.**

Com a tentativa de correlacionar os dados obtidos *in silico* com dados fisiológicos, foram mensurados os níveis de ACC livre e de produção de etileno, tanto em raiz quanto em folha das cultivares de soja BR16 e EMBRAPA48, respectivamente suscetível e tolerante à seca. As plantas foram germinadas e cultivadas em condições similares em que foram obtidos os dados dos transcriptomas analisados (germinadas em estufa e cultivadas em hidroponia), inclusive nos mesmos tempos de exposição ao déficit hídrico. De acordo com a análise de parâmetros fisiológicos experimentais (Figura S9), observou-se que as duas cultivares sofreram com a escassez hídrica, mas que a cultivar tolerante à seca foi a que melhor respondeu ao déficit hídrico, sendo um indício da veracidade desta afirmação o aumento da taxa fotossintética, da condutância estomática e da taxa de transpiração, verificados aos 75 minutos de estresse nesta cultivar. (Figuras S9-A, S9-E e S9-F). Além disso, quanto à eficiência de consumo de água (*WUE*) (Figura S9-H), observou-se que antes dos 75 minutos, a cultivar suscetível à seca obteve um melhor aproveitamento de seus recursos hídricos quando comparado com a cultivar tolerante, mas após este tempo, esta situação se inverteu, evidenciando um maior impacto ocasionado por este estresse na cultivar suscetível.

Ao se comparar os níveis de ACC livre e produção de etileno entre os tecidos, observa-se que os maiores níveis do precursor do fitormônio foram detectados em folhas, e as raízes apresentaram os maiores valores mensurados de produção de etileno (Figura 14). Comparando-se as cultivares, a EMBRAPA48 apresentou em suas folhas os maiores índices de ACC livre enquanto que os níveis de produção de etileno foram variáveis. Na raiz das plantas não estressadas (tempo zero), a cultivar BR16 apresenta um maior valor de ACC livre e a cultivar EMBRAPA48 uma maior produção de etileno (Figura 14-A e 14-B, respectivamente), fato este que não se observa na folha, onde os níveis de ACC livre entre



**Figura 14. Níveis de Produção de Etileno e ACC Livre em Soja Sob Déficit Hídrico.** Foram determinados os valores de produção de etileno e ACC (1-aminociclopropano-1-ácido carboxílico) livre em raiz e folha das cultivares de soja BR16 e EMBRAPA48 em diferentes tempos de déficit hídrico. Os códigos nos gráficos representam: **A1, A2 e A3** - níveis de ACC livre; **B1, B2 e B3** - níveis de produção de etileno. Nos gráficos de barra **A1, A2, B1 e B2**, a análise estatística foi obtida comparando-se os dados entre tecidos semelhantes das duas cultivares em um mesmo tempo de déficit hídrico. Já nas curvas de tendência **A3 e B3**, a estatística aplicada foi por comparação dos dados da planta não estressada (tempo zero) com a estressada (demais tempos de déficit hídrico). Sendo assim, a diferença entre os gráficos **1 e 2** com o **3** (tanto em **A** quanto em **B**) é a análise estatística. Os asteriscos representam o nível de significância estatística: (\*)  $p \leq 0,05$ ; (\*\*)  $0,01 \leq p < 0,05$ ; (\*\*\*)  $0,001 \leq p < 0,01$ . Cada ponto representa o valor médio ( $\pm$  desvio padrão) de três repetições, em plantas diferentes. A ausência da representação gráfica do desvio padrão em alguns pontos ocorre pelo fato de seus valores absolutos serem inferiores à escala.

as cultivares é praticamente o mesmo, mas a produção de etileno é máxima na cultivar EMBRAPA48 (Figura 14-A2/B2). Com exceção do intervalo de tempo de 25 a 50 minutos de déficit hídrico, foi observado tanto em raiz quanto em folha da cultivar BR16, que os níveis de ACC livre e de produção de etileno se comportam em ciclos de elevação e redução, onde um aumento no nível de ACC livre coincide com uma redução na produção

de etileno e vice-versa. Ainda com relação à cultivar suscetível à seca, observa-se em ambos os tecidos que o pico de produção de etileno (75 minutos de déficit hídrico) coincide com o menor índice de ACC livre na cultivar. Já com relação à cultivar EMBRAPA48, este padrão cíclico se observa muito menos do que na cultivar suscetível à seca. Observa-se que os picos mais elevados de ACC livre na cultivar tolerante encontram-se respectivamente nos tempos de 75 minutos em raiz e 125 minutos de seca em folha, enquanto que a produção máxima de etileno na raiz da mesma cultivar é detectada nos tempos de 75 e 150 minutos de déficit hídrico (os valores em ambas as condições são iguais, considerando o desvio padrão), sendo que em folhas da mesma cultivar o pico máximo de produção do gás foi no tempo zero (planta não estressada).

## 6. DISCUSSÃO.

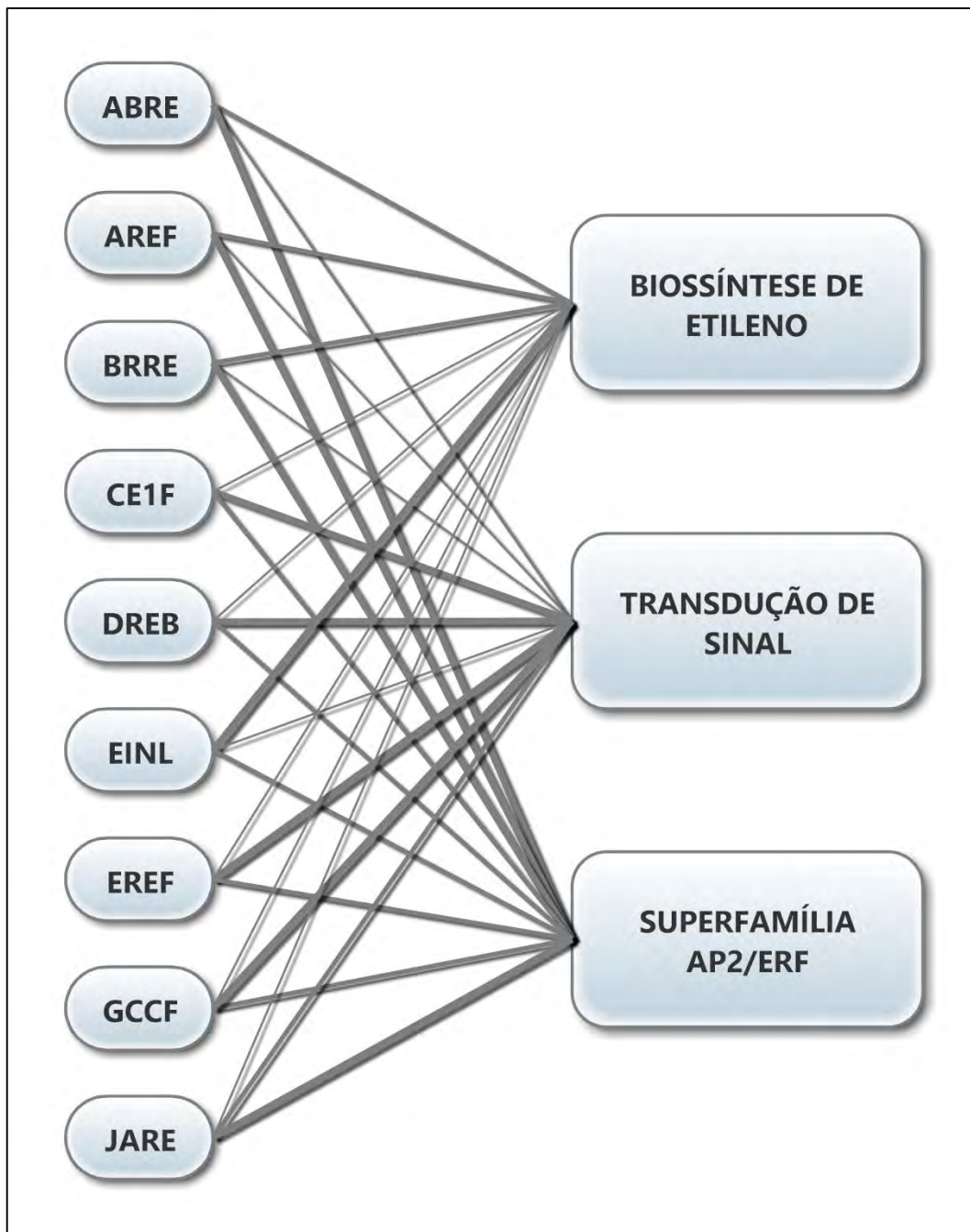
As possíveis proteínas identificadas em soja que fazem parte das rotas metabólicas de biossíntese de etileno e da sinalização por ele mediada são extremamente conservadas, apresentando domínios já descritos para seus homólogos em organismos modelo. O experimento de BBH mostrou que 69,89 % do total de eventos BBH positivos possuem possível ortologia com sequências similares de *A. thaliana* e 58,52 % com *O. sativa*, sendo que 28,41 % do total de eventos positivos são relativos à ortologia com as duas espécies. Além disso, 41,67 % dos genes diferencialmente expressos são BBH positivos sendo que, deste total, 40,00 % são positivos exclusivamente com *A. thaliana*, 28,00 % com *O. sativa* e 32,00 % com ambas as plantas modelo. Desta forma, é possível constatar neste caso uma maior proximidade filogenética da soja com *A. thaliana*, o que corrobora com o fato das duas serem classificadas como dicotiledôneas, mas uma parcela significativa destas proteínas é conservada nas três espécies vegetais. A análise ontológica mostrou o que era esperado tanto para função quanto para os processos moleculares dos quais possivelmente estas proteínas possam participar, assim como a possível citolocalização de cada uma delas.

A análise destes possíveis promotores gênicos de soja mostrou que a indução e/ou repressão da transcrição de um determinado gene nesta espécie vegetal provavelmente não é regulada somente por fatores de transcrição isolados, mas pela interação de vários elementos proteicos com um conjunto de sequências de DNA reguladoras. Associado à esta hipótese, esta análise *in silico* corroborou com dados experimentais da literatura, obtidos principalmente em plantas modelo, que propuseram a existência de interconexões (*crosstalk*) entre a regulação do metabolismo do etileno com outras vias de desenvolvimento, homeostase e resposta à diversos tipos de estresse. Esta afirmação é comprovada nitidamente ao se observar nestes possíveis promotores analisados a presença de elementos *cis*-atuantes importantes em respostas à outros fitormônios (principalmente ácido abscísico, auxinas, brassinoesteróides e ácido jasmônico), além de elementos importantes em resposta à diversas classes de estresses, tanto bióticos e abióticos (choque térmico, defesa contra patógenos, injúrias mecânicas, dentre outros). Comparando-se os níveis percentuais de possíveis promotores que apresentam elementos *cis*-atuantes para estes fitormônios entre os 502 genes analisados e os 120 diferencialmente expressos nos

transcriptomas de soja com déficit hídrico, observa-se que, ao se comparar as duas populações de possíveis promotores a representação de cada família de fitormônios permanece constante, onde possíveis promotores que apresentam elementos que respondem diretamente à ácido jasmônico (família *JARE*) são os mais abundantes, seguidos dos que possuem elementos responsivos ao ácido abscísico (sendo a família *ABRE* a mais abundante), auxinas (família *AREF*), brassinoesteróides (família *BRRE*) e, por fim a família de elementos *CEIF*, também responsiva à ABA. Com base nestas análises, cada uma destas famílias de elementos *cis*-atuantes pode contribuir de forma distinta na regulação da transcrição de cada grupo de genes de soja analisado, sendo que a família *EINL* é melhor representada em possíveis promotores de genes referentes à biossíntese de etileno, as famílias *CEIF*, *DREB*, *EREF* e *GCCF* com a transdução de sinal mediada por este fitormônio, e as famílias *ABRE*, *AREF*, *BRRE* e *JARE* em possíveis promotores de genes de fatores de transcrição da superfamília AP2/ERF (Figura 15).

Com relação à família clássica de elementos responsivos à etileno (*GCCF*), observou-se um baixo índice de possíveis promotores de soja que apresentaram estes elementos (10,36 %), mas um índice mais elevado (48,80 %) de possíveis promotores com elementos da família *DREB*, similar à *GCCF*, mas que responde inicialmente à fatores de transcrição importantes na resposta à seca (*DREB*), fato este que também se confirma entre os genes diferencialmente expressos de soja, que apresentam elementos das famílias *GCCF* e *DREB* em 14,17 % e 46,67 % de seus possíveis promotores, respectivamente. Estudos com ERF1 de *A. thaliana* mostraram que este fator de transcrição pode reconhecer elementos de ambas as famílias, de acordo com a situação de estresse. Experimentos com imunoprecipitação de cromatina mostraram que ERF1 interage exclusivamente com elementos da família *GCCF* em resposta à estresse biótico e exclusivamente com elementos da família *DREB* em resposta à estresse abiótico, onde o fator de transcrição interagiu com elementos de uma ou outra família, nunca das duas ao mesmo tempo (Cheng et al., 2013).

Um ponto importante a ser analisado é o fato de que a maioria dos trabalhos que envolvem a biossíntese de etileno e a sinalização por ele mediada correlacionam a atividade de EIN3/EIL principalmente com a indução da transcrição de genes da superfamília AP2/ERF. Ao serem analisadas estas 502 possíveis regiões promotoras de soja, observou-se a presença de elementos *cis*-atuantes responsáveis a este fator de



**Figura 15. Contribuição das Principais Famílias de Elementos *cis*-Atuantes Presentes em Possíveis Promotores de Soja Importantes na Sinalização Mediada por Fitormônios e em Respostas à Estresses.** O diagrama correlaciona o número de possíveis promotores de soja e o número de elementos *cis*-atuantes presentes em cada grupo analisado: biossíntese de etileno, transdução de sinal mediada por este fitormônio e fatores de transcrição da superfamília AP2/ERF. As espessuras das linhas estão diretamente relacionadas com a contribuição de cada família de elementos *cis*-atuantes em cada grupo, sendo que as menos espessas correspondem à um menor número de elementos e um menor número de possíveis promotores que os apresentam, e as mais espessas apresentam um maior número de elementos e um maior número de possíveis promotores que os apresentam. *ABRE* - elementos de resposta à ABA; *AREF* - elementos de resposta à auxina; *BRRE* - elementos de resposta à brassinoesteróides; *CE1F* - elementos de acoplamento 1; *DREB* - proteínas com elementos de interação responsivos à desidratação; *EINL* - elementos de interação com fatores semelhantes à EIN3; *EREF* - elementos responsivos à etileno; *GCCF* - elementos *GCC-box*; *JARE* - elementos de resposta à jasmonato.

transcrição em 63,35 % desse total (repetindo-se este índice entre os genes diferencialmente expressos identificados nos transcriptomas). Mesmo que a presença por si só deste elemento não seja suficiente para afirmar categoricamente que EIN3/EIL regulem diretamente a expressão de outros genes por rotas independentes de AP2/ERF, é importante considerar que sua presença em um grande número de possíveis promotores possa ser pelo menos um indício da veracidade desta observação. Análises de transcriptomas de *A. thaliana* identificaram genes diferencialmente expressos regulados em mutantes *ein3/eil1* confrontados com a planta selvagem suplementadas com ferro. Dentre estes genes, foi observado que o *fator de transcrição induzido por deficiência em ferro 1* (FIT) pode ser um possível parceiro de EIN3/EIL, que por sua vez contribuem no aumento dos níveis de expressão dos genes regulados por FIT, além de promover a expressão do próprio FIT (Bauer & Blondet, 2011). Associado á este pensamento observa-se também que EIN3/EIL são capazes de regular a expressão de EBF2, importante proteína *F-box* responsável pela regulação da atividade de EIN3/EIL (Vandenbussche et al., 2012).

Uma observação interessante é que 95,22 % dos possíveis promotores de soja analisados possuem elementos *cis*-atuantes da família *LREM* (elementos responsivos à luz, não modulados por diferentes tipos de luz) e mais de 85,00 % deles apresentam elementos da família *HEAT* (elementos de choque térmico) (Figura 13). Este fato corrobora com estudos que demonstram em *A. thaliana* que a resposta à baixa intensidade de luz pode ser regulada por etileno e auxinas (por intermédio do aumento da expressão dos genes *AUX22*, *ACS6*, *ACS8* e *ACS9*), além de que a biossíntese de etileno e a sinalização por ele mediada, também reguladas por fitocromo B, são afetadas por anti-fase de luz (dia e noite) e ciclos de temperatura (Vandenbussche et al., 2003; Bours et al., 2013). Estudos complementares com caules de ervilhas estioladas mostram que além da intensidade, a qualidade da luz também pode influenciar rotas envolvendo etileno, uma vez que a luz vermelha também regula a biossíntese de etileno e o gravitropismo nesta espécie vegetal (Steed et al., 2004). Além disso, estudos com mutantes insensíveis à etileno, principalmente relacionados com seus receptores ou ortólogos à EIN2, tipicamente apresentam altos níveis de produção de etileno, enquanto que mutantes *ctr1-1* exibem produção de etileno mais baixas do que as plantas selvagens (Thain et al., 2004). Entretanto, em duplos mutantes *ein3/ein11*, apesar de apresentarem fenótipos similares a mutantes *ein2*, produzem baixos níveis de etileno quando crescidos em dias longos mas

produzem níveis mais altos deste fitormônio quando crescidos em escuridão contínua, quando comparados com plantas selvagens, mas ainda mais baixos do que os mensurados nos mutantes *etr1* e *ein2* (An et al., 2010). Desta forma, sugere-se que exista uma rota paralela à EIN3/EIL que seja responsável pelo controle negativo da biossíntese de etileno, sendo este mecanismo dependente de luz. Mecanismos de regulação transcricional, associados com fatores de transcrição responsivos à luz mostram que, elementos da família *LREM* podem induzir ou reprimir genes em resposta à estímulos luminosos de acordo com variações na sequências que flanqueiam elementos da família *GBOX*, que encontram-se adjacentes à estes elementos de resposta à luz (Jiao et al., 2007). Como mais de 83,00 % dos possíveis promotores de soja analisados que são relacionados com a biossíntese de etileno apresentam elementos *cis*-atuantes da família *GBOX*, associados com o alto índice de elementos da família *LREM*, acredita-se que mecanismos envolvendo EIN3/EIL [ou algum(s) fator(es) que reconhece(m) elemento(s) *G-box* regulado(s) por eles] e fatores de transcrição responsivos à luz possam fazer parte dos mecanismos de regulação da biossíntese de etileno em soja.

Dentre as enzimas relacionadas com a biossíntese de etileno, as responsáveis pela produção de AdoMet são as que mais contribuem com rotas metabólicas que são, à princípio, independentes de etileno. Estudos com poliaminas vegetais apontam o seu envolvimento em respostas à diferentes estresses ambientais em *A. thaliana*, onde elas podem atuar em processos de sinalização celular em um complexo *crostalk* mediado por fitormônios, incluindo a regulação por ácido abscísico em estresses abióticos (Alcazar et al., 2010). Sendo as enzimas MAT (SAMS) uma das responsáveis por fornecer substrato para a biossíntese de poliaminas, é lógico pensar que estas enzimas sejam induzidas também por ABA em resposta à estresses abióticos, fato este que é observado em estudos com *Lycopersicon esculentum* (tomate) onde foram detectados altos níveis de transcritos da referida enzima em condições de alta salinidade e tratamento com ABA exógeno (Espartero et al., 1994). Desta forma, pode-se supor que altos níveis de ácido abscísico podem estar relacionados com baixos níveis de etileno por um possível redirecionamento de AdoMet para rotas biossintéticas de poliaminas. No presente estudo, observou-se que dentre as enzimas MAT de soja, 54, 55 % delas possuem elementos *cis*-atuantes da família *ABRE* (resposta à ABA) em seus possíveis promotores, o que indica uma possível indução destes genes em resposta a diversos tipos de estresses abióticos mediadas por este



fitormônio. Dentre os genes diferencialmente expressos em transcriptomas de soja sob déficit hídrico, GmMAT#003, GmMAT#010 e GmMAT#011 possuem em seus promotores tanto elementos de resposta à ABA quanto à etileno, o que indica a participação destas enzimas nestas duas rotas de sinalização.

A presença de elementos *cis*-atuantes responsivos à fitormônios é de relevância significativa na regulação da biossíntese de etileno. Zhang e colaboradores (2009) demonstraram que em tomate o ácido abscísico pode induzir a expressão tanto de genes que codificam as enzimas ACC sintase (ACS) e ACC oxidase (ACO), promovendo a biossíntese de etileno e o amadurecimento de frutos (Zhang et al., 2009). Além disso, estudos mostram que um dos primeiros efeitos da auxina é o aumento da biossíntese deste fitormônio gasoso por aumento da expressão dos genes que codificam ACSs, enzima limitante nesta rota biossintética (Abel & Theologis, 1996). Paralelamente, estudos com milho, mostram que os brassinoesteróides induzem a produção de etileno principalmente via ACO e interagem com auxinas aumentando a produção deste fitormônio gasoso (Lim et al., 2002). Por fim, estudos mostram que o metil-jasmonato pode induzir a atividade de enzimas ACO, diminuindo assim os níveis de ACC livre e elevando a produção de etileno em folhas de oliveiras (Sanz et al., 1993).

Com relação à biossíntese de etileno, algumas questões podem ser levantadas. Correlacionando os dados de transcriptoma, qPCR, produção de etileno constata-se que, nas situações analisadas, os níveis de produção de etileno são distintos quando se comparam diferentes tecidos (radicular e foliar). Confrontando-se os valores de produção de etileno aos 75 minutos de déficit hídrico, observou-se em folhas o ápice de produção do fitormônio na cultivar suscetível à seca (BR16), e uma redução significativa do mesmo parâmetro no mesmo tecido da cultivar tolerante (quando comparado com a mesma cultivar não estressada), sendo que em raízes, as duas cultivares apresentaram picos de máxima produção de etileno (Figura 14). Associado ao fato de que somente a cultivar tolerante apresentou um aumento na condutância estomática e nas taxas de fotossíntese e transpiração nesta mesma condição de estresse (75 minutos de déficit hídrico) (Figura S9), pode-se inferir que, na situação analisada, raiz e folha podem apresentar mecanismos distintos de resposta à etileno e que a produção deste fitormônio em folhas podem estar associada com possíveis mecanismos de tolerância à seca, uma vez que, o etileno pode regular o fechamento de estômatos (Chen et al., 2013). Além disso, a associação predita do

aumento da produção de etileno com a redução dos níveis de ACC livre não ocorre de forma global. Duas situações adversas são observadas: a primeira onde um aumento do ACC livre é acompanhado por um aumento da produção de etileno; e a segunda pelo inverso da primeira (redução dos níveis de ACC livre acompanhada por redução da produção de etileno). Mesmo que a hidrólise de conjugados de ACC seja um assunto ainda contraditório, o primeiro caso poderia ser justificado pelo aumento da produção de ACC livre, tanto pelo aumento significativo da conversão direta de AdoMet em ACC pela enzima ACS quanto pela reação de hidrólise de conjugados de ACC (malonil-ACC), de tal forma que os níveis de ACC livre sejam superiores à capacidade das enzimas ACO disponíveis de produzir etileno, que por sua vez atingiria seu nível máximo na referida condição (Hoffman et al., 1982; Yang & Hoffman, 1984; Jiao et al., 1986; Fluhr et al., 1996). Já o segundo caso pode ocorrer pelo aumento da formação de conjugados de ACC (malonil-ACC e glutamil-ACC) quanto pela degradação do precursor de etileno, por ação das enzimas ACD. Como principal indício da ocorrência desta hipótese pode-se apresentar a expressão diferencial de GmGGT#002 (tanto em raiz quanto em folha de ambas as cultivares) e de GmACD#001 (raiz da cultivar EMBRAPA48) nos momentos em que ocorre redução dos níveis de ACC livre, acompanhada por redução da produção de etileno. Para melhor compreensão destes fatos, é de extrema importância a identificação e caracterização *in vitro* e *in vivo* das rotas moleculares que envolvem a conjugação e degradação de ACC, pois só assim será possível entender os mecanismos mais refinados de regulação da biossíntese de etileno em resposta aos mais variados sinais, além de identificar e o real papel da formação dos conjugados de ACC nestes processos.

Uma questão importante que deve ser considerada é o fato de existir um aumento da produção de etileno nas cultivares de soja sob déficit hídrico, uma vez que estudos mostram que, isoladamente, os níveis deste fitormônio em plantas submetidas a tal estresse geralmente são baixos (Spollen et al., 2000; Sharp, 2002). Para melhor compreender esta observação antagônica, é preciso levar em consideração o sistema em que estas cultivares de soja foram cultivadas, o hidropônico, onde as raízes encontram-se submersas em solução contendo nutrientes. Estudos mostram que quando se compara esta condição com sistemas de solo, mesmo sob situações idênticas de estresse, ocorrem variações nos padrões da expressão de diversos genes (Guimaraes-Dias et al., 2012). Em paralelo, diversos trabalhos mostram que em plantas sob condições de alagamento, os índices de

produção de etileno são mais elevados, quando comparados em situações de seca (Bailey-Serres & Voesenek, 2008). Sendo assim, acredita-se que as raízes das cultivares de soja crescidas em hidroponia possam apresentar respostas moleculares semelhantes às condições de alagamento e, ao serem submetidas imediatamente à um déficit hídrico severo, possam apresentar respostas diferentes das associadas com crescimento em solo, como por exemplo o aumento na produção de etileno. Este fato pode explicar também a menor produção deste gás nas folhas de ambas as cultivares, quando comparadas com raízes nas mesmas condições, uma vez que em tal sistema de crescimento a parte aérea da planta não se encontra submersa, além do baixo desenvolvimento destas cultivares de soja, quando comparado com sistemas de solo.

Quanto à transdução de sinal mediada por etileno estudos com *Rosa hybrida* que mostram que a presença de etileno em flores está diretamente associada com a elevação dos transcritos de CTR (Ma et al., 2006). Em soja, observou-se a presença de elementos *cis*-atuantes pertencentes à família *GCCF* na maioria dos possíveis promotores dos genes homólogos à CTR e EIN2. Este fato induz à suposição de que a rota de sinalização mediada por este fitormônio via ERF1 regule a expressão dos transcritos destes dois elementos importantes na propagação do sinal, mantendo os níveis de CTR elevados para que na ausência do sinal a rota seja prontamente inibida. Associado à este fato, observa-se que dentre os transcritos referentes à CTR detectados como diferencialmente expressos (GmCTR#001, GmCTR#003 e GmCTR#004), identificados exclusivamente em transcriptomas de folhas, somente GmCTR#001 não apresenta elementos *cis*-atuantes da família *GCCF* em seu possível promotor (Tabela S87). O fato da expressão de GmCTR#001 em raiz ser maior em plantas não estressadas (Figura 11), acompanhada de uma redução gradativa no decorrer do aumento do tempo de seca mesmo com o aumento da produção de etileno (Figura 14), pode indicar mecanismos adicionais em sua regulação transcricional, como a influência de outros fitormônios, como por exemplo ácido abscísico e auxinas, além da participação de outros parálogos de CTR presentes em soja (GmCTR#003 e GmCTR#004). Em paralelo, o fato de homólogos à EIN3, EIN5 e EBF terem sido detectados como diferencialmente expressos, indica que, em vias gerais, a rota de sinalização mediada por etileno esteja ativa nas situações analisada, principalmente nos tempos iniciais de estresse, pelo menos em raízes da cultivar EMBRAPA48 (tolerante à seca) e folhas da cultivar BR16 (suscetível à seca), uma vez que estudos indicam que na

presença deste fitormônio a expressão transcricional destes genes encontra-se elevada (Vandenbussche et al., 2012). A não detecção da expressão diferencial de alguns genes ou pequenas discrepâncias quando se compara a quantificação de transcritos com os dados obtidos pelos transcriptomas pode estar associada com limitações da metodologia adotada para a construção das bibliotecas subtrativas de cDNA, utilizadas na geração dos transcriptomas, principalmente se a expressão dos genes for mais elevada nas plantas controle (no caso, as não estressadas - tempo zero).

Estudos similares aos descritos neste trabalho foram realizados por Zhang e colaboradores (2008) para membros da superfamília de fatores de transcrição AP2/ERF de soja, onde 148 unigenes que possuem domínios AP2/ERF foram identificados e classificados como integrantes das famílias AP2 (26 genes), ERF (120 genes) e RAV (2 genes). Com relação à família ERF, Zhang também subdividiu os 98 unigenes desse grupo que possuíam seu domínio AP2 completo em 12 grupos de acordo com a classificação proposta por Sakuma (2002) para *A. thaliana* [grupos *A1* até *A6* - subfamília CBF/DREB (36 genes); e *B1* até *B6* - subfamília ERF (62 genes)]. Vale a pena ressaltar que Zhang também não identificou uma sequência homóloga à *Soloísta* de *A. thaliana* em soja (Zhang et al., 2008). Ao se comparar os números obtidos no presente estudo com os obtidos por Zhang, observa-se que a proporção entre eles é mantida, sendo que dentre os grupos analisados, o melhor representado nos dois trabalhos (com um maior número de membros) seria o referente ao grupo *B* (subfamília ERF) (Figura S3). Zhang também identificou a presença de motivos proteicos característicos em ERFs, mas sua pesquisa foi realizada de tal forma que foram identificados motivos em comum com *A. thaliana* e, não necessariamente ao mesmo tempo, com *O. sativa*. Sendo assim, ele selecionou em seu estudo 74 possíveis motivos compartilhados com a dicotiledônea *A. thaliana* e 61 motivos em comum com a monocotiledônea *O. sativa* (Zhang et al., 2008), dentre os quais se destacam os motivos EAR e LWSY, também identificados no presente estudo, sendo o primeiro associado com possíveis mecanismos de repressão transcricional (Kagale & Rozwadowski, 2011) e o segundo ainda sem função definida. Além disso, a similaridade identificada destes possíveis motivos com sequências procarióticas e virais corrobora com análises evolutivas do domínio AP2 que propõem que este domínio possivelmente possa ter origem em cianobactérias (*Tetrahymena thermophila* e *Trichodesmium erythraeum*) e bacteriófagos (*Rb49* e *Felix 01*) (Magnani et al., 2004).

Dentre os fatores de transcrição da superfamília AP2/ERF diferencialmente expressos e que possuem elementos *cis*-atuantes das famílias *EINL* e/ou *GCCF*, destacam-se alguns com relação a sua possível função em soja. Estudos demonstram que plantas transgênicas (incluindo a própria soja), transformadas com o fator de transcrição DREB1D de *A. thaliana* sob controle do promotor gênico CaMV35S (*cauliflower mosaic virus*), apresentam maior tolerância à seca e frio (Guttikonda et al., 2014). No presente trabalho, observou-se a expressão diferencial de um homólogo à DREB1D de *A. thaliana* (GmERF#223), sendo que seu nível de expressão pode chegar a ser elevado em mais de 600 em folhas de soja nas condições analisadas (Figura 12; Figura S7). Um fato interessante é que a expressão deste fator de transcrição em *A. thaliana* é dependente de ácido abscísico (Guttikonda et al., 2014), mas o possível promotor do seu homólogo diferencialmente expresso em soja não possui elementos *cis*-atuantes que respondam diretamente à este e a qualquer outro fitormônio, exceto à etileno (família *EINL*) (Tabela S87). Esta observação leva a algumas suposições, dentre as quais se destacam a de que este homólogo de DREB1D presente em soja seja regulado por etileno, ou que outros sinais possam regular a expressão deste gene por mecanismos ainda não descritos.

Pandey e colaboradores (2005) identificaram em *A. thaliana* um fator de transcrição com domínio AP2 que atua como repressor da resposta molecular mediada por ácido abscísico durante a germinação de sementes e condições de estresse, denominado de ABR1 (repressor ABA 1), sendo que mutantes *abr1* possuem hipersensibilidade à ABA quando submetidos à estas condições (Pandey et al., 2005). Dados de qPCR para um homólogo de ABR1 presente em soja (GmERF#069) mostraram que este gene é diferencialmente expresso nas condições experimentais analisadas, principalmente em raízes, onde o nível de expressão deste transcrito pode chegar a ser elevado em mais de 600 vezes quando comparado com cultivares de soja não estressadas (Figura 12). Os dados de qPCR também mostraram que os níveis deste transcrito são mais elevados em raízes da cultivar suscetível à seca (BR16), o que pode causar uma menor sensibilidade desta cultivar ao ácido abscísico e, o que pode ser uma das explicações da menor tolerância desta cultivar à seca.

Tanto a luz quanto etileno apresentam rotas metabólicas que se interconectam e regulam em muitos aspectos do desenvolvimento vegetal. Uma importante proteína caracterizada funcionalmente em *A. thaliana* que atua à jusante do ponto de convergência entre as rotas de sinalização mediadas por etileno e pela luz seria o fator de transcrição

com domínio AP2 RAP2.4, coordenando múltiplos processos do desenvolvimento desta planta, sendo regulado negativamente pela luz e positivamente por estresses abióticos (salino e seca). Além disso, no núcleo, este fator de transcrição pode interagir com elementos *cis*-atuantes do tipo *GCC-box* (família *GCCF*) e *DRE* (elemento responsivo à desidratação - família *DREB*) (Lin et al., 2008). Em soja, foram identificados quatro homólogos ao RAP2.4 de *A. thaliana* como diferencialmente expressos nos transcriptomas analisados (GmERF#051, GmERF#079, GmERF#161 e GmERF#0198), sendo que dois deles (GmERF#051 e GmERF#079) possuem em seus possíveis promotores elementos *cis*-atuantes que podem responder à sinalização mediada por etileno, dando assim indícios da participação deste fitormônio também neste mecanismo de resposta (Tabela S87).

Outro fator de transcrição desta família detectado como diferencial nos transcriptomas de soja analisados foram dois homólogos ao fator ORA47 (GmERF#047 e GmERF075), elemento chave na biossíntese de ácido jasmônico em *A. thaliana* (Khurshid, 2012), sendo que um deles (GmERF#075) possui na estrutura de seu possível promotor elementos *cis*-atuantes da família *GCCF*, e ambos apresentam elementos da família *JARE*, conhecida por responder diretamente à ácido jasmônico (Tabela S87). Este pode ser mais um indício da ação sinérgica destes dois fitormônios em mecanismos de resposta à seca.

Por fim, não foram somente membros da família ERF de soja que foram detectados como diferencialmente expressos nos transcriptomas analisados, mas também um membro da família RAV (GmRAV#001) (Figura S7). Estudos com RAV1 e seus homólogos mais próximos, RAV1L e RAV2 de *A. thaliana*, mostraram que estes fatores podem ser regulados por fitormônios, sendo reprimidos significativamente por brassinoesteróides e ácido abscísico, que mediam respostas à vários estresses abióticos em plantas. Além disso, em condições de escassez de água, plantas transgênicas que expressam altos níveis de RAV apresentam altos níveis de transpiração, e em condições de alta salinidade, as plantas transgênicas que apresentam altos níveis destes fatores, apresentam a inibição da germinação de sementes, enquanto que mutantes *rav* promovem este evento. Sendo assim, estes estudos sugerem que os fatores de transcrição RAV são reguladores negativos do crescimento e de respostas à seca e alta salinidade, e que seus efeitos de regulação negativa em resposta à estresses bióticos podem atuar independentemente de ABA (Fu et al., 2014). Análises do possível promotor de GmRAV#001 mostram que dentre as principais famílias de elementos *cis*-atuantes nele identificadas que podem estar diretamente relacionadas com

a mediação direta por fitormônios, destacam-se a família *EIN3* (etileno), *ABRE* e *CEIF* (ácido abscísico) e *JARE* (ácido jasmônico). Desta forma, a ausência de elementos responsivos à brassinoesteróides é um indício de que, diferentemente de seu homólogo em *A. thaliana*, GmRAV#001 possivelmente não responderia diretamente a este fitormônio (Tabela S87). Além disso, resultados de qPCR, associados com dados de transcriptoma, mostram que a expressão do transcrito referente à este fator é mais elevada na cultivar suscetível à seca (Figura 12; Figura S7). Desta forma, tendo-se como base os estudos de Fu e colaboradores (2014), pode-se inferir que os altos níveis de expressão deste transcrito também possam estar associados com mecanismos moleculares que tornam esta cultivar (BR16) mais suscetível à esse tipo de estresse.

## 7. CONCLUSÕES FINAIS.

As análises dos bancos de dados públicos e disponibilizados pelo Projeto GENOSOJA propiciaram a elaboração de modelos *in silico* de rotas metabólicas relacionadas com a biossíntese de etileno e com a sinalização por ele mediada, podendo assim mostrar o quão conservadas são estas rotas em espécies vegetais como *Arabidopsis thaliana*, *Oryza sativa* e *Glycine max*. Desta forma, também foi possível evidenciar pontos que ainda necessitam ser melhor estudados, tanto nas três espécies, como caracterização funcional de genes responsáveis pela conjugação de ACC com grupamentos malonil e glutamil, quanto exclusivamente em soja, como caracterização de um ou mais mecanismos de regulação da atividade dos homólogos de EIN2 na ausência de etileno.

A análise dos bancos de dados dos transcriptomas de soja fornecidos pelo Projeto GENOSOJA foi importante para a criação de hipóteses que puderam ser parcialmente elucidadas, mas muitas dúvidas ainda perduraram devido principalmente às limitações da técnica utilizada para a criação das bibliotecas de cDNA, associadas com o sistema de cultivo hidropônico, que em vias de fato, para a *commodity* estudada, também poderia ser considerado como uma situação de estresse.

Supondo-se que, pelo menos em raízes, o sistema hidropônico apresente um comportamento similar ao de alagamento (onde segundo a literatura os níveis de etileno são altos), acredita-se que exista a possibilidade dos níveis deste fitormônio serem elevados já durante o cultivo, e que o mesmo se encontre dissolvido na solução nutritiva (Bailey-Serres & Voesenek, 2008). Acredita-se também que a elevação natural de temperatura ocasionada pela perda de água brusca possa ser uma explicação para a elevação da produção de etileno em alguns momentos, uma vez que enzimas da rota de biossíntese de etileno podem ter um aumento de atividade com a elevação da temperatura (de 17 °C até 35 °C) (Antunes & Sfakiotakis, 2000). Como a difusão do etileno no ar é maior do que na solução nutritiva e o déficit hídrico é mais intenso e a desidratação mais rápida nestas condições experimentais, não existe tempo hábil para efetivar respostas moleculares que envolvam a transcrição de genes e posterior tradução em proteínas antes da dessecação irreversível. Seguindo esta linha de raciocínio, sugere-se que nos curtos intervalos de déficit hídrico (25 - 50 minutos), a biossíntese deste fitormônio e a sinalização por ele mediada seja ainda reflexo das condições de cultivo. Na medida em que



o tempo de estresse aumenta, a sinalização é gradativamente reduzida, dando lugar a uma resposta mais intensa à seca. Como as plantas trocaram uma situação de estresse por outra, acredita-se que as rotas metabólicas de resposta à seca ativadas sejam distintas das que são habitualmente descritas para sistemas de solo, por exemplo (dependentes e independentes de ABA, com baixa produção de etileno), uma vez que as rotas de sinalização ativas antes do déficit hídrico também seriam distintas. Desta forma, acredita-se que o etileno, associado com outros fitormônios (como por exemplo ácido jasmônico), possuam um grande papel no início da resposta à seca nestas condições. Confirmando-se esta hipótese, os níveis de expressão de vários genes identificados neste estudo poderiam ser mais elevados nas plantas antes do déficit hídrico, o que dificultaria a detecção dos mesmos em transcriptomas obtidos a partir de bibliotecas subtrativas, além de possíveis variações de seus perfis de expressão.

Em paralelo, os dados sugeriram que a maior suscetibilidade da cultivar BR16 à seca pode estar associada à uma menor sensibilidade da mesma ao ácido abscísico, hipótese esta que foi levantada devido aos experimentos de qPCR mostrarem que a expressão dos transcritos homólogos aos fatores de transcrição ABR1 e RAV ser mais elevada nesta cultivar.

Por fim, este estudo mostrou *in silico* que a regulação da resposta mediada por etileno em soja pode sofrer influências de diversos fatores endógenos e exógenos, e o balanço entre as sinalizações mediadas por estes fatores pode determinar a qualidade e a intensidade da resposta a estes estímulos. Por isso, é imprescindível a elucidação funcional *in vitro* e *in vivo* dos mecanismos moleculares e redes metabólicas que este fitormônio gasoso coordena nesta *commodity* de grande valor econômico agregado, para que no futuro seja possível obter cultivares com maior tolerância aos mais variados tipos de estresses com maior facilidade.

## 8. REFERÊNCIAS BIBLIOGRÁFICAS.

- Abe H, Yamaguchi-Shinozaki K, Urao T, Iwasaki T, Hosokawa D, Shinozaki K** (1997) Role of Arabidopsis MYC and MYB homologs in drought and abscisic acid-regulated gene expression. *The Plant Cell* **9**: 1859-1868
- Abel S, Nguyen MD, Chow W, Theologis A** (1995) ASC4, a primary indoleacetic acid-responsive gene encoding 1-aminocyclopropane-1-carboxylate synthase in *Arabidopsis thaliana*: Structural characterization, expression in *Escherichia coli*, and expression characteristics in response to auxin. *Journal of Biological Chemistry* **270**: 19093-19099
- Abel S, Theologis A** (1996) Early genes and auxin action. *Plant Physiology* **111**: 9-17
- Acaster MA, Kende H** (1983) Properties and partial purification of 1-aminocyclopropane-1-carboxylate synthase. *Plant Physiology* **72**: 139-145
- Achard P, Baghour M, Chapple A, Hedden P, Van Der Straeten D, Genschik P, Moritz T, Harberd NP** (2007) The plant stress hormone ethylene controls floral transition via DELLA-dependent regulation of floral meristem-identity genes. *Proceedings of the National Academy of Sciences* **104**: 6484-6489
- Achard P, Cheng H, De Grauwe L, Decat J, Schoutteten H, Moritz T, Van Der Straeten D, Peng J, Harberd NP** (2006) Integration of plant responses to environmentally activated phytohormonal signals. *Science* **311**: 91-94
- Achard P, Herr A, Baulcombe DC, Harberd NP** (2004) Modulation of floral development by a gibberellin-regulated microRNA. *Development* **131**: 3357-3365
- Achard P, Vriezen WH, Van Der Straeten D, Harberd NP** (2003) Ethylene regulates Arabidopsis development via the modulation of DELLA protein growth repressor function. *The Plant Cell* **15**: 2816-2825
- Addicott FT, Lyon JL** (1969) Physiology of abscisic acid and related substances. *Annual Review of Plant Physiology* **20**: 139-164
- Agarwal P, Agarwal PK, Nair S, Sopory SK, Reddy MK** (2007) Stress-inducible DREB2A transcription factor from *Pennisetum glaucum* is a phosphoprotein and its phosphorylation negatively regulates its DNA-binding activity. *Molecular Genetics and Genomics* **277**: 189-198
- Ahuja I, De Vos RCH, Bones AM, Hall RD** (2010) Plant molecular stress responses face climate change. *Trends in Plant Science* **15**: 664-674
- Aida M, Beis D, Heidstra R, Willemsen V, Blilou I, Galinha C, Nussaume L, Noh Y-S, Amasino R, Scheres B** (2004) The *Plethora* genes mediate patterning of the Arabidopsis root stem cell niche. *Cell* **119**: 109-120
- Alcazar R, Altabella T, Marco F, Bortolotti C, Reymond M, Koncz C, Carrasco P, Tiburcio AF** (2010) Polyamines: Molecules with regulatory functions in plant abiotic stress tolerance. *Planta* **231**: 1237-1249
- Allen MD, Yamasaki K, Ohme-Takagi M, Tateno M, Suzuki M** (1998) A novel mode of DNA recognition by a  $\beta$ -sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *17*
- Alonso JM, Hirayama T, Roman G, Nourizadeh S, Ecker JR** (1999) EIN2, a bifunctional transducer of ethylene and stress responses in Arabidopsis. *Science* **284**: 2148-2152
- Alonso JM, Stepanova AN, Solano R, Wisman E, Ferrari S, Ausubel FM, Ecker JR** (2003) Five components of the ethylene-response pathway identified in a screen for

- weak ethylene-insensitive mutants in Arabidopsis. Proceedings of the National Academy of Sciences **100**: 2992-2997
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ** (1990) Basic local alignment search tool. Journal of Molecular Biology **215**: 403-410
- Amrhein N, Dorzok U, Kionka C, Kondziolka U, Skorupka H, Tophof S** (1984) The biochemistry and physiology of 1-aminocyclopropane-1-carboxylic acid conjugation. In Y Fuchs, E Chalutz, eds, Ethylene, Vol 9. Springer Netherlands, pp 11-20
- Amthor JS** (2003) Efficiency of lignin biosynthesis: A quantitative analysis. Annals of Botany **91**: 673-695
- An F, Zhao Q, Ji Y, Li W, Jiang Z, Yu X, Zhang C, Han Y, He W, Liu Y, Zhang S, Ecker JR, Guo H** (2010) Ethylene-induced stabilization of ethylene insensitive 3 and EIN3-like 1 is mediated by proteasomal degradation of EIN3 binding F-box 1 and 2 that requires EIN2 in Arabidopsis. The Plant Cell: 1-18
- Antunes MDC, Sfakiotakis EM** (2000) Effect of high temperature stress on ethylene biosynthesis, respiration and ripening of 'Hayward' kiwifruit. Postharvest Biology and Technology **20**: 251-259
- Apelbaum A, Yang SF** (1981) Biosynthesis of stress ethylene induced by water deficit. Plant Physiology **68**: 594-596
- Aranda MA, Escaler M, Thomas CL, Maule AJ** (1999) A heat shock transcription factor in pea is differentially controlled by heat and virus replication. The Plant Journal **20**: 153-161
- Aravind L, Ponting CP** (1997) The GAF domain: An evolutionary link between diverse phototransducing proteins. Trends in Biochemical Sciences **22**: 458-459
- Arc E, Sechet J, Corbineau F, Rajjou L, Marion-Poll A** (2013) ABA crosstalk with ethylene and nitric oxide in seed dormancy and germination. Frontiers in Plant Science **4**
- Argueso CT, Hansen M, Kieber JJ** (2007) Regulation of ethylene biosynthesis. Journal of Plant Growth Regulation **26**: 92-105
- Arthurs S, Heinz KM, Prasifka JR** (2004) An analysis of using entomopathogenic nematodes against above-ground pests. Bulletin of Entomological Research **94**: 297-306
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G** (2000) Gene ontology: Tool for the unification of biology. Nature Genetics **25**: 25-29
- Avila CJ, Vivian LM, Tomquelski GV** (2013) Ocorrência, aspectos biológicos, danos e estratégias de manejo de *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae) nos sistemas de produção agrícolas. Circular Técnica - Embrapa **23**: 1-12
- Bailey-Serres J, Fukao T, Gibbs DJ, Holdsworth MJ, Lee SC, Licausi F, Perata P, Voosenek LACJ, van Dongen JT** (2012) Making sense of low oxygen sensing. Trends in Plant Science **17**: 129-138
- Bailey-Serres J, Voosenek LACJ** (2008) Flooding stress: Acclimations and genetic diversity. Annual Review of Plant Biology **59**: 313-339
- Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS** (2009) MEME suite: Tools for motif discovery and searching. Nucleic Acids Research **37**: W202-W208

- Bari R, Jones JDG** (2009) Role of plant hormones in plant defence responses. *Plant Molecular Biology* **69**: 473-488
- Bauer P, Blondet E** (2011) Transcriptome analysis of *ein3 eil1* mutants in response to iron deficiency. *Plant Signaling & Behavior* **6**: 1669-1671
- Beaudoin N, Serizet C, Gosti F, Giraudat J** (2000) Interactions between abscisic acid and ethylene signaling cascades. *The Plant Cell* **12**: 1103-1115
- Belin C, Megies C, Hauserova E, Lopez-Molina L** (2009) Abscisic acid represses growth of the Arabidopsis embryonic axis after germination by enhancing auxin signaling. *The Plant Cell* **21**: 2253-2268
- Benjamin JG, Nielsen DC** (2006) Water deficit effects on root distribution of soybean, field pea and chickpea. *Field Crops Research* **97**: 248-253
- Benko-Iseppon AM, Nepomuceno AL, Abdelnoor RV** (2012) GENOSOJA - the Brazilian soybean genome consortium: High throughput omics and beyond. *Genetics and Molecular Biology* **35**: i-iv
- Benschop JJ, Bou J, Peeters AJM, Wagemaker N, Guhl K, Ward D, Hedden P, Moritz T, Voeselek LACJ** (2006) Long-term submergence-induced elongation in *Rumex palustris* requires abscisic acid-dependent biosynthesis of gibberellin. *Plant Physiology* **141**: 1644-1652
- Bethke G, Unthan T, Uhrig JF, Poschl Y, Gust AA, Scheel D, Lee J** (2009) FLG22 regulates the release of an ethylene response factor substrate from MAP kinase 6 in *Arabidopsis thaliana* via ethylene signaling. *Proceedings of the National Academy of Sciences* **106**: 8067-8072
- Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, Schwede T** (2014) SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research* **42**: W252-W258
- Binder BM, Rodriguez FI, Bleecker AB** (2010) The copper transporter responsive-to-antagonist 1 (RAN1) is essential for the biogenesis of ethylene receptors in Arabidopsis. *Journal of Biological Chemistry*
- Bisson MM, Bleckmann A, Allekotte S, Groth G** (2009) EIN2, the central regulator of ethylene signalling, is localized at the ER membrane where it interacts with the ethylene receptor ETR1. *Biochemical Journal* **424**: 1-6
- Bisson MMA, Groth G** (2010) New insight in ethylene signaling: Autokinase activity of ETR1 modulates the interaction of receptors and EIN2. *Molecular Plant* **3**: 882-889
- Bjorklund S, Antti H, Uddestrand I, Moritz T, Sundberg B** (2007) Cross-talk between gibberellin and auxin in development of *Populus* wood: Gibberellin stimulates polar auxin transport and has a common transcriptome with auxin. *The Plant Journal* **52**: 499-511
- Bleecker AB** (1999) Ethylene perception and signaling: An evolutionary perspective. *Trends in Plant Science* **4**: 269-274
- Blom N, Gammeltoft S, Brunak S** (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of Molecular Biology* **294**: 1351-1362
- Blum A** (2011) Plant water relations, plant stress and plant production. *In* Plant Breeding for Water-Limited Environments. Springer New York, pp 11-52
- Boerma HR, Specht JE** (2004) Soybeans: Improvement, production, and uses. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America

- Bol JF, Buchel AS, Knoester M, Baladin T, Van Loon LC, Linthorst HJM** (1996) Regulation of the expression of plant defence genes. *In* AR Smith, AW Berry, NVJ Harpham, IE Moshkov, GV Novikova, ON Kulaeva, MA Hall, eds, Plant Hormone Signal Perception and Transduction. Springer Netherlands, pp 135-139
- Bolton MD, Thomma BPHJ, Nelson BD** (2006) *Sclerotinia sclerotiorum* (Lib.) de Bary: Biology and molecular traits of a cosmopolitan pathogen. *Molecular Plant Pathology* **7**: 1-16
- Booth HS, Campbell MB** (1929) Studies of anesthetic ethylene: I. The odor of ethylene. *Anesthesia & Analgesia* **8**: 221-226
- Boter M, Ruiz-Rivero O, Abdeen A, Prat S** (2004) Conserved MYC transcription factors play a key role in jasmonate signaling both in tomato and Arabidopsis. *Genes & Development* **18**: 1577-1591
- Bours R, Van Zanten M, Pierik R, Bouwmeester H, Van Der Krol A** (2013) Antiphase light and temperature cycles affect phytochrome B-controlled ethylene sensitivity and biosynthesis, limiting leaf movement and growth of Arabidopsis. *Plant Physiology* **163**: 882-895
- Bouzayen M, Latche A, Alibert G, Pech J-C** (1988) Intracellular sites of synthesis and storage of 1-(malonylamino)cyclopropane-1-carboxylic acid in *Acer pseudoplatanus* cells. *Plant Physiology* **88**: 613-617
- Bouzayen M, Latche A, Pech J-C, Marigo G** (1989) Carrier-mediated uptake of 1-(malonylamino)cyclopropane-1-carboxylic acid in vacuoles isolated from *Catharanthus roseus* cells. *Plant Physiology* **91**: 1317-1322
- Boyer JS** (1982) Plant productivity and environment. *Science* **218**: 443-448
- Boyle B, Brisson N** (2001) Repression of the defense gene PR-10a by the single-stranded DNA binding protein SEBF. *The Plant Cell* **13**: 2525-2537
- Brenner WG, Romanov GA, Kollmer I, Burkle L, Schmülling T** (2005) Immediate-early and delayed cytokinin response genes of *Arabidopsis thaliana* identified by genome-wide expression profiling reveal novel cytokinin-sensitive processes and suggest cytokinin action through transcriptional cascades. *The Plant Journal* **44**: 314-333
- Brugiere N, Jiao S, Hantke S, Zinselmeier C, Roessler JA, Niu X, Jones RJ, Habben JE** (2003) Cytokinin oxidase gene expression in maize is localized to the vasculature, and is induced by cytokinins, abscisic acid, and abiotic stress. *Plant Physiology* **132**: 1228-1240
- Burg SP, Burg EA** (1967) Molecular requirements for the biological activity of ethylene. *Plant Physiology* **42**: 144-152
- Busk PK, Pages M** (1998) Regulation of abscisic acid-induced transcription. *Plant Molecular Biology* **37**: 425-435
- Caillaud M-C, Dubreuil G, Quentin M, Perfus-Barbeoch L, Lecomte P, de Almeida Engler J, Abad P, Rosso M-N, Favery B** (2008) Root-knot nematodes manipulate plant cell functions during a compatible interaction. *Journal of Plant Physiology* **165**: 104-113
- Cano-Delgado A, Yin Y, Yu C, Vafeados D, Mora-Garcia S, Cheng J-C, Nam KH, Li J, Chory J** (2004) BRL1 and BRL3 are novel brassinosteroid receptors that function in vascular differentiation in Arabidopsis. *Development* **131**: 5341-5351
- Capitani G, Hohenester E, Feng L, Storici P, Kirsch JF, Jansonius JN** (1999) Structure of 1-aminocyclopropane-1-carboxylate synthase, a key enzyme in the biosynthesis of the plant hormone ethylene. *Journal of Molecular Biology* **294**: 745-756

- Cara B, Giovannoni JJ** (2008) Molecular biology of ethylene during tomato fruit development and maturation. *Plant Science* **175**: 106-113
- Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M, Bayerlein M, Werner T** (2005) MatInspector and beyond: Promoter analysis based on transcription factor binding sites. *Bioinformatics* **21**: 2933-2942
- Cary AJ, Liu W, Howell SH** (1995) Cytokinin action is coupled to ethylene in its effects on the inhibition of root and hypocotyl elongation in *Arabidopsis thaliana* seedlings. *Plant Physiology* **107**: 1075-1082
- Casagrande EC, Farias JRB, Neumaier N, Oya T, Pedrosa J, Martins PK, Breton MC, Nepomuceno AL** (2001) Expressão gênica diferencial durante déficit hídrico em soja. *Revista Brasileira de Fisiologia Vegetal* **13**: 168-184
- Cavalett O, Ortega E** (2009) Emergy, nutrients balance, and economic assessment of soybean production and industrialization in Brazil. *Journal of Cleaner Production* **17**: 762-771
- Chae HS, Faure F, Kieber JJ** (2003) The *eto1*, *eto2*, and *eto3* mutations and cytokinin treatment increase ethylene biosynthesis in *Arabidopsis* by increasing the stability of ACS protein. *The Plant Cell* **15**: 545-559
- Chandler JW** (2009) Auxin as compere in plant hormone crosstalk. *Planta* **231**: 1-12
- Chang C, Kwok SF, Bleecker AB, Meyerowitz EM** (1993) *Arabidopsis* ethylene-response gene ETR1: Similarity of product to two-component regulators. *Science* **262**: 539-544
- Chao Q, Rothenberg M, Solano R, Roman G, Terzaghi W, Ecker† JR** (1997) Activation of the ethylene gas response pathway in *Arabidopsis* by the nuclear protein Ethylene-Insensitive 3 and related proteins. *Cell* **89**: 1133-1144
- Chen H, Banerjee AK, Hannapel DJ** (2004) The tandem complex of BEL and KNOX partners is required for transcriptional repression of *ga20ox1*. *The Plant Journal* **38**: 276-284
- Chen LIN, Dodd IC, Davies WJ, Wilkinson S** (2013) Ethylene limits abscisic acid- or soil drying-induced stomatal closure in aged wheat leaves. *Plant, Cell & Environment* **36**: 1850-1859
- Chen R, Binder BM, Garrett WM, Tucker ML, Chang C, Cooper B** (2011) Proteomic responses in *Arabidopsis thaliana* seedlings treated with ethylene. *Molecular BioSystems* **7**: 2637-2650
- Chen Y-F, Etheridge N, Schaller GE** (2005) Ethylene signal transduction. *Annals of Botany* **95**: 901-915
- Cheng M-C, Liao P-M, Kuo W-W, Lin T-P** (2013) The *Arabidopsis* ethylene response factor 1 regulates abiotic stress-responsive gene expression by binding to different *cis*-acting elements in response to different stress signals. *Plant Physiology* **162**: 1566-1582
- Chomczynski P, Sacchi N** (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry* **162**: 156-159
- Chung G, Singh RJ** (2008) Broadening the genetic base of soybean: A multidisciplinary approach. *Critical Reviews in Plant Sciences* **27**: 295-341
- Chung M-C, Chou S-J, Kuang L-Y, Charng Y-y, Yang SF** (2002) Subcellular localization of 1-aminocyclopropane-1-carboxylic acid oxidase in apple fruit. *Plant and Cell Physiology* **43**: 549-554

- Clark KL, Larsen PB, Wang X, Chang C** (1998) Association of the Arabidopsis CTR1 Raf-like kinase with the ETR1 and ERS ethylene receptors. *Proceedings of the National Academy of Sciences* **95**: 5401-5406
- Clouse SD, Sasse JM** (1998) Brassinosteroids: Essential regulators of plant growth and development. *Annual Review of Plant Physiology and Plant Molecular Biology* **49**: 427-451
- Conesa A, Gotz S, García-Gomez JM, Terol J, Talon M, Robles M** (2005) Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**: 3674-3676
- Cunningham JP, Zalucki MP, West SA** (1999) Learning in *Helicoverpa armigera* (Lepidoptera: Noctuidae): A new look at the behaviour and control of a polyphagous pest. *Bulletin of Entomological Research* **89**: 201-207
- Czepak C, Albernaz KC, Vivan LM, Guimaraes HO, Carvalhais T** (2013) Primeiro registro de ocorrência de *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae) no Brasil. *Pesquisa Agropecuária Tropical* **43**: 110-113
- Dai Z, Gao J, An K, Lee JM, Edwards GE, An G** (1996) Promoter elements controlling developmental and environmental regulation of a tobacco ribosomal protein gene L34. *Plant Molecular Biology* **32**: 1055-1065
- Dalquen DA, Dessimoz C** (2013) Bidirectional best hits miss many orthologs in duplication-rich clades such as plants and animals. *Genome Biology and Evolution* **5**: 1800-1806
- Darwin C, Darwin F** (1881) *The power of movement in plants*. Appleton, New York
- Datta K, Baisakh N, Ganguly M, Krishnan S, Yamaguchi Shinozaki K, Datta SK** (2012) Overexpression of Arabidopsis and rice stress genes inducible transcription factor confers drought and salinity tolerance to rice. *Plant Biotechnology Journal* **10**: 579-586
- Davies PJ** (1995) *The plant hormones: Their nature, occurrence, and functions*. In P Davies, ed, *Plant Hormones*. Springer Netherlands
- Davis EL, Hussey RS, Baum TJ** (2004) Getting to the roots of parasitism by nematodes. *Trends in Parasitology* **20**: 134-141
- Deikman J, Xu R, Kneissl ML, Ciardi JA, Kim K-N, Pelah D** (1998) Separation of *cis*-elements responsive to ethylene, fruit development, and ripening in the 5'-flanking region of the ripening-related E8 gene. *Plant Molecular Biology* **37**: 1001-1011
- Delker C, Stenzel I, Hause B, Miersch O, Feussner I, Wasternack C** (2006) Jasmonate biosynthesis in *Arabidopsis thaliana* - enzymes, products, regulation. *Plant Biology* **8**: 297-306
- Dello Ioio R, Linhares FS, Sabatini S** (2008) Emerging role of cytokinin as a regulator of cellular differentiation. *Current Opinion in Plant Biology* **11**: 23-27
- Diatchenko L, Lau YF, Campbell AP, Chenchik A, Moqadam F, Huang B, Lukyanov S, Lukyanov K, Gurskaya N, Sverdlov ED, Siebert PD** (1996) Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proceedings of the National Academy of Sciences* **93**: 6025-6030
- Diaz-Martin J, Almoguera C, Prieto-Dapena P, Espinosa JM, Jordano J** (2005) Functional interaction between two transcription factors involved in the developmental regulation of a small heat stress protein gene promoter. *Plant Physiology* **139**: 1483-1494

- Dinh TT, Girke T, Liu X, Yant L, Schmid M, Chen X** (2012) The floral homeotic protein APETALA 2 recognizes and acts through an AT-rich sequence element. *Development* **139**: 1978-1986
- Dong C-H, Rivarola M, Resnick JS, Maggin BD, Chang C** (2008) Subcellular colocalization of Arabidopsis RTE1 and ETR1 supports a regulatory role for RTE1 in ETR1 ethylene signaling. *The Plant Journal* **53**: 275-286
- Doubt SL** (1917) The response of plants to illuminating gas. *Botanical Gazette* **63**: 209-224
- Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L** (2010) MYB transcription factors in Arabidopsis. *Trends in Plant Science* **15**: 573-581
- Dubouzet JG, Sakuma Y, Ito Y, Kasuga M, Dubouzet EG, Miura S, Seki M, Shinozaki K, Yamaguchi-Shinozaki K** (2003) OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *The Plant Journal* **33**: 751-763
- Ecker JR** (2004) Reentry of the ethylene MPK6 module. *The Plant Cell* **16**: 3169-3173
- Egawa C, Kobayashi F, Ishibashi M, Nakamura T, Nakamura C, Takumi S** (2006) Differential regulation of transcript accumulation and alternative splicing of a DREB2 homolog under abiotic stress conditions in common wheat. *Genes & Genetic Systems* **81**: 77-91
- Eisenhart C** (1947) The assumptions underlying the analysis of variance. *Biometrics* **3**: 1-21
- Emery RJN, Reid DM, Chinnappa CC** (1994) Phenotypic plasticity of stem elongation in two ecotypes of *Stellaria longipes*: The role of ethylene and response to wind. *Plant, Cell & Environment* **17**: 691-700
- Erb M, Meldau S, Howe GA** (2012) Role of phytohormones in insect-specific plant reactions. *Trends in Plant Science* **17**: 250-259
- Espartero J, Pintor-Toro JA, Pardo JM** (1994) Differential accumulation of S-adenosylmethionine synthetase transcripts in response to salt stress. *Plant Molecular Biology* **25**: 217-227
- Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen M-y, Pieper U, Sali A** (2001) Comparative protein structure modeling using MODELLER. *In* Current Protocols in Protein Science. John Wiley & Sons, Inc.
- Fehr WR, Caviness CE, Burmood DT, Pennington JS** (1971) Stage of development descriptions for soybeans, *Glycine Max* (L.) Merrill. *Crop Science* **11**: 929-931
- Figuroa P, Browse J** (2012) The Arabidopsis JAZ2 promoter contains a G-box and thymidine-rich module that are necessary and sufficient for jasmonate-dependent activation by MYC transcription factors and repression by JAZ proteins. *Plant and Cell Physiology* **53**: 330-343
- Finlayson SA, Foster KR, Reid DM** (1991) Transport and metabolism of 1-aminocyclopropane-1-carboxylic acid in sunflower (*Helianthus annuus* L.) seedlings. *Plant Physiology* **96**: 1360-1367
- Fluhr R, Mattoo AK, Dilley DR** (1996) Ethylene — biosynthesis and perception. *Critical Reviews in Plant Sciences* **15**: 479-523
- Forcat S, Bennett M, Mansfield J, Grant M** (2008) A rapid and robust method for simultaneously measuring changes in the phytohormones ABA, JA and SA in plants following biotic and abiotic stress. *Plant Methods* **4**: 16-23



- Frederick JR, Camp CR, Bauer PJ** (2001) Drought-stress effects on branch and mainstem seed yield and yield components of determinate soybean *Crop Science* **41**: 759-763
- Friedman M, Brandon DL** (2001) Nutritional and health benefits of soy proteins. *Journal of Agricultural and Food Chemistry* **49**: 1069-1086
- Frigerio M, Alabadí D, Perez-Gómez J, Garcia-Carcel L, Phillips AL, Hedden P, Blazquez MA** (2006) Transcriptional regulation of gibberellin metabolism genes by auxin signaling in *Arabidopsis*. *Plant Physiology* **142**: 553-563
- Fu M, Kang HK, Son S-H, Kim S-K, Nam KH** (2014) A subset of RAV transcription factors modulates drought and salt stress responses ABA-independently in *Arabidopsis*. *Plant and Cell Physiology*
- Fu X, Harberd NP** (2003) Auxin promotes *Arabidopsis* root growth by modulating gibberellin response. *Nature* **421**: 740-743
- Fujita Y, Yoshida T, Yamaguchi-Shinozaki K** (2013) Pivotal role of the AREB/ABF-SnRK2 pathway in ABRE-mediated transcription in response to osmotic stress in plants. *Physiologia Plantarum* **147**: 15-27
- Gao Z, Chen Y-F, Randlett MD, Zhao X-C, Findell JL, Kieber JJ, Schaller GE** (2003) Localization of the Raf-like kinase CTR1 to the endoplasmic reticulum of *Arabidopsis* through participation in ethylene receptor signaling complexes. *Journal of Biological Chemistry* **278**: 34725-34732
- Gardner G** (2009) Plant hormone signaling. *HortScience* **44**: 222-223
- Garrido F, Estrela S, Alves C, Sánchez-Pérez GF, Sillero A, Pajares MA** (2011) Refolding and characterization of methionine adenosyltransferase from *Euglena gracilis*. *Protein Expression and Purification* **79**: 128-136
- Gerashchenkov GA, Rozhnova NA** (2013) The involvement of phytohormones in the plant sex regulation. *Russian Journal of Plant Physiology* **60**: 597-610
- Ghassemian M, Nambara E, Cutler S, Kawaide H, Kamiya Y, McCourt P** (2000) Regulation of abscisic acid signaling by the ethylene response pathway in *Arabidopsis*. *The Plant Cell* **12**: 1117-1126
- Gilmour SJ, Zarka DG, Stockinger EJ, Salazar MP, Houghton JM, Thomashow MF** (1998) Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression. *The Plant Journal* **16**: 433-442
- Glick BR** (2005) Modulation of plant ethylene levels by the bacterial enzyme ACC deaminase. *FEMS Microbiology Letters* **251**: 1-7
- Goellner K, Loehrer M, Langenbach C, Conrath UWE, Koch E, Schaffrath U** (2010) *Phakopsora pachyrhizi*, the causal agent of Asian soybean rust. *Molecular Plant Pathology* **11**: 169-177
- Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS** (2012) Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research* **40**: D1178-D1186
- Gou J, Strauss SH, Tsai CJ, Fang K, Chen Y, Jiang X, Busov VB** (2010) Gibberellins regulate lateral root formation in *Populus* through interactions with auxin and other hormones. *The Plant Cell* **22**: 623-639
- Greenboim-Wainberg Y, Maymon I, Borochoy R, Alvarez J, Olszewski N, Ori N, Eshed Y, Weiss D** (2005) Cross talk between gibberellin and cytokinin: The *Arabidopsis* GA response inhibitor SPINDLY plays a positive role in cytokinin signaling. *The Plant Cell* **17**: 92-102

- Grefen C, Stadele K, Ruzicka K, Obrdlik P, Harter K, Horak J** (2008) Subcellular localization and *in vivo* interactions of the *Arabidopsis thaliana* ethylene receptor family members. *Molecular Plant* **1**: 308-320
- Grimmig B, Gonzalez-Perez MN, Leubner-Metzger G, Vogeli-Lange R, Meins F, Hain R, Penuelas J, Heidenreich B, Langebartels C, Ernst D, Sandermann H** (2003) Ozone-induced gene expression occurs via ethylene-dependent and -independent signalling. *Plant Molecular Biology* **51**: 599-607
- Grossmann K** (2000) Mode of action of auxin herbicides: A new ending to a long, drawn out story. *Trends in Plant Science* **5**: 506-508
- Grossmann K, Kwiatkowski J, Tresch S** (2001) Auxin herbicides induce H<sub>2</sub>O<sub>2</sub> overproduction and tissue damage in cleavers (*Galium aparine* L.). *Journal of Experimental Botany* **52**: 1811-1816
- Gubler F, Jacobsen JV** (1992) Gibberellin-responsive elements in the promoter of a barley high-pI  $\alpha$ -amylase gene. *The Plant Cell* **4**: 1435-1441
- Guilfoyle T, Hagen G, Ulmasov T, Murfett J** (1998) How does auxin turn on genes? *Plant Physiology* **118**: 341-347
- Guimaraes-Dias F, Neves-Borges AC, Viana AAB, Mesquita RO, Romano E, Grossi-de-Sa MdF, Nepomuceno AL, Loureiro ME, Alves-Ferreira M** (2012) Expression analysis in response to drought stress in soybean: Shedding light on the regulation of metabolic pathway genes. *Genetics and Molecular Biology* **35**: 222-232
- Guo H, Ecker JR** (2003) Plant responses to ethylene gas are mediated by SCF<sup>EBF1/EBF2</sup>-dependent proteolysis of EIN3 transcription factor. *Cell* **115**: 667-677
- Guo M, Thomas J, Collins G, Timmermans MCP** (2008) Direct repression of KNOX loci by the asymmetric leaves 1 complex of Arabidopsis. *The Plant Cell* **20**: 48-58
- Gurley WB, Hepburn AG, Key JL** (1979) Sequence organization of the soybean genome. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis* **561**: 167-183
- Guttikonda SK, Valliyodan B, Neelakandan AK, Tran L-SP, Kumar R, Quach TN, Voothuluru P, Gutierrez-Gonzalez JJ, Aldrich DL, Pallardy SG, Sharp RE, Ho T-HD, Nguyen HT** (2014) Overexpression of AtDREB1D transcription factor improves drought tolerance in soybean. *Molecular Biology Reports*: 1-14
- Hahn A, Harter K** (2009) Mitogen-activated protein kinase cascades and ethylene: Signaling, biosynthesis, or both? *Plant Physiology* **149**: 1207-1210
- Hall BP, Shakeel SN, Schaller GE** (2007) Ethylene receptors: Ethylene perception and signal transduction. *Journal of Plant Growth Regulation* **26**: 118-130
- Hall TA** (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**: 95-98
- Hansen H, Grossmann K** (2000) Auxin-induced ethylene triggers abscisic acid biosynthesis and growth inhibition. *Plant Physiology* **124**: 1437-1448
- Haralampidis K, Milioni D, Rigas S, Hatzopoulos P** (2002) Combinatorial interaction of *cis*-elements specifies the expression of the Arabidopsis AtHsp90-1 gene. *Plant Physiology* **129**: 1138-1149
- He J-X, Gendron JM, Sun Y, Gampala SSL, Gendron N, Sun CQ, Wang Z-Y** (2005) BZR1 is a transcriptional repressor with dual roles in brassinosteroid homeostasis and growth responses. *Science* **307**: 1634-1638

- Hegg EL, Que-Jr L** (1997) The 2-his-1-carboxylate facial triad — an emerging structural motif in mononuclear non-heme iron(II) enzymes. *European Journal of Biochemistry* **250**: 625-629
- Heim MA, Jakoby M, Werber M, Martin C, Weisshaar B, Bailey PC** (2003) The basic helix–loop–helix transcription factor family in plants: A genome-wide study of protein structure and functional diversity. *Molecular Biology and Evolution* **20**: 735-747
- Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J** (2007) qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biology* **8**: R19
- Henikoff JG, Greene EA, Pietrokovski S, Henikoff S** (2000) Increased coverage of protein families with the blocks database servers. *Nucleic Acids Research* **28**: 228-230
- Henikoff S, Henikoff JG** (1994) Protein family classification based on searching a database of blocks. *Genomics* **19**: 97-107
- Henikoff S, Henikoff JG, Alford WJ, Pietrokovski S** (1995) Automated construction and graphical presentation of protein blocks from unaligned sequences. *Gene* **163**: GC17-GC26
- Hernandez D, François P, Farinelli L, Osteras M, Schrenzel J** (2008) De novo bacterial genome sequencing: Millions of very short reads assembled on a desktop computer. *Genome Research*: 1-8
- Hiratsu K, Matsui K, Koyama T, Ohme-Takagi M** (2003) Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in *Arabidopsis*. *The Plant Journal* **34**: 733-739
- Hirayama T, Kieber JJ, Hirayama N, Kogan M, Guzman P, Nourizadeh S, Alonso JM, Dailey WP, Dancis A, Ecker JR** (1999) Responsive to agonist 1, a menkes/wilson disease–related copper transporter, is required for ethylene signaling in *Arabidopsis*. *Cell* **97**: 383-393
- Hirayama T, Shinozaki K** (1996) A *cdc5<sup>+</sup>* homolog of a higher plant, *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* **93**: 13371-13376
- Hoad GV** (1995) Transport of hormones in the phloem of higher plants. *Plant Growth Regulation* **16**: 173-182
- Hoagland DR, Arnon DI** (1938) The water-culture method for growing plants without soil. University of California
- Hobo T, Asada M, Kowyama Y, Hattori T** (1999) ACGT-containing abscisic acid response element (ABRE) and coupling element 3 (CE3) are functionally equivalent. *The Plant Journal* **19**: 679-689
- Hoffman NE, Yang SF, McKeon T** (1982) Identification of 1-(malonylamino)cyclopropane-1-carboxylic acid as a major conjugate of 1-aminocyclopropane-1-carboxylic acid, an ethylene precursor in higher plants. *Biochemical and Biophysical Research Communications* **104**: 765-770
- Hoffman T, Schmidt JS, Zheng X, Bent AF** (1999) Isolation of ethylene-insensitive soybean mutants that are altered in pathogen susceptibility and gene-for-gene disease resistance. *Plant Physiology* **119**: 935-950
- Hosoda K, Imamura A, Katoh E, Hatta T, Tachiki M, Yamada H, Mizuno T, Yamazaki T** (2002) Molecular structure of the GARP family of plant MYB-related

- DNA binding motifs of the Arabidopsis response regulators. *The Plant Cell* **14**: 2015-2029
- Hua J, Chang C, Sun Q, Meyerowitz EM** (1995) Ethylene insensitivity conferred by Arabidopsis ERS gene. *Science* **269**: 1712-1714
- Hua J, Meyerowitz EM** (1998) Ethylene responses are negatively regulated by a receptor gene family in *Arabidopsis thaliana*. *Cell* **94**: 261-271
- Hua J, Sakai H, Nourizadeh S, Chen QG, Bleecker AB, Ecker JR, Meyerowitz EM** (1998) EIN4 and ERS2 are members of the putative ethylene receptor gene family in Arabidopsis. *The Plant Cell* **10**: 1321-1332
- Huh SU, Lee S-B, Kim HH, Paek K-H** (2012) ATAF2, a NAC transcription factor, binds to the promoter and regulates NIT2 gene expression involved in auxin biosynthesis. *Molecules and Cells* **34**: 305-313
- Hwang I, Sheen J, Muller B** (2012) Cytokinin signaling networks. *Annual Review of Plant Biology* **63**: 353-380
- Hymowitz T, Newell CA** (1981) Taxonomy of the genus *Glycine*, domestication and uses of soybeans. *Economic Botany* **35**: 272-288
- Ibrahim H, Hosseini P, Alkharouf N, Hussein E, Gamal El-Din AEK, Aly M, Matthews B** (2011) Analysis of gene expression in soybean (*Glycine max*) roots in response to the root knot nematode *Meloidogyne incognita* using microarrays and KEGG pathways. *BMC Genomics* **12**: 1-16
- Ichimura K, Shinozaki K, Tena G, Sheen J, Henry Y, Champion A, Kreis M, Zhang S, Hirt H, Wilson C, Heberle-Bors E, Ellis BE, Morris PC, Innes RW, Ecker JR, Scheel D, Klessig DF, Machida Y, Mundy J, Ohashi Y, Walker JC** (2002) Mitogen-activated protein kinase cascades in plants: A new nomenclature. *Trends in Plant Science* **7**: 301-308
- Ikeda M, Ohme-Takagi M** (2009) A novel group of transcriptional repressors in Arabidopsis. *Plant and Cell Physiology* **50**: 970-975
- Irsigler AST, Costa MDL, Zhang P, Reis PAB, Dewey RE, Boston RS, Fontes EPB** (2007) Expression profiling on soybean leaves reveals integration of ER- and osmotic-stress pathways. *BMC Genomics* **8**: 1-15
- Izawa T, Foster R, Chua N-H** (1993) Plant bZIP protein DNA binding specificity. *Journal of Molecular Biology* **230**: 1131-1144
- Jaglo-Ottosen KR, Gilmour SJ, Zarka DG, Schabenberger O, Thomashow MF** (1998) Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science* **280**: 104-106
- Jasinski S, Piazza P, Craft J, Hay A, Woolley L, Rieu I, Phillips A, Hedden P, Tsiantis M** (2005) KNOX action in Arabidopsis is mediated by coordinate regulation of cytokinin and gibberellin activities. *Current Biology* **15**: 1560-1565
- Jiao X-Z, Philosoph-Hadas S, Su L-Y, Yang SF** (1986) The conversion of 1-(malonylamino)cyclopropane-1-carboxylic acid to 1-aminocyclopropane-1-carboxylic acid in plant tissues. *Plant Physiology* **81**: 637-641
- Jiao Y, Lau OS, Deng XW** (2007) Light-regulated transcriptional networks in higher plants. *Nature Reviews Genetics* **8**: 217-230
- Jin H, Xu G, Meng Q, Huang F, Yu D** (2013) GmNAC5, a NAC transcription factor, is a transient response regulator induced by abiotic stress in soybean. *The Scientific World Journal* **2013**: 1-6
- Johnson PR, Ecker JR** (1998) The ethylene gas signal transduction pathway: A molecular perspective. *Annual Review of Genetics* **32**: 227-254

- Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, Hunter S** (2014) InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**: 1236-1240
- Joo S, Liu Y, Lueth A, Zhang S** (2008) MAPK phosphorylation-induced stabilization of ACS6 protein is mediated by the non-catalytic C-terminal domain, which also contains the *cis*-determinant for rapid degradation by the 26S proteasome pathway. *The Plant Journal* **54**: 129-140
- Jouannic S, Hamal A, Leprince A-S, Tregear JW, Kreis M, Henry Y** (1999) Characterization of novel plant genes encoding MEKK/STE11 and RAF-related protein kinases. *Gene* **229**: 171-181
- Ju C, Chang C** (2012) Advances in ethylene signalling: Protein complexes at the endoplasmic reticulum membrane. *AoB Plants* **2012**
- Ju C, Yoon GM, Shemansky JM, Lin DY, Ying ZI, Chang J, Garrett WM, Kessenbrock M, Groth G, Tucker ML, Cooper B, Kieber JJ, Chang C** (2012) CTR1 phosphorylates the central regulator EIN2 to control ethylene hormone signaling from the ER membrane to the nucleus in Arabidopsis. *Proceedings of the National Academy of Sciences* **109**: 19486-19491
- Kagale S, Rozwadowski K** (2011) EAR motif-mediated transcriptional repression in plants. *Epigenetics* **6**: 141-146
- Kagaya Y, Ohmiya K, Hattori T** (1999) RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Research* **27**: 470-478
- Kanehisa M** (1997) Linking databases and organisms: GenomeNet resources in Japan. *Trends in Biochemical Sciences* **22**: 442-444
- Kanehisa M, Goto S** (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* **28**: 27-30
- Kawahara Y, de la Bastide M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S, Schwartz DC, Tanaka T, Wu J, Zhou S, Childs KL, Davidson RM, Lin H, Quesada-Ocampo L, Vaillancourt B, Sakai H, Lee S, Kim J, Numa H, Itoh T, Buell CR, Matsumoto T** (2013) Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice* **6**: 1-10
- Kaya C, Tuna AL, Yokas I** (2009) The role of plant hormones in plants under salinity stress. *In* M Ashraf, M Ozturk, HR Athar, eds, *Salinity and Water Stress*, Vol 44. Springer Netherlands, pp 45-50
- Kende H** (1993) Ethylene biosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology* **44**: 283-307
- Kendrick MD, Chang C** (2008) Ethylene signaling: New levels of complexity and regulation. *Current Opinion in Plant Biology* **11**: 479-485
- Khurshid M** (2012) Functional analysis of ORA47, a key regulator of jasmonate biosynthesis in Arabidopsis, Vol 1. *Wohrmann Print Service*
- Kieber JJ, Rothenberg M, Roman G, Feldmann KA, Ecker JR** (1993) CTR1, a negative regulator of the ethylene response pathway in Arabidopsis, encodes a member of the Raf family of protein kinases. *Cell* **72**: 427-441
- Kieber JJ, Schaller GE** (2010) The perception of cytokinin: A story 50 years in the making. *Plant Physiology* **154**: 487-492

- Kim HS, Park BO, Yoo JH, Jung MS, Lee SM, Han HJ, Kim KE, Kim SH, Lim CO, Yun D-J, Lee SY, Chung WS** (2007) Identification of a calmodulin-binding NAC protein as a transcriptional repressor in Arabidopsis. *Journal of Biological Chemistry* **282**: 36292-36302
- Kim SY, Chung H-J, Thomas TL** (1997) Isolation of a novel class of bZIP transcription factors that interact with ABA-responsive and embryo-specification elements in the Dc3 promoter using a modified yeast one-hybrid system. *The Plant Journal* **11**: 1237-1251
- Kizis D, Pages M** (2002) Maize DRE-binding proteins DBF1 and DBF2 are involved in *rab17* regulation through the drought-responsive element in an ABA-dependent pathway. *The Plant Journal* **30**: 679-689
- Klee HJ, Hayford MB, Kretzmer KA, Barry GF, Kishore GM** (1991) Control of ethylene synthesis by expression of a bacterial enzyme in transgenic tomato plants. *The Plant Cell* **3**: 1187-1193
- Knoester M, Van Loon LC, Van Den Heuvel J, Hennig J, Bol JF, Linthorst HJM** (1998) Ethylene-insensitive tobacco lacks nonhost resistance against soil-borne fungi. *Proceedings of the National Academy of Sciences* **95**: 1933-1937
- Koenning SR, Overstreet C, Noling JW, Donald PA, Becker JO, Fortnum BA** (1999) Survey of crop losses in response to phytoparasitic nematodes in the United States for 1994. *Journal of Nematology* **31**: 587-618
- Kohli A, Sreenivasulu N, Lakshmanan P, Kumar PP** (2013) The phytohormone crosstalk paradigm takes center stage in understanding how plants respond to abiotic stresses. *Plant Cell Reports* **32**: 945-957
- Kosugi S, Ohashi Y** (2000) Cloning and DNA-binding properties of a tobacco ethylene-insensitive 3 (EIN3) homolog. *Nucleic Acids Research* **28**: 960-967
- Kotewicz ML, D'Alessio JM, Driftmier KM, Blodgett KP, Gerard GF** (1985) Cloning and overexpression of Moloney murine leukemia virus reverse transcriptase in *Escherichia coli*. *Gene* **35**: 249-258
- Kulcheski FR, Marcelino-Guimaraes FC, Nepomuceno AL, Abdelnoor RV, Margis R** (2010) The use of microRNAs as reference genes for quantitative polymerase chain reaction in soybean. *Analytical Biochemistry* **406**: 185-192
- Kusnetsov V, Landsberger M, Meurer J, Oelmuller R** (1999) The assembly of the CAAT-box binding complex at a photosynthesis gene promoter is regulated by light, cytokinin, and the stage of the plastids. *Journal of Biological Chemistry* **274**: 36009-36014
- Lanahan MB, Ho TH, Rogers SW, Rogers JC** (1992) A gibberellin response complex in cereal  $\alpha$ -amylase gene promoters. *The Plant Cell* **4**: 203-211
- Lata C, Yadav A, Prasad M** (2011) Role of plant transcription factors in abiotic stress tolerance,
- Leubner-Metzger G, Petruzzelli L, Waldvogel R, Vogeli-Lange R, Meins F, Jr.** (1998) Ethylene-responsive element binding protein (EREBP) expression and the transcriptional regulation of class I  $\beta$ -1,3-glucanase during tobacco seed germination. *Plant Molecular Biology* **38**: 785-795
- Levitt J** (1980) Responses of plants to environmental stresses. Volume II. Water, radiation, salt, and other stresses. Academic Press.
- Lewis DR, Negi S, Sukumar P, Muday GK** (2011) Ethylene inhibits lateral root development, increases IAA transport and expression of PIN3 and PIN7 auxin efflux carriers. *Development* **138**: 3485-3495

- Li H, Johnson P, Stepanova A, Alonso JM, Ecker JR** (2004) Convergence of signaling pathways in the control of differential cell growth in Arabidopsis. *Developmental Cell* **7**: 193-204
- Li J-F, Qu L-H, Li N** (2005) Tyr152 plays a central role in the catalysis of 1-aminocyclopropane-1-carboxylate synthase. *Journal of Experimental Botany* **56**: 2203-2210
- Li J, Dai X, Zhao Y** (2006) A role for auxin response factor 19 in auxin and ethylene signaling in Arabidopsis. *Plant Physiology* **140**: 899-908
- Li L, Yu X, Thompson A, Guo M, Yoshida S, Asami T, Chory J, Yin Y** (2009) Arabidopsis MYB30 is a direct target of BES1 and cooperates with BES1 to regulate brassinosteroid-induced gene expression. *The Plant Journal* **58**: 275-286
- Li R, Li Y, Kristiansen K, Wang J** (2008) SOAP: Short oligonucleotide alignment program. *Bioinformatics* **24**: 713-714
- Liang X, Abel S, Keller JA, Shen NF, Theologis A** (1992) The 1-aminocyclopropane-1-carboxylate synthase gene family of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* **89**: 11046-11050
- Liang X, Oono Y, Shen NF, Köhler C, Li K, Scolnik PA, Theologis A** (1995) Characterization of two members (ACS1 and ACS3) of the 1-aminocyclopropane-1-carboxylate synthase gene family of *Arabidopsis thaliana*. *Gene* **167**: 17-24
- Licausi F, Ohme-Takagi M, Perata P** (2013) APETALA 2/ethylene responsive factor (AP2/ERF) transcription factors: Mediators of stress responses and developmental programs. *New Phytologist* **199**: 639-649
- Lim SH, Chang SC, Lee JS, Kim S-K, Kim SY** (2002) Brassinosteroids affect ethylene production in the primary roots of maize (*Zea mays* L.). *Journal of Plant Biology* **45**: 148-153
- Lin R-C, Park H-J, Wang H-Y** (2008) Role of Arabidopsis RAP2.4 in regulating light- and ethylene-mediated developmental processes and drought stress tolerance. *Molecular Plant* **1**: 42-57
- Lin Z, Zhong S, Grierson D** (2009) Recent advances in ethylene research. *Journal of Experimental Botany* **60**: 3311-3336
- Liu J-X, Howell SH** (2010) bZIP28 and NF-Y transcription factors are activated by ER stress and assemble into a transcriptional complex to regulate stress response genes in Arabidopsis. *The Plant Cell* **22**: 782-796
- Liu K** (1997) Chemistry and nutritional value of soybean components. *In* Soybeans. Springer US, pp 25-113
- Liu M, Zhu S, Zhou J** (2012) Determination of 1-aminocyclopropane-1-carboxylic acid in apple and peach extracts by high performance liquid chromatography coupled to a fluorescence detector. *Analytical Letters* **45**: 2324-2333
- Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K** (1998) 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* **10**: 1391-1406
- Liu Y, Gai JY, Lu HN, Wang YJ, Chen SY** (2005) Identification of drought tolerant germplasm and inheritance and QTL mapping of related root traits in soybean (*Glycine max* (L.) Merr.). *Yi Chuan Xue Bao* **32**: 855-863

- Liu Y, Su L-Y, Yang SF** (1985) Ethylene promotes the capability to malonylate 1-aminocyclopropane-1-carboxylic acid and D-amino acids in preclimacteric tomato fruits. *Plant Physiology* **77**: 891-895
- Liu Y, Zhang S** (2004) Phosphorylation of 1-aminocyclopropane-1-carboxylic acid synthase by MPK6, a stress-responsive mitogen-activated protein kinase, induces ethylene biosynthesis in Arabidopsis. *The Plant Cell* **16**: 3386-3399
- Lopes-Caitar V, Carvalho M, Darben L, Kuwahara M, Nepomuceno A, Dias W, Abdelnoor R, Marcelino-Guimaraes F** (2013) Genome-wide analysis of the Hsp20 gene family in soybean: Comprehensive sequence, genomic organization and expression profile analysis under abiotic and biotic stresses. *BMC Genomics* **14**: 1-17
- Lorenzo O, Piqueras R, Sanchez-Serrano JJ, Solano R** (2003) Ethylene response factor 1 integrates signals from ethylene and jasmonate pathways in plant defense. *The Plant Cell* **15**: 165-178
- Lu C-A, Ho T-hD, Ho S-L, Yu S-M** (2002) Three novel MYB proteins with one DNA binding repeat mediate sugar and hormone regulation of  $\alpha$ -amylase gene expression. *The Plant Cell* **14**: 1963-1980
- Lu C, Fedoroff N** (2000) A mutation in the Arabidopsis HYL1 gene encoding a dsRNA binding protein affects responses to abscisic acid, auxin, and cytokinin. *The Plant Cell* **12**: 2351-2365
- Lund ST, Stall RE, Klee HJ** (1998) Ethylene regulates the susceptible response to pathogen infection in tomato. *The Plant Cell* **10**: 371-382
- Lurie S, Handros A, Fallik E, Shapira R** (1996) Reversible inhibition of tomato fruit gene expression at high temperature (effects on tomato fruit ripening). *Plant Physiology* **110**: 1207-1214
- Ma N, Tan H, Liu X, Xue J, Li Y, Gao J** (2006) Transcriptional regulation of ethylene receptor and CTR genes involved in ethylene-induced flower opening in cut rose (*Rosa hybrida*) cv. Samantha. *Journal of Experimental Botany* **57**: 2763-2773
- MacMillan J, Suter PJ** (1958) The occurrence of gibberellin A1 in higher plants: Isolation from the seed of runner bean (*Phaseolus multiflorus*). *Naturwissenschaften* **45**: 46
- Magnani E, Sjolander K, Hake S** (2004) From endonucleases to transcription factors: Evolution of the AP2 DNA binding domain in plants. *The Plant Cell* **16**: 2265–2277
- Mainardi JA, Purgatto E, Vieira AJ, Bastos WA, Cordenunsi BR, Oliveira do Nascimento JR, Lajolo FM** (2006) Effects of ethylene and 1-methylcyclopropene (1-MCP) on gene expression and activity profile of  $\alpha$ -1,4-glucan-phosphorylase during banana ripening. *Journal of Agricultural and Food Chemistry* **54**: 7294-7299
- Manavalan LP, Guttikonda SK, Tran L-SP, Nguyen HT** (2009) Physiological and molecular approaches to improve drought resistance in soybean. *Plant and Cell Physiology* **50**: 1260-1276
- Martin MN, Cohen JD, Saftner RA** (1995) A new 1-aminocyclopropane-1-carboxylic acid-conjugating activity in tomato fruit. *Plant Physiology* **109**: 917-926
- Maruyama K, Todaka D, Mizoi J, Yoshida T, Kidokoro S, Matsukura S, Takasaki H, Sakurai T, Yamamoto YY, Yoshiwara K, Kojima M, Sakakibara H, Shinozaki K, Yamaguchi-Shinozaki K** (2012) Identification of *cis*-acting promoter elements in cold and dehydration induced transcriptional pathways in Arabidopsis, rice, and soybean. *DNA Research* **19**: 37-49



- Maruyama Y, Yamoto N, Suzuki Y, Chiba Y, Yamazaki K-i, Sato T, Yamaguchi J** (2013) The Arabidopsis transcriptional repressor ERF9 participates in resistance against necrotrophic fungi. *Plant Science* **213**: 79-87
- Masoumi H, Masoumi M, Darvish F, Daneshian J, Nourmohammadi G, Habibi D** (2010) Change in several antioxidant enzymes activity and seed yield by water deficit stress in soybean (*Glycine max* L.) cultivars, Vol 38
- Mathieu J, Yant LJ, Mürdter F, Küttner F, Schmid M** (2009) Repression of flowering by the miR172 target SMZ. *PLoS Biology* **7**: e1000148
- Matias-Hernandez L, Aguilar-Jaramillo AE, Marín-Gonzalez E, Suarez-Lopez P, Pelaz S** (2014) RAV genes: Regulation of floral induction and beyond. *Annals of Botany*: 1-12
- Matsui K, Umemura Y, Ohme-Takagi M** (2008) AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in Arabidopsis. *The Plant Journal* **55**: 954-967
- Matsukura S, Mizoi J, Yoshida T, Todaka D, Ito Y, Maruyama K, Shinozaki K, Yamaguchi-Shinozaki K** (2010) Comprehensive analysis of rice DREB2-type genes that encode transcription factors involved in the expression of abiotic stress-responsive genes. *Molecular Genetics and Genomics* **283**: 185-196
- Mayerhofer H, Panneerselvam S, Mueller-Dieckmann J** (2012) Protein kinase domain of CTR1 from *Arabidopsis thaliana* promotes ethylene receptor cross talk. *Journal of Molecular Biology* **415**: 768-779
- Mayne RG, Kende H** (1986) Ethylene biosynthesis in isolated vacuoles of *Vicia faba* L. - requirement for membrane integrity. *Planta* **167**: 159-165
- McClellan CA, Chang C** (2008) The role of protein turnover in ethylene biosynthesis and response. *Plant Science* **175**: 24-31
- McDonnell L, Plett JM, Andersson-Gunneras S, Kozela C, Dugardeyn J, Van Der Straeten D, Glick BR, Sundberg B, Regan S** (2009) Ethylene levels are regulated by a plant encoded 1-aminocyclopropane-1-carboxylic acid deaminase. *Physiologia Plantarum* **136**: 94-109
- Mehrotra R, Sethi S, Zutshi I, Bhalothia P, Mehrotra S** (2013) Patterns and evolution of ACGT repeat *cis*-element landscape across four plant genomes. *BMC Genomics* **14**: 2-11
- Menke FLH, Champion A, Kijne JW, Memelink J** (1999) A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene *Str* interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2. *The EMBO Journal* **18**: 4455-4463
- Miranda VdJ, Coelho RR, Viana AA, de Oliveira Neto OB, Carneiro RM, Rocha TL, de Sa MF, Fragoso RR** (2013) Validation of reference genes aiming accurate normalization of qPCR data in soybean upon nematode parasitism and insect attack. *BMC Research Notes* **6**: 196-205
- Miyasaka S, Medina JC** (1981) A soja no Brasil. Secretaria de Tecnologia Industrial do Ministério da Indústria e Comércio
- Miyawaki K, Matsumoto-Kitano M, Kakimoto T** (2004) Expression of cytokinin biosynthetic isopentenyltransferase genes in Arabidopsis: Tissue specificity and regulation by auxin, cytokinin, and nitrate. *The Plant Journal* **37**: 128-138
- Miyazaki JH, Yang SF** (1987) Metabolism of 5-methylthioribose to methionine. *Plant Physiology* **84**: 277-281

- Morgan PW, Drew MC** (1997) Ethylene and plant responses to stress. *Physiologia Plantarum* **100**: 620-630
- Morrissey DV, Collins ML** (1989) Nucleic acid hybridization assays employing dA-tailed capture probes. Single capture methods. *Molecular and Cellular Probes* **3**: 189-207
- Moubayidin L, Di Mambro R, Sabatini S** (2009) Cytokinin-auxin crosstalk. *Trends in Plant Science* **14**: 557-562
- Muday GK, Rahman A** (2008) Auxin transport and the integration of gravitropic growth. *In Plant Tropisms*. Blackwell Publishing Ltd, pp 47-77
- Nakano T, Suzuki K, Fujimura T, Shinshi H** (2006) Genome-wide analysis of the ERF gene family in Arabidopsis and rice. *Plant Physiology* **140**: 411-432
- Nakashima K, Yamaguchi-Shinozaki K** (2010) Promoters and transcription factors in abiotic stress-responsive gene expression. *In A Pareek, SK Sopory, HJ Bohnert, eds, Abiotic stress adaptation in plants*. Springer Netherlands, pp 199-216
- Nakashima K, Yamaguchi-Shinozaki K** (2013) ABA signaling in stress-response and seed development. *Plant Cell Reports* **32**: 959-970
- Nascimento LCD, Costa GGL, Binneck E, Pereira GAG, Carazzolle MF** (2012) A web-based bioinformatics interface applied to the GENOSOJA project: Databases and pipelines. *Genetics and Molecular Biology* **35**: 203-211
- Nath P, Trivedi PK, Sane VA, Sane AP** (2006) Role of ethylene in fruit ripening. *In N Khan, ed, Ethylene Action in Plants*. Springer Berlin Heidelberg, pp 151-184
- Neljubow DN** (1901) Ueber die horizontale nutation der stengel von *Pisum sativum* und einiger anderer pflanzen. *Beihefte Zum Botanischen Centralblatt*. **Bd. 10**: 128-139
- Nemhauser JL, Mockler TC, Chory J** (2004) Interdependency of brassinosteroid and auxin signaling in Arabidopsis. *PLoS Biology* **2**: e258-e270
- Niblack TL, Lambert KN, Tylka GL** (2006) A model plant pathogen from the kingdom Animalia: *Heterodera glycines*, the soybean cyst nematode. *Annual Review of Phytopathology* **44**: 283-303
- Niu X, Helentjaris T, Bate NJ** (2002) Maize ABI4 binds coupling element 1 in abscisic acid and sugar response genes. *The Plant Cell* **14**: 2565-2575
- Nole-Wilson S, Krizek BA** (2000) DNA binding properties of the Arabidopsis floral development protein AINTEGUMENTA. *Nucleic Acids Research* **28**: 4076-4082
- O'Donnell PJ, Calvert C, Atzorn R, Wasternack C, Leyser HMO, Bowles DJ** (1996) Ethylene as a signal mediating the wound response of tomato plants. *Science* **274**: 1914-1917
- O'Neill DP, Davidson SE, Clarke VC, Yamauchi Y, Yamaguchi S, Kamiya Y, Reid JB, Ross JJ** (2010) Regulation of the gibberellin pathway by auxin and DELLA proteins. *Planta* **232**: 1141-1149
- Ohme-Takagi M, Shinshi H** (1995) Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *The Plant Cell* **7**: 173-182
- Ohta M, Matsui K, Hiratsu K, Shinshi H, Ohme-Takagi M** (2001) Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. *The Plant Cell* **13**: 1959-1968
- Ohta M, Ohme-Takagi M, Shinshi H** (2000) Three ethylene-responsive transcription factors in tobacco with distinct transactivation functions. *The Plant Journal* **22**: 29-38
- Olmedo G, Guo H, Gregory BD, Nourizadeh SD, Aguilar-Henonin L, Li H, An F, Guzman P, Ecker JR** (2006) Ethylene-insensitive 5 encodes a 5'→3'

- exoribonuclease required for regulation of the EIN3-targeting F-box proteins EBF1/2. *Proceedings of the National Academy of Sciences* **103**: 13286-13293
- Olsen AN, Ernst HA, Leggio LL, Skriver K** (2005) NAC transcription factors: Structurally distinct, functionally diverse. *Trends in Plant Science* **10**: 79-87
- Ortega E, Cavalett O, Bonifácio R, Watanabe M** (2005) Brazilian soybean production: Emergency analysis with an expanded scope. *Bulletin of Science, Technology & Society* **25**: 323-334
- Paiva Rolla AA, Fatima Correa Carvalho J, Fuganti-Pagliarini R, Engels C, Rio A, Marin SRR, Oliveira MCN, Beneventi MA, Marcelino-Guimaraes FC, Farias JRB, Neumaier N, Nakashima K, Yamaguchi-Shinozaki K, Nepomuceno AL** (2014) Phenotyping soybean plants transformed with rd29A:AtDREB1A for drought tolerance in the greenhouse and field. *Transgenic Research* **23**: 75-87
- Pandey GK, Grant JJ, Cheong YH, Kim BG, Li L, Luan S** (2005) ABR1, an APETALA2-domain transcription factor that functions as a repressor of ABA response in Arabidopsis. *Plant Physiology* **139**: 1185-1193
- Papin JF, Vahrson W, Dittmer DP** (2004) SYBR green-based real-time quantitative PCR assay for detection of *West Nile Virus* circumvents false-negative results due to strain variability. *Journal of Clinical Microbiology* **42**: 1511-1518
- Pedreno MA, Bouzayen M, Pech JC, Marigo G, Latche A** (1991) Vacuolar release of 1-(malonylamino)cyclopropane-1-carboxylic acid, the conjugated form of the ethylene precursor. *Plant Physiology* **97**: 1483-1486
- Peiser GD, Wang T-T, Hoffman NE, Yang SF, Liu H-w, Walsh CT** (1984) Formation of cyanide from carbon 1 of 1-aminocyclopropane-1-carboxylic acid during its conversion to ethylene. *Proceedings of the National Academy of Sciences* **81**: 3059-3063
- Peleg Z, Blumwald E** (2011) Hormone balance and abiotic stress tolerance in crop plants. *Current Opinion in Plant Biology* **14**: 290-295
- Penninckx IA, Eggermont K, Terras FR, Thomma BP, De Samblanx GW, Buchala A, Métraux JP, Manners JM, Broekaert WF** (1996) Pathogen-induced systemic activation of a plant defensin gene in Arabidopsis follows a salicylic acid-independent pathway. *The Plant Cell* **8**: 2309-2323
- Penninckx IAMA, Thomma BPHJ, Buchala A, Metraux J-P, Broekaert WF** (1998) Concomitant activation of jasmonate and ethylene response pathways is required for induction of a plant defensin gene in Arabidopsis. *The Plant Cell* **10**: 2103-2113
- Pernisova M, Klíma P, Horak J, Valkova M, Malbeck J, Soucek P, Reichman P, Hoyerova K, Dubova J, Friml J, Zazimalova E, Hejatko J** (2009) Cytokinins modulate auxin-induced organogenesis in plants via regulation of the auxin efflux. *Proceedings of the National Academy of Sciences*: 1-6
- Pfaffl MW, Horgan GW, Dempfle L** (2002) Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Research* **30**: e36-e45
- Pieterse CMJ, Van Loon LC** (1999) Salicylic acid-independent plant defence pathways. *Trends in Plant Science* **4**: 52-58
- Pieterse CMJ, Van Wees SCM, Van Pelt JA, Knoester M, Laan R, Gerrits H, Weisbeek PJ, Van Loon LC** (1998) A novel signaling pathway controlling induced systemic resistance in Arabidopsis. *The Plant Cell* **10**: 1571-1580
- Petrokovski S** (1996) Searching databases of conserved sequence regions by aligning protein multiple-alignments. *Nucleic Acids Research* **24**: 3836-3845

- Pirbalouti AG, Sajjadi SE, Parang K** (2014) A review (research and patents) on jasmonic acid and its derivatives. *Archiv der Pharmazie* **347**: 229-239
- Plett MJ, McDonnell L, Regan S** (2009) Plant encoded 1-aminocyclopropane-1-carboxylic acid deaminase activity implicated in different aspects of plant development. *Plant Signaling & Behavior* **4**: 1186-1189
- Pommerrenig B, Feussner K, Zierer W, Rabinovych V, Klebl F, Feussner I, Sauer N** (2011) Phloem-specific expression of Yang cycle genes and identification of novel Yang cycle enzymes in *Plantago* and *Arabidopsis*. *The Plant Cell* **23**: 1904-1919
- Pospisilova J, Synkova H, Haisel D, Batkova P** (2009) Effect of abscisic acid on photosynthetic parameters during *ex vitro* transfer of micropropagated tobacco plantlets. *Biologia Plantarum* **53**: 11-20
- Pre M, Atallah M, Champion A, De Vos M, Pieterse CMJ, Memelink J** (2008) The AP2/ERF domain transcription factor ORA59 integrates jasmonic acid and ethylene signals in plant defense. *Plant Physiology* **147**: 1347-1357
- Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Bournsnel C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, Finn RD** (2012) The PFAM protein families database. *Nucleic Acids Research* **40**: D290-D301
- Qiao H, Chang KN, Yazaki J, Ecker JR** (2009) Interplay between ethylene, ETP1/ETP2 F-box proteins, and degradation of EIN2 triggers ethylene responses in *Arabidopsis*. *Genes & Development* **23**: 512-521
- Qiao H, Shen Z, Huang S-SC, Schmitz RJ, Urich MA, Briggs SP, Ecker JR** (2012) Processing and subcellular trafficking of ER-tethered EIN2 control response to ethylene gas. *Science* **338**: 390-393
- Qu X, Hall B, Gao Z, Schaller GE** (2007) A strong constitutive ethylene-response phenotype conferred on *Arabidopsis* plants containing null mutations in the ethylene receptors ETR1 and ERS1. *BMC Plant Biology* **7**: 1-15
- Qu X, Schaller GE** (2004) Requirement of the histidine kinase domain for signal transduction by the ethylene receptor ETR1. *Plant Physiology* **136**: 2961-2970
- Raskin I, Beyer EM** (1989) Role of ethylene metabolism in *Amaranthus retroflexus*. *Plant Physiology* **90**: 1-5
- Reinhardt D, Kende H, Boiler T** (1994) Subcellular localization of 1-aminocyclopropane-1-carboxylate oxidase in tomato cells. *Planta* **195**: 142-146
- Resnick JS, Wen C-K, Shockey JA, Chang C** (2006) Reversion to ethylene sensitivity 1, a conserved gene that regulates ethylene receptor function in *Arabidopsis*. *Proceedings of the National Academy of Sciences* **103**: 7917-7922
- Rodrigues FA, Marcolino-Gomes J, Carvalho JdFC, Nascimento LCd, Neumaier N, Farias JRB, Carazzolle MF, Marcelino FC, Nepomuceno AL** (2012) Subtractive libraries for prospecting differentially expressed genes in the soybean under water deficit. *Genetics and Molecular Biology* **35**: 304-314
- Rodriguez FI, Esch JJ, Hall AE, Binder BM, Schaller GE, Bleeker AB** (1999) A copper cofactor for the ethylene receptor ETR1 from *Arabidopsis*. *Science* **283**: 996-998
- Roje S** (2006) S-adenosyl-L-methionine: Beyond the universal methyl group donor. *Phytochemistry* **67**: 1686-1698
- Roman G, Lubarsky B, Kieber JJ, Rothenberg M, Ecker JR** (1995) Genetic analysis of ethylene signal transduction in *Arabidopsis thaliana*: Five novel mutant loci integrated into a stress response pathway. *Genetics* **139**: 1393-1409

- Romeis J, Shanower TG** (1996) Arthropod natural enemies of *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae) in India. *Biocontrol Science and Technology* **6**: 481-508
- Romero I, Fuertes A, Benito MJ, Malpica JM, Leyva A, Paz-Ares J** (1998) More than 80 R2R3-MYB regulatory genes in the genome of *Arabidopsis thaliana*. *The Plant Journal* **14**: 273-284
- Ross JJ, O'Neill DP, Smith JJ, Kerckhoffs LHJ, Elliott RC** (2000) Evidence that auxin promotes gibberellin A1 biosynthesis in pea. *The Plant Journal* **21**: 547-552
- Ruzicka K, Ljung K, Vanneste S, Podhorska R, Beeckman T, Friml J, Benkova E** (2007) Ethylene regulates root growth through effects on auxin biosynthesis and transport-dependent auxin distribution. *The Plant Cell* **19**: 2197-2212
- Ruzicka K, Simaskova M, Duclercq J, Petrasek J, Zazimalova E, Simon S, Friml J, Van Montagu MCE, Benkova E** (2009) Cytokinin regulates root meristem activity via modulation of the polar auxin transport. *Proceedings of the National Academy of Sciences* **106**: 4284-4289
- Rzewuski G, Cornell KA, Rooney L, Bürstenbinder K, Wirtz M, Hell R, Sauter M** (2007) OsMTN encodes a 5'-methylthioadenosine nucleosidase that is upregulated during submergence-induced ethylene synthesis in rice (*Oryza sativa* L.). *Journal of Experimental Botany* **58**: 1505-1514
- Sadeghipour O, Abbasi S** (2012) Soybean response to drought and seed inoculation. *World Applied Sciences Journal* **17**: 55-60
- Saibo NJM, Lourenço T, Oliveira MM** (2009) Transcription factors and regulation of photosynthetic and related metabolism under environmental stresses. *Annals of Botany* **103**: 609-623
- Sakai H, Hua J, Chen QG, Chang C, Medrano LJ, Bleecker AB, Meyerowitz EM** (1998) ETR2 is an ETR1-like gene involved in ethylene signaling in Arabidopsis. *Proceedings of the National Academy of Sciences* **95**: 5812-5817
- Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, Yamaguchi-Shinozaki K** (2002) DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochemical and Biophysical Research Communications* **290**: 998-1009
- Santner A, Estelle M** (2009) Recent advances and emerging trends in plant hormone signalling. *Nature* **459**: 1071-1078
- Santner A, Estelle M** (2010) The ubiquitin-proteasome system regulates plant hormone signaling. *The Plant Journal* **61**: 1029-1040
- Sanz LC, Fernandez-Maculet JC, Gomez E, Vioque B, Olias JM** (1993) Effect of methyl jasmonate on ethylene biosynthesis and stomatal closure in olive leaves. *Phytochemistry* **33**: 285-289
- Sauter M, Cornell KA, Beszteri S, Rzewuski G** (2004) Functional analysis of methylthioribose kinase genes in plants. *Plant Physiology* **136**: 4061-4071
- Sauter M, Lorbiecke R, OuYang B, Pochapsky TC, Rzewuski G** (2005) The immediate-early ethylene response gene OsARD1 encodes an acireductone dioxygenase involved in recycling of the ethylene precursor S-adenosylmethionine. *The Plant Journal* **44**: 718-729
- Schaller GE, Bleecker AB** (1995) Ethylene-binding sites generated in yeast expressing the Arabidopsis ETR1 gene. *Science* **270**: 1809-1811

- Schindler U, Terzaghi W, Beckmann H, Kadesch T, Cashmore AR** (1992) DNA binding site preferences and transcriptional activation properties of the Arabidopsis transcription factor GBF1. *The EMBO Journal* **11**: 1275-1289
- Schlotzhauer SD, Littell RC** (1997) SAS system for elementary statistical analysis. SAS institute Cary, North Carolina
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang XC, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA** (2010) Genome sequence of the palaeopolyploid soybean. *Nature* **463**: 178-183
- Schnepf RD, Dohlman E, Bolling C** (2001) Soybeans, agriculture, and policy in Brazil. *In* Agriculture in Brazil and Argentina: Developments and prospects for major field crops, Ed 1 Vol 1. Economic Research Service/USDA, pp 35-52
- Schweizer P, Felix G, Buchala A, Muller C, Metraux J-P** (1996) Perception of free cutin monomers by plant cells. *The Plant Journal* **10**: 331-341
- Sebastia CH, Hardin SC, Clouse SD, Kieber JJ, Huber SC** (2004) Identification of a new motif for CDPK phosphorylation *in vitro* that suggests ACC synthase may be a CDPK substrate. *Archives of Biochemistry and Biophysics* **428**: 81-91
- Sediyama T** (1985) Cultura da soja. UFV - Centro de Ciências Agrárias. Departamento de Fitotecnia
- Selkov E, Grechkin Y, Mikhailova N** (1998) MPW: The metabolic pathways database. *Nucleic Acids Research* **26**: 43-45
- Sessa G, Meller Y, Fluhr R** (1995) A GCC element and a G-box motif participate in ethylene-induced expression of the PRB-1b gene. *Plant Molecular Biology* **28**: 145-153
- Shaikhali J, Heiber I, Seidel T, Stroher E, Hiltcher H, Birkmann S, Dietz K-J, Baier M** (2008) The redox-sensitive transcription factor Rap2.4a controls nuclear expression of 2-Cys peroxiredoxin A and other chloroplast antioxidant enzymes. *BMC Plant Biology* **8**: 1-14
- Sharma HC, Dhillon MK, Arora R** (2008) Effects of *Bacillus thuringiensis*  $\delta$ -endotoxin-fed *Helicoverpa armigera* on the survival and development of the parasitoid *Campoletis chloridae*. *Entomologia Experimentalis et Applicata* **126**: 1-8
- Sharp RE** (2002) Interaction with ethylene: Changing views on the role of abscisic acid in root and shoot growth responses to water stress. *Plant, Cell & Environment* **25**: 211-222
- Shen R, Fan JB, Campbell D, Chang W, Chen J, Doucet D, Yeakley J, Bibikova M, Wickham Garcia E, McBride C, Steemers F, Garcia F, Kermani BG, Gunderson K, Oliphant A** (2005) High-throughput SNP genotyping on universal bead arrays. *Mutation Research* **573**: 70-82
- Shi H, Chen L, Ye T, Liu X, Ding K, Chan Z** (2014) Modulation of auxin content in Arabidopsis confers improved drought stress resistance. *Plant Physiology and Biochemistry* **82**: 209-217
- Shigyo M, Ito M** (2004) Analysis of gymnosperm two-AP2-domain-containing genes. *Development Genes and Evolution* **214**: 105-114

- Shinozaki K, Yamaguchi-Shinozaki K, Seki M** (2003) Regulatory network of gene expression in the drought and cold stress responses. *Current Opinion in Plant Biology* **6**: 410-417
- Shinwari ZK, Nakashima K, Miura S, Kasuga M, Seki M, Yamaguchi-Shinozaki K, Shinozaki K** (1998) An Arabidopsis gene family encoding DRE/CRT binding proteins involved in low-temperature-responsive gene expression. *Biochemical and Biophysical Research Communications* **250**: 161-170
- Shkolnik-Inbar D, Bar-Zvi D** (2010) ABI4 mediates abscisic acid and cytokinin inhibition of lateral root formation by reducing polar auxin transport in Arabidopsis. *The Plant Cell* **22**: 3560-3573
- Shoji T, Hashimoto T** (2011) Tobacco MYC2 regulates jasmonate-inducible nicotine biosynthesis genes directly and by way of the NIC2-locus ERF genes. *Plant and Cell Physiology* **52**: 1117-1130
- Silverstone AL, Chang C, Krol E, Sun TP** (1997) Developmental regulation of the gibberellin biosynthetic gene GA1 in *Arabidopsis thaliana*. *The Plant Journal* **12**: 9-19
- Silverstone AL, Tseng T-S, Swain SM, Dill A, Jeong SY, Olszewski NE, Sun T-p** (2007) Functional analysis of SPINDLY in gibberellin signaling in Arabidopsis. *Plant Physiology* **143**: 987-1000
- Singh SR, Rachie KO, Dashiell KE** (1987) Soybeans for the tropics: Research, production, and utilization. Wiley
- Skorupska H, Albertsen MC, Langholz KD, Palmer RG** (1989) Detection of ribosomal RNA genes in soybean, *Glycine max* (L.) Merr., by *in situ* hybridization. *Genome* **32**: 1091-1095
- Sloane RJ, Patterson RP, Carter TE** (1990) Field drought tolerance of a soybean plant introduction. *Crop Science* **30**: 118-123
- Solano R, Stepanova A, Chao Q, Ecker JR** (1998) Nuclear events in ethylene signaling: A transcriptional cascade mediated by ethylene-insensitive 3 and ethylene-response-factor 1. *Genes & Development* **12**: 3703-3714
- Somerville C, Briscoe J** (2001) Genetic engineering and water. *Science* **292**: 2217
- Spollen WG, LeNoble ME, Samuels TD, Bernstein N, Sharp RE** (2000) Abscisic acid accumulation maintains maize primary root elongation at low water potentials by restricting ethylene production. *Plant Physiology* **122**: 967-976
- Steed CL, Taylor LK, Harrison MA** (2004) Red light regulation of ethylene biosynthesis and gravitropism in etiolated pea stems. *Plant Growth Regulation* **43**: 117-125
- Stepanova AN, Alonso JM** (2009) Ethylene signaling and response: Where different regulatory modules meet. *Current Opinion in Plant Biology* **12**: 548-555
- Stepanova AN, Hoyt JM, Hamilton AA, Alonso JM** (2005) A link between ethylene and auxin uncovered by the characterization of two root-specific ethylene-insensitive mutants in Arabidopsis. *The Plant Cell* **17**: 2230-2242
- Stepanova AN, Robertson-Hoyt J, Yun J, Benavente LM, Xie DY, Dolezal K, Schlereth A, Jurgens G, Alonso JM** (2008) TAA1-mediated auxin biosynthesis is essential for hormone crosstalk and plant development. *Cell* **133**: 177-191
- Stockinger EJ, Gilmour SJ, Thomashow MF** (1997) *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a *cis*-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proceedings of the National Academy of Sciences* **94**: 1035-1040

- Stolf-Moreira R, Medri ME, Neumaier N, Lemos NG, Pimenta JA, Tobita S, Brogin RL, Marcelino-Guimaraes FC, Oliveira MC, Farias JR, Abdelnoor RV, Nepomuceno AL** (2010) Soybean physiology and gene expression during drought. *Genetics and Molecular Research* **9**: 1946-1956
- Stone BB, Cohen SP, Breton GL, Nietupski RM, Pelletier DA, Fiandaca MJ, Moe JG, Smith JH, Shah JS, Weisburg WG** (1996) Detection of rRNA from four respiratory pathogens using an automated Q $\beta$  replicase assay. *Molecular and Cellular Probes* **10**: 359-370
- Sun S, Yu J-P, Chen F, Zhao T-J, Fang X-H, Li Y-Q, Sui S-F** (2008) TINY, a dehydration-responsive element (DRE)-binding protein-like transcription factor connecting the DRE- and ethylene-responsive element-mediated signaling pathways in Arabidopsis. *Journal of Biological Chemistry* **283**: 6261-6271
- Sutoh K, Yamauchi D** (2003) Two *cis*-acting elements necessary and sufficient for gibberellin-upregulated proteinase expression in rice seeds. *The Plant Journal* **34**: 635-645
- Sutton DH, Conn GL, Brown T, Lane AN** (1997) The dependence of DNase I activity on the conformation of oligodeoxynucleotides. *Biochemical Journal* **321**: 481-486
- Swaminathan K, Peterson K, Jack T** (2008) The plant B3 superfamily. *Trends in Plant Science* **13**: 647-655
- Swarbreck D, Wilks C, Lamesch P, Berardini TZ, Garcia-Hernandez M, Foerster H, Li D, Meyer T, Muller R, Ploetz L, Radenbaugh A, Singh S, Swing V, Tissier C, Zhang P, Huala E** (2008) The Arabidopsis information resource (TAIR): Gene structure and function annotation. *Nucleic Acids Research* **36**: D1009-D1014
- Swarup R, Parry G, Graham N, Allen T, Bennett M** (2002) Auxin cross-talk: Integration of signalling pathways to control plant development. *Plant Molecular Biology* **49**: 409-424
- Swarup R, Perry P, Hagenbeek D, Van Der Straeten D, Beemster GT, Sandberg G, Bhalerao R, Ljung K, Bennett MJ** (2007) Ethylene upregulates auxin biosynthesis in Arabidopsis seedlings to enhance inhibition of root cell elongation. *The Plant Cell* **19**: 2186-2196
- Takei K, Yamaya T, Sakakibara H** (2004) Arabidopsis CYP735A1 and CYP735A2 encode cytokinin hydroxylases that catalyze the biosynthesis of *trans*-zeatin. *Journal of Biological Chemistry* **279**: 41866-41872
- Takusagawa F, Kamitori S, Misaki S, Markham GD** (1996) Crystal structure of S-adenosylmethionine synthetase. *Journal of Biological Chemistry* **271**: 136-147
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S** (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**: 2731-2739
- Tanaka Y, Sano T, Tamaoki M, Nakajima N, Kondo N, Hasezawa S** (2005) Ethylene inhibits abscisic acid-induced stomatal closure in Arabidopsis. *Plant Physiology* **138**: 2337-2343
- Tatsuki M, Mori H** (2001) Phosphorylation of tomato 1-aminocyclopropane-1-carboxylic acid synthase, Le-ACS2, at the C-terminal region. *Journal of Biological Chemistry* **276**: 28051-28057
- Texeira LR, Braccini AdLe, Sperandio D, Scapim CA, Schuster I, Vigano J** (2008) Avaliação de cultivares de soja quanto à tolerância ao estresse hídrico em substrato contendo polietileno glicol, Vol 30



- Thain SC, Vandenbussche F, Laarhoven LJJ, Dowson-Day MJ, Wang Z-Y, Tobin EM, Harren FJM, Millar AJ, Van Der Straeten D** (2004) Circadian rhythms of ethylene emission in *Arabidopsis*. *Plant Physiology* **136**: 3751-3761
- Thu NBA, Hoang XLT, Doan H, Nguyen T-H, Bui D, Thao NP, Tran L-SP** (2014) Differential expression analysis of a subset of GmNAC genes in shoots of two contrasting drought-responsive soybean cultivars DT51 and MTD720 under normal and drought conditions. *Molecular Biology Reports* **41**: 5563-5569
- Tiwari SB, Belachew A, Ma SF, Young M, Ade J, Shen Y, Marion CM, Holtan HE, Bailey A, Stone JK, Edwards L, Wallace AD, Canales RD, Adam L, Ratcliffe OJ, Repetti PP** (2012) The EDLL motif: A potent plant transcriptional activation domain from AP2/ERF transcription factors. *The Plant Journal* **70**: 855-865
- Toyofuku K, Umemura T-a, Yamaguchi J** (1998) Promoter elements required for sugar-repression of the RAmy3D gene for  $\alpha$ -amylase in rice. *FEBS Letters* **428**: 275-280
- Trusov Y, Botella JR** (2006) Silencing of the ACC synthase gene ACACS2 causes delayed flowering in pineapple [*Ananas comosus* (L.) Merr.]. *Journal of Experimental Botany* **57**: 3953-3960
- Tsuchisaka A, Yu G, Jin H, Alonso JM, Ecker JR, Zhang X, Gao S, Theologis A** (2009) A combinatorial interplay among the 1-aminocyclopropane-1-carboxylate isoforms regulates ethylene biosynthesis in *Arabidopsis thaliana*. *Genetics* **183**: 979-1003
- Tucker ML, Xue P, Yang R** (2010) 1-aminocyclopropane-1-carboxylic acid (ACC) concentration and ACC synthase expression in soybean roots, root tips, and soybean cyst nematode (*Heterodera glycines*)-infected roots. *Journal of Experimental Botany* **61**: 463-472
- Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM** (2007) Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research* **35**: W71-W74
- Valente MAS, Faria JAQA, Soares-Ramos JRL, Reis PAB, Pinheiro GL, Piovesan ND, Morais AT, Menezes CC, Cano MAO, Fietto LG, Loureiro ME, Aragao FJL, Fontes EPB** (2009) The ER luminal binding protein (BiP) mediates an increase in drought tolerance in soybean and delays drought-induced leaf senescence in soybean and tobacco. *Journal of Experimental Botany* **60**: 533-546
- Vandenbussche F, Vaseva I, Vissenberg K, Van Der Straeten D** (2012) Ethylene in vegetative development: A tale with a riddle. *New Phytologist* **194**: 895-909
- Vandenbussche F, Vriezen WH, Smalle J, Laarhoven LJJ, Harren FJM, Van Der Straeten D** (2003) Ethylene and auxin control the *Arabidopsis* response to decreased light intensity. *Plant Physiology* **133**: 517-527
- Vasudevan P, Briggs M** (2008) Biodiesel production - current state of the art and challenges. *Journal of Industrial Microbiology & Biotechnology* **35**: 421-430
- Verslues PE, Zhu JK** (2005) Before and beyond ABA: Upstream sensing and internal signals that determine ABA accumulation and response under abiotic stress. *Biochemical Society Transactions* **33**: 375-379
- Vlot AC, Dempsey DMA, Klessig DF** (2009) Salicylic acid, a multifaceted hormone to combat disease. *Annual Review of Phytopathology* **47**: 177-206
- Vriezen WH, Achard P, Harberd NP, Van Der Straeten D** (2004) Ethylene-mediated enhancement of apical hook formation in etiolated *Arabidopsis thaliana* seedlings is gibberellin dependent. *The Plant Journal* **37**: 505-516

- Wang H, Ding C, Du H, Liu H, Wang Y, Yu D** (2014) Identification of soybean MYC2-like transcription factors and overexpression of GmMYC1 could stimulate defense mechanism against common cutworm in transgenic tobacco. *Biotechnology Letters* **36**: 1881-1892
- Wang KLC, Yoshida H, Lurin C, Ecker JR** (2004) Regulation of ethylene gas biosynthesis by the Arabidopsis ETO1 protein. *Nature* **428**: 945-950
- Wang W, Hall AE, O'Malley R, Bleecker AB** (2003) Canonical histidine kinase activity of the transmitter domain of the ETR1 ethylene receptor from Arabidopsis is not required for signal transmission. *Proceedings of the National Academy of Sciences* **100**: 352-357
- Wei W, Huang J, Hao Y-J, Zou H-F, Wang H-W, Zhao J-Y, Liu X-Y, Zhang W-K, Ma B, Zhang J-S, Chen S-Y** (2009) Soybean GmPHD-type transcription regulators improve stress tolerance in transgenic Arabidopsis plants. *PLoS One* **4**: e7209
- Weiss D, Ori N** (2007) Mechanisms of cross talk between gibberellin and other hormones. *Plant Physiology* **144**: 1240-1246
- Welsch R, Maass D, Voegel T, DellaPenna D, Beyer P** (2007) Transcription factor RAP2.2 and its interacting partner SINAT2: Stable elements in the carotenogenesis of Arabidopsis leaves. *Plant Physiology* **145**: 1073-1085
- Welsch R, Medina J, Giuliano G, Beyer P, Von Lintig J** (2003) Structural and functional characterization of the phytoene synthase promoter from *Arabidopsis thaliana*. *Planta* **216**: 523-534
- Wen X, Zhang C, Ji Y, Zhao Q, He W, An F, Jiang L, Guo H** (2012) Activation of ethylene signaling is mediated by nuclear translocation of the cleaved EIN2 carboxyl terminus. *Cell Research* **22**: 1613-1616
- Went FW** (1927) Wuchsstoff und wachstum, "proefschrift... door frits warmolt went". J. H. de Bussy, Amsterdam
- Went FW, Thimann KV** (1937) Phytohormones. The Macmillan Company, New York
- Wery J, Silim SN, Knights EJ, Malhotra RS, Cousin R** (1993) Screening techniques and sources of tolerance to extremes of moisture and air temperature in cool season food legumes. *Euphytica* **73**: 73-83
- Willige BC, Isono E, Richter R, Zourelidou M, Schwechheimer C** (2011) Gibberellin regulates Pin-formed abundance and is required for auxin transport-dependent growth and development in *Arabidopsis thaliana*. *The Plant Cell* **23**: 2184-2195
- Wilmowicz E, Keszy J, Kopcewicz J** (2008) Ethylene and ABA interactions in the regulation of flower induction in *Pharbitis nil*. *Journal of Plant Physiology* **165**: 1917-1928
- Winicov I, Bastola DR** (1999) Transgenic overexpression of the transcription factor Alfin1 enhances expression of the endogenous MsPRP2 gene in alfalfa and improves salinity tolerance of the plants. *Plant Physiology* **120**: 473-480
- Wittkopp PJ, Kalay G** (2012) *cis*-regulatory elements: Molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews Genetics* **13**: 59-69
- Wolbang CM, Ross JJ** (2001) Auxin promotes gibberellin biosynthesis in decapitated tobacco plants. *Planta* **214**: 153-157
- Worrall D, Holroyd GH, Moore JP, Glowacz M, Croft P, Taylor JE, Paul ND, Roberts MR** (2012) Treating seeds with activators of plant defence generates long-lasting priming of resistance to pests and pathogens. *New Phytologist* **193**: 770-778

- Wu L, Zhang Z, Zhang H, Wang X-C, Huang R** (2008) Transcriptional modulation of ethylene response factor protein JERF3 in the oxidative stress response enhances tolerance of tobacco seedlings to salt, drought, and freezing. *Plant Physiology* **148**: 1953-1963
- Wu Y, Cosgrove DJ** (2000) Adaptation of roots to low water potentials by changes in cell wall extensibility and cell wall proteins. *Journal of Experimental Botany* **51**: 1543-1553
- Xie F, Liu Q, Wen C-K** (2006) Receptor signal output mediated by the ETR1 N-terminus is primarily subfamily I receptor dependent. *Plant Physiology* **142**: 492-508
- Xie Z, Zhang Z-L, Zou X, Yang G, Komatsu S, Shen QJ** (2006) Interactions of two abscisic-acid induced WRKY genes in repressing gibberellin signaling in aleurone cells. *The Plant Journal* **46**: 231-242
- Xu J, Li Y, Wang Y, Liu H, Lei L, Yang H, Liu G, Ren D** (2008) Activation of MAPK kinase 9 induces ethylene and camalexin biosynthesis and enhances sensitivity to salt stress in *Arabidopsis*. *Journal of Biological Chemistry* **283**: 26996-27006
- Xu Z-S, Xia L-Q, Chen M, Cheng X-G, Zhang R-Y, Li L-C, Zhao Y-X, Lu Y, Ni Z-Y, Liu L, Qiu Z-G, Ma Y-Z** (2007) Isolation and molecular characterization of the *Triticum aestivum* L. ethylene-responsive factor 1 (TaERF1) that increases multiple stress tolerance. *Plant Molecular Biology* **65**: 719-732
- Yamagami T, Tsuchisaka A, Yamada K, Haddon WF, Harden LA, Theologis A** (2003) Biochemical diversity among the 1-amino-cyclopropane-1-carboxylate synthase isozymes encoded by the *Arabidopsis* gene family. *Journal of Biological Chemistry* **278**: 49102-49112
- Yamaguchi-Shinozaki K, Shinozaki K** (1994) A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *The Plant Cell* **6**: 251-264
- Yamaguchi-Shinozaki K, Shinozaki K** (2005) Organization of *cis*-acting regulatory elements in osmotic- and cold-stress-responsive promoters. *Trends in Plant Science* **10**: 88-94
- Yamasaki K, Kigawa T, Inoue M, Tateno M, Yamasaki T, Yabuki T, Aoki M, Seki E, Matsuda T, Tomo Y, Hayami N, Terada T, Shirouzu M, Osanai T, Tanaka A, Seki M, Shinozaki K, Yokoyama S** (2004) Solution structure of the B3 DNA binding domain of the *Arabidopsis* cold-responsive transcription factor RAV1. *The Plant Cell* **16**: 3448-3459
- Yamasaki K, Kigawa T, Inoue M, Yamasaki T, Yabuki T, Aoki M, Seki E, Matsuda T, Tomo Y, Terada T, Shirouzu M, Tanaka A, Seki M, Shinozaki K, Yokoyama S** (2005) Solution structure of the major DNA-binding domain of *Arabidopsis thaliana* ethylene-insensitive 3-like 3. *Journal of Molecular Biology* **348**: 253-264
- Yang SF, Hoffman NE** (1984) Ethylene biosynthesis and its regulation in higher plants. *Annual Review of Plant Physiology* **35**: 155-189
- Yip WK, Moore T, Yang SF** (1992) Differential accumulation of transcripts for four tomato 1-aminocyclopropane-1-carboxylate synthase homologs under various conditions. *Proceedings of the National Academy of Sciences* **89**: 2475-2479
- Yoo S-D, Cho Y-H, Tena G, Xiong Y, Sheen J** (2008) Dual control of nuclear EIN3 by bifurcate MAPK cascades in C<sub>2</sub>H<sub>4</sub> signalling. *Nature* **451**: 789-795

- Zhai Y, Wang Y, Li Y, Lei T, Yan F, Su L, Li X, Zhao Y, Sun X, Li J, Wang Q** (2013) Isolation and molecular characterization of GmERF7, a soybean ethylene-response factor that increases salt stress tolerance in tobacco. *Gene* **513**: 174-183
- Zhang G, Chen M, Chen X, Xu Z, Guan S, Li L-C, Li A, Guo J, Mao L, Ma Y** (2008) Phylogeny, gene structures, and expression patterns of the ERF gene family in soybean (*Glycine max* L.). *Journal of Experimental Botany* **59**: 4095-4107
- Zhang G, Chen M, Li L, Xu Z, Chen X, Guo J, Ma Y** (2009) Overexpression of the soybean GmERF3 gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. *Journal of Experimental Botany* **60**: 3781-3796
- Zhang H, Yang Y, Zhang Z, Chen J, Wang X-C, Huang R** (2008) Expression of the ethylene response factor gene TSRF1 enhances abscisic acid responses during seedling development in tobacco. *Planta* **228**: 777-787
- Zhang M, Yuan B, Leng P** (2009) The role of ABA in triggering ethylene biosynthesis and ripening of tomato fruit. *Journal of Experimental Botany* **60**: 1579-1588
- Zhang S-W, Li C-H, Cao J, Zhang Y-C, Zhang S-Q, Xia Y-F, Sun D-Y, Sun Y** (2009) Altered architecture and enhanced drought tolerance in rice via the down-regulation of indole-3-acetic acid by TLD1/OsGH3.13 activation. *Plant Physiology* **151**: 1889-1901
- Zhao S, Fernald RD** (2005) Comprehensive algorithm for quantitative real-time polymerase chain reaction. *Journal of Computational Biology* **12**: 1047-1064
- Zhong R, Lee C, Ye Z-H** (2010) Global analysis of direct targets of secondary wall NAC master switches in Arabidopsis. *Molecular Plant* **3**: 1087-1103
- Zhu J-K** (2002) Salt and drought stress signal transduction in plants. *Annual Review of Plant Biology* **53**: 247-273
- Zhu Y, Cai X-L, Wang Z-Y, Hong M-M** (2003) An interaction between a MYC protein and an EREBP protein is involved in transcriptional regulation of the rice WX gene. *Journal of Biological Chemistry* **278**: 47803-47811
- Zhu Z** (2014) Molecular basis for jasmonate and ethylene signal interactions in Arabidopsis. *Journal of Experimental Botany*: 1-6
- Zonta EP** (1984) SANEST: Sistema de análise estatística relação dos módulos disponíveis,

# APÊNDICES

**Tabela S1.** Lista de Proteínas Relacionadas com a Biossíntese de Etileno em *Arabidopsis thaliana*

Proteína de <i>Arabidopsis thaliana</i>			Domínio Proteico			Via Metabólica
Nome Genérico	Locus ID	R.A. <sup>1</sup>	Código PFAM	Descrição	Posição	
AtAAT	At1g77670	440	PF00155	Aminotransferase classe I e II	079 - 432	Ciclo de Yang
AtACD1	At1g48420	401	PF00291	Enzima dependente de piridoxal fosfato	054 - 363	Degradação de ACC
AtACO1	At2g19590	310	PF03171	Superfamília de oxigenases 2OG-Fe(II)	158 - 259	Biossíntese de Etileno
AtACO2	At1g62380	320	PF03171	Superfamília de oxigenases 2OG-Fe(II)	156 - 256	Biossíntese de Etileno
AtACO4	At1g05010	323	PF03171	Superfamília de oxigenases 2OG-Fe(II)	153 - 254	Biossíntese de Etileno
AtACO#001	At1g12010	320	PF03171	Superfamília de oxigenases 2OG-Fe(II)	156 - 256	Biossíntese de Etileno
AtACO#002	At1g77330	307	PF03171	Superfamília de oxigenases 2OG-Fe(II)	156 - 256	Biossíntese de Etileno
AtACS1	At3g61510	488	PF00155	Aminotransferase classe I e II	047 - 425	Biossíntese de Etileno
AtACS2	At1g01480	496	PF00155	Aminotransferase classe I e II	048 - 432	Biossíntese de Etileno
AtACS4	At2g22810	474	PF00155	Aminotransferase classe I e II	040 - 425	Biossíntese de Etileno
AtACS5	At5g65800	470	PF00155	Aminotransferase classe I e II	040 - 424	Biossíntese de Etileno
AtACS6	At4g11280	495	PF00155	Aminotransferase classe I e II	051 - 432	Biossíntese de Etileno
AtACS7	At4g26200	447	PF00155	Aminotransferase classe I e II	054 - 437	Biossíntese de Etileno
AtACS8	At4g37770	469	PF00155	Aminotransferase classe I e II	040 - 424	Biossíntese de Etileno
AtACS9	At3g49700	470	PF00155	Aminotransferase classe I e II	040 - 424	Biossíntese de Etileno
AtACS10	At1g62960	557	PF00155	Aminotransferase classe I e II	169 - 546	Biossíntese de Etileno <sup>2</sup>
AtACS11	At4g08040	460	PF00155	Aminotransferase classe I e II	038 - 419	Biossíntese de Etileno
AtACS12	At5g51690	495	PF00155	Aminotransferase classe I e II	105 - 486	Biossíntese de Etileno <sup>2</sup>
AtACT#001	At1g03940	469	PF02458	Família de transferases	003 - 463	Síntese de MACC <sup>3</sup>
AtACT#002	At1g03495	465	PF02458	Família de transferases	006 - 459	Síntese de MACC <sup>3</sup>
AtACT#003	At2g30200	393	PF02458	Família de transferases	091 - 390	Síntese de MACC <sup>3</sup>
AtACT#004	At3g29670	451	PF02458	Família de transferases	001 - 448	Síntese de MACC <sup>3</sup>
AtACT#005	At5g39050	469	PF02458	Família de transferases	008 - 466	Síntese de MACC <sup>3</sup>
AtACT#006	At5g39090	448	PF02458	Família de transferases	003 - 443	Síntese de MACC <sup>3</sup>
AtARD1	At4g14716	192	PF03079	Família ARD/ARD'	014 - 168	Ciclo de Yang
			PF07883	Domínio cupin	086 - 162	
AtARD2	At4g14710	199	PF03079	Família ARD/ARD'	014 - 168	Ciclo de Yang
			PF07883	Domínio cupin	086 - 162	
AtARD3	At2g26400	199	PF03079	Família ARD/ARD'	014 - 168	Ciclo de Yang
			PF07883	Domínio cupin	086 - 162	

**Tabela S1.** Lista de Proteínas Relacionadas com a Biossíntese de Etileno em *Arabidopsis thaliana* (cont.)

Proteína de <i>Arabidopsis thaliana</i>			Domínio Proteico			Via Metabólica
Nome Genérico	Locus ID	R.A. <sup>1</sup>	Código PFAM	Descrição	Posição	
AtARD4	At5g43850	187	PF03079	Família ARD/ARD'	004 - 158	Ciclo de Yang
			PF07883	Domínio cupin	078 - 152	
AtASP2	At5g19550	405	PF00155	Aminotransferase classe I e II	029 - 397	Ciclo de Yang
ATASP3	At5g11520	449	PF00155	Aminotransferase classe I e II	073 - 441	Ciclo de Yang
AtASP5	At4g31990	453	PF00155	Aminotransferase classe I e II	077 - 445	Ciclo de Yang
AtDEP1	At5g53850	507	PF00596	Domínio N-terminal de aldolases e aducinas classe II	026 - 229	Ciclo de Yang
			PF00702	Hidrolase semelhante à dehalogenase haloácida	265 - 478	
AtGGT1	At4g39640	572	PF01019	Gama-glutamiltranspeptidase	048 - 567	Síntese de GACC <sup>3</sup>
AtGGT2	At4g39650	578	PF01019	Gama-glutamiltranspeptidase	052 - 573	Síntese de GACC <sup>3</sup>
AtGGT3	At1g69820	191	PF01019	Gama-glutamiltranspeptidase	001 - 191	Síntese de GACC <sup>3</sup>
AtGGT4	At4g29210	637	PF01019	Gama-glutamiltranspeptidase	095 - 632	Síntese de GACC <sup>3</sup>
AtMAT1	At1g02500	393	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	002 - 101	Biossíntese de Etileno
			PF02772	S-adenosil-metionina sintetase, domínio central	116 - 238	
			PF02773	S-adenosil-metionina sintetase, domínio C-terminal	240 - 382	
AtMAT2	At4g01850	393	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	002 - 101	Biossíntese de Etileno
			PF02772	S-adenosil-metionina sintetase, domínio central	116 - 238	
			PF02773	S-adenosil-metionina sintetase, domínio C-terminal	240 - 382	
AtMAT3	At2g36880	390	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	002 - 101	Biossíntese de Etileno
			PF02772	S-adenosil-metionina sintetase, domínio central	116 - 238	
			PF02773	S-adenosil-metionina sintetase, domínio C-terminal	240 - 382	
AtMAT4	At3g17390	393	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	002 - 101	Biossíntese de Etileno
			PF02772	S-adenosil-metionina sintetase, domínio central	116 - 238	
			PF02773	S-adenosil-metionina sintetase, domínio C-terminal	240 - 382	
AtMTI1	At2g05830	374	PF01008	Família da subunidade do fator de iniciação 2	050 - 353	Ciclo de Yang
AtMTK1	At1g49820	420	PF01636	Família da enzima fosfotransferase	038 - 295	Ciclo de Yang
AtMTN1	At4g38800	267	PF01048	Superfamília das fosforilases	028 - 257	Ciclo de Yang
AtMTN2	At4g34840	254	PF01048	Superfamília das fosforilases	015 - 244	Ciclo de Yang

<sup>1</sup> Tamanho em resíduos de aminoácidos.<sup>2</sup> Possível aminoácido transferase (AAT). Não possui atividade de ACC sintase comprovada.<sup>3</sup> A função primária desta proteína não está inicialmente associada a esta rota metabólica.

**Tabela S2.** Lista de Proteínas Relacionadas com a Transdução de Sinal Mediada por Etileno em *Arabidopsis thaliana*

Proteína de <i>Arabidopsis thaliana</i>			Domínio Proteico			Função
Nome Genérico	Locus ID	R.A. <sup>1</sup>	Código PFAM	Descrição	Posição	
AtCDPK2	At1g35670	495	PF00069	Domínio proteína cinase	028 - 284	Ativação de ACS Tipo II
			PF00036	“EF hand”	331 - 392 398 - 462	
AtCTR1	At5g03730	821	PF14381	Proteína cinase responsiva à etileno Le-CTR1	205 - 414	
			PF07714	Proteína tirosina cinase	551 - 805	
AtEBF1	At2g25490	628	PF00646	Domínio F-box	064 - 104	Direciona EIN3 para Proteólise
AtEBF2	At5g25350	625	PF00646	Domínio F-box	054 - 095	Direciona EIN3 para Proteólise
AtEIL1	At2g27050	584	PF04873	Insensível à etileno 3	049 - 431	Fator de Transcrição Nuclear
AtEIL2	At5g21120	518	PF04873	Insensível à etileno 3	048 - 408	Fator de Transcrição Nuclear
AtEIN2	At5g03280	1294	PF01566	Proteína associada à resistência natural em macrófagos	038 - 390	Regulação após CTR1
AtEIN3	At3g20770	628	PF04873	Insensível à etileno 3	049 - 422	Fator de Transcrição Nuclear
AtEIN5	At1g54490	947	PF03159	N-terminal da exonuclease 5’-3’ XRN	001 - 256	Silenciamento Gênico de EBF
			PF00098	Articulação de zinco	263 - 280	
AtEOL1	At4g02680	888	PF00515	Repetição tetratricopeptídeo	519 - 544	Regulador Negativo de ACS Tipo II
AtEOL2	At5g58550	925	PF00515	Repetição tetratricopeptídeo	738 - 771	Regulador Negativo de ACS Tipo II
AtETO1	At3g51770	959	PF13414	Repetição TPR	875 - 940	Regulador Negativo de ACS Tipo II
AtEIN4	At3g04580	766	PF01590	Domínio GAF	181 - 331	Receptor de Etileno
			PF00512	Domínio histidina cinase A (fosfoceptor)	367 - 432	
			PF00072	Domínio regulador de resposta de receptor	644 - 758	
AtERS1	At2g40940	613	PF01590	Domínio GAF	158 - 307	Receptor de Etileno
			PF00512	Domínio histidina cinase A (fosfoceptor)	343 - 408	
			PF02518	Histidina cinase, DNA girase B e semelhante à ATPase	455 - 586	
AtERS2	At1g04310	645	PF01590	Domínio GAF	190 - 346	Receptor de Etileno
AtETP1	At3g18980	415	PF00646	Domínio F-box	001 - 043	Direciona EIN2 para Proteólise
			PF07734	Associado a F-box	214 - 398	
AtETP2	At3g18910	388	PF00646	Domínio F-box	007 - 044	Direciona EIN2 para Proteólise
			PF07734	Associado a F-box	206 - 373	
AtETR1	At1g66340	738	PF01590	Domínio GAF	158 - 307	Receptor de Etileno
			PF00512	Domínio histidina cinase A (fosfoceptor)	343 - 408	
			PF02518	Histidina cinase, DNA girase B e semelhante à ATPase	455 - 582	
			PF00072	Domínio regulador de resposta de receptor	612 - 726	



**Tabela S2.** Lista de Proteínas Relacionadas com a Transdução de Sinal Mediada por Etileno em *Arabidopsis thaliana* (cont.)

Proteína de <i>Arabidopsis thaliana</i>			Domínio Proteico			Função
Nome Genérico	Locus ID	R.A. <sup>1</sup>	Código PFAM	Descrição	Posição	
AtETR2	At3g23150	773	PF01590	Domínio GAF	187 - 331	Receptor de Etileno
			PF00512	Domínio histidina cinase A (fosfoceptor)	367 - 427	
			PF00072	Domínio regulador de resposta de receptor	648 - 763	
AtMKK4	At1g51660	366	PF00069	Domínio proteína cinase	079 - 334	Ativação de ACS Tipo I
AtMKK5	At3g21220	348	PF00069	Domínio proteína cinase	070 - 325	Ativação de ACS Tipo I
AtMKK9	At1g73500	310	PF00069	Domínio proteína cinase	047 - 306	Ativação de EIN3
AtMPK3	At3g45640	370	PF00069	Domínio proteína cinase	041 - 324	Ativação de EIN3
AtMPK6	At2g43790	395	PF00069	Domínio proteína cinase	065 - 348	Ativação de EIN3 e ACS Tipo I
AtRAN1	At5g20010	221	PF00071	Família ras	015 - 173	Transporte de Cobre
AtRAN2	At5g20020	221	PF00071	Família ras	015 - 173	Transporte de Cobre
AtRAN3	At5g55190	221	PF00071	Família ras	015 - 173	Transporte de Cobre
AtRAN4	At5g55080	222	PF00071	Família ras	015 - 173	Transporte de Cobre
AtRTE1	At2g26070	250	PF05608	Proteína sem função conhecida (DUF778)	060 - 195	Regulador Positivo do Receptor de Etileno
AtRTH1	At3g51040	231	PF05608	Proteína sem função conhecida (DUF778)	038 - 175	Regulador Positivo do Receptor de Etileno

<sup>1</sup> Tamanho em resíduos de aminoácidos.

**Tabela S3.** Lista de Proteínas AP2 de *Arabidopsis thaliana*<sup>1</sup>

Proteína de <i>Arabidopsis thaliana</i>			Domínio AP2		Função e/ou Nome(s) Completo(s) da Proteína
Nome Genérico	Locus ID	R.A. <sup>2</sup>	Nº <sup>3</sup>	Posição	
AtAP2#001	At1g16060	345	2	057 - 117 159 - 211	Proteína com Duplo Domínio AP2 que Interage com ARIA (ADAP); “Wrinkled” 3 (WRI3)
AtAP2#002	At1g51190	568	2	189 - 249 291 - 343	“Plethora” 2 (PLT2)
AtAP2#003	At1g72570	415	2	222 - 282 324 - 376	Proteína da Superfamília de um Tipo de Integrase que se Liga ao DNA
AtAP2#004	At1g79700	303	2	051 - 111 153 - 195	“Wrinkled” 4 (WRI4)
AtAP2#005	At2g28550	449	1	152 - 202	Alvo Marcado de Início de Ativação 1 (TOE1); RAP2.7
AtAP2#006	At2g39250	325	1	107 - 157	“Schnarchzapfen” (SNZ)
AtAP2#007	At2g41710	423	1	069 - 129	Proteína da Superfamília de um Tipo de Integrase que se Liga ao DNA
AtAP2#008	At3g20840	574	2	180 - 240 282 - 334	“Plethora” 1 (PLT1)
AtAP2#009	At3g54320	438	2	064 - 124 166 - 218	“Wrinkled” 1 (WRI1)
AtAP2#010	At3g54990	346	1	121 - 171	“Schlafmutze” (SMZ)
AtAP2#011	At4g36920	432	2	130 - 180 222 - 273	Mutante Floral 2 (FLO2; AP2; FL1; ATAP2)
AtAP2#012	At4g37750	555	2	282 - 342 384 - 436	Complementa Mutante da Proteína Cinase C 1 (CKC1); “Aintegumenta” (ANT)
AtAP2#013	At5g10510	569	2	252 - 312 354 - 406	Semelhante à “Aintegumenta” 6 (AIL6); “Plethora” 3 (PLT3)
AtAP2#014	At5g17430	584	2	209 - 269 311 - 363	“Baby Boom” (BBM)
AtAP2#015	At5g57390	558	2	202 - 262 304 - 356	Semelhante à “Aintegumenta” 6 (AIL5); “Plethora” 5 (PLT5); “Chotto” 1 (CHO1)
AtAP2#016	At5g60120	485	1	159 - 209	Alvo Marcado de Início de Ativação 2 (TOE2)

**Tabela S3.** Lista de Proteínas AP2 de *Arabidopsis thaliana* (cont.)<sup>1</sup>

Proteína de <i>Arabidopsis thaliana</i>			Domínio AP2		Função e/ou Nome(s) Completo(s) da Proteína
Nome Genérico	Locus ID	R.A. <sup>2</sup>	Nº <sup>3</sup>	Posição	
AtAP2#017	At5g65510	498	2	230 - 290 332 - 384	Semelhante à “Aintegumenta” 6 (AIL7); “Plethora” 7 (PLT7)
AtAP2#018	At5g67180	352	2	095 - 145 187 - 238	Alvo Marcado de Início de Ativação 3 (TOE3)
AtAP2#019	At4g13040	246	1	137 - 179	Proteína da Superfamília de um Tipo de Integrase que Interage com DNA – Proteína Soloísta <sup>4</sup>

<sup>1</sup> Lista de proteínas AP2 de *Arabidopsis thaliana* obtida no trabalho de Nakano et al. (2006).

<sup>2</sup> Tamanho em resíduos de aminoácidos.

<sup>3</sup> Número de domínios AP2 (PF00847) presentes na sequência de aminoácidos em questão.

<sup>4</sup> Esta proteína diverge filogeneticamente das outras proteínas AP2 de *Arabidopsis thaliana*.

**Tabela S4.** Lista dos Fatores Transcricionais Responsivos à Etileno (ERF) de *Arabidopsis thaliana*<sup>1</sup>

Nome Genérico	Locus ID	C-I <sup>2</sup>	C-II <sup>3</sup>	Nome Genérico	Locus ID	C-I <sup>2</sup>	C-II <sup>3</sup>
AtERF#001	At1g15360	B6	V	AtERF#062	At4g13620	A6	I
AtERF#002	At5g19790	B6	V	AtERF#063	At4g11140	B5	VI
AtERF#003	At5g25190	B6	V	AtERF#064	At4g23750	B5	VI
AtERF#004	At5g11190	B6	V	AtERF#065	At5g53290	B5	VI
AtERF#005	At5g25390	B6	V	AtERF#066	At4g27950	B5	VI
AtERF#006	At1g46768	A5	II	AtERF#067	At3g61630	B5	VI
AtERF#007	At4g06746	A5	II	AtERF#068	At2g46310	B5	VI
AtERF#008	At2g23340	A5	II	AtERF#069	At1g22985	B5	VI
AtERF#009	At4g36900	A5	II	AtERF#070	At1g71130	B5	VI
AtERF#010	At5g67190	A5	II	AtERF#071	At2g47520	B2	VII
AtERF#011	At3g50260	A5	II	AtERF#072	At3g16770	B2	VII
AtERF#012	At1g21910	A5	II	AtERF#073	At1g72360	B2	VII
AtERF#013	At1g77640	A5	II	AtERF#074	At1g53910	B2	VII
AtERF#014	At1g44830	A5	II	AtERF#075	At3g14230	B2	VII
AtERF#015	At4g31060	A5	II	AtERF#076	At1g28370	B1	VIII
AtERF#016	At5g21960	A5	II	AtERF#077	At1g03800	B1	VIII
AtERF#017	At1g19210	A5	II	AtERF#078	At3g15210	B1	VIII
AtERF#018	At1g74930	A5	II	AtERF#079	At1g53170	B1	VIII
AtERF#019	At1g22810	A5	II	AtERF#080	At5g44210	B1	VIII
AtERF#020	At1g71520	A5	II	AtERF#081	At1g28360	B1	VIII
AtERF#021	At1g71450	A4	III	AtERF#082	At1g50640	B1	VIII
AtERF#022	At1g33760	A4	III	AtERF#083	At3g20310	B1	VIII
AtERF#023	At1g01250	A4	III	AtERF#084	At1g80580	B1	VIII
AtERF#024	At2g36450	A4	III	AtERF#085	At5g13910	B1	VIII
AtERF#025	At5g52020	A4	III	AtERF#086	At5g18560	B1	VIII
AtERF#026	At1g63040	A4	III	AtERF#087	At1g28160	B1	VIII
AtERF#027	At1g12630	A4	III	AtERF#088	At1g12890	B1	VIII
AtERF#028	At5g51990	A1	III	AtERF#089	At1g12980	B1	VIII
AtERF#029	At4g25490	A1	III	AtERF#090	At1g24590	B1	VIII
AtERF#030	At4g25470	A1	III	AtERF#091	At4g18450	B3	IX
AtERF#031	At4g25480	A1	III	AtERF#092	At3g23240	B3	IX
AtERF#032	At1g63030	A1	III	AtERF#093	At2g31230	B3	IX
AtERF#033	At1g12610	A1	III	AtERF#094	At1g06160	B3	IX
AtERF#034	At2g44940	A4	III	AtERF#095	At3g23220	B3	IX
AtERF#035	At3g60490	A4	III	AtERF#096	At5g43410	B3	IX
AtERF#036	At3g16280	A4	III	AtERF#097	At1g04370	B3	IX
AtERF#037	At1g77200	A4	III	AtERF#098	At3g23230	B3	IX
AtERF#038	At2g35700	A4	III	AtERF#099	At2g44840	B3	IX
AtERF#039	At4g16750	A4	III	AtERF#100	At4g17500	B3	IX
AtERF#040	At5g25810	A4	III	AtERF#101	At5g47220	B3	IX
AtERF#041	At5g11590	A4	III	AtERF#102	At5g47230	B3	IX
AtERF#042	At2g25820	A4	III	AtERF#103	At4g17490	B3	IX
AtERF#043	At4g32800	A4	III	AtERF#104	At5g61600	B3	IX
AtERF#044	At3g11020	A2	IV	AtERF#105	At5g51190	B3	IX
AtERF#045	At5g05410	A2	IV	AtERF#106	At5g07580	B3	IX
AtERF#046	At2g38340	A2	IV	AtERF#107	At5g61590	B3	IX
AtERF#047	At2g40350	A2	IV	AtERF#108	At1g43160	B4	X
AtERF#048	At2g40340	A2	IV	AtERF#109	At4g34410	B4	X
AtERF#049	At1g75490	A2	IV	AtERF#110	At5g50080	B4	X
AtERF#050	At5g18450	A2	IV	AtERF#111	At5g64750	B4	X
AtERF#051	At3g57600	A2	IV	AtERF#112	At2g33710	B4	X
AtERF#052	At2g40220	A3	IV	AtERF#113	At5g13330	B4	X
AtERF#053	At2g20880	A6	I	AtERF#114	At5g61890	B4	X
AtERF#054	At4g28140	A6	I	AtERF#115	At5g07310	B4	X
AtERF#055	At1g36060	A6	I	AtERF#116	At1g25470	B6	VI-L
AtERF#056	At2g22200	A6	I	AtERF#117	At1g49120	B6	VI-L
AtERF#057	At5g65130	A6	I	AtERF#118	At1g68550	B6	VI-L
AtERF#058	At1g22190	A6	I	AtERF#119	At3g25890	B6	VI-L
AtERF#059	At1g78080	A6	I	AtERF#120	At2g20350	B6	Xb-L
AtERF#060	At4g39780	A6	I	AtERF#121	At5g67010	B6	Xb-L
AtERF#061	At1g64380	A6	I	AtERF#122	At5g67000	B6	Xb-L

<sup>1</sup> Lista de ERFs de *Arabidopsis thaliana* obtida no trabalho de Nakano et al. (2006).<sup>2</sup> Classificação realizada por Sakuma et al. (2002).<sup>3</sup> Classificação realizada por Nakano et al. (2006).

**Tabela S5.** Lista de Proteínas RAV de *Arabidopsis thaliana*<sup>1</sup>

Proteína de <i>Arabidopsis thaliana</i>			Domínios AP2/B3		Função e/ou Nome(s) Completo(s) da Proteína
Nome Genérico	Locus ID	R.A. <sup>2</sup>	Código PFAM <sup>3</sup>	Posição	
AtRAV1	At1g13260	344	PF00847	060 - 107	Relacionado à ABI3/VP1 1 (RAV1); Fator de Interação com DNA em Resposta à Etileno 4 (EDF4)
			PF02362	188 - 291	
AtRAV2	At1g68840	352	PF00847	063 - 111	Relacionado à ABI3/VP1 2 (RAV2); “Tempranillo” 2 (TEM2)
			PF02362	188 - 288	
AtRAV#001	At1g25560	361	PF00847	070 - 118	Fator de Interação com DNA em Resposta à Etileno 1 (EDF1); “Tempranillo” 1 (TEM1)
			PF02362	195 - 302	
AtRAV#002	At1g50680	337	PF00847	027 - 073	Proteína Relacionada à ABI3/VP1 (RAV)
			PF02362	157 - 256	
AtRAV#003	At1g51120	352	PF00847	045 - 092	Proteína Relacionada à ABI3/VP1 (RAV)
			PF02362	178 - 274	
AtRAV#004	At3g25730	333	PF00847	060 - 108	Fator de Interação com DNA em Resposta à Etileno 3 (EDF3)
			PF02362	183 - 288	

<sup>1</sup> Lista de proteínas RAV de *Arabidopsis thaliana* obtida no trabalho de Nakano et al. (2006).

<sup>2</sup> Tamanho em resíduos de aminoácidos.

<sup>3</sup> PF00847: Domínio AP2; PF02362: Domínio B3 de interação com DNA.

**Tabela S6.** Lista de Proteínas Relacionadas com a Biossíntese de Etileno em *Oryza sativa*

Proteína de <i>Oryza sativa</i>			Domínio Proteico			Via Metabólica <sup>2</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	Código PFAM	Descrição	Posição	
OsAAT#001	Os09g28050	466	PF00155	Aminotransferase classe I e II	109 - 458	Ciclo de Yang
OsAAT#002	Os10g34350	412	PF00155	Aminotransferase classe I e II	030 - 381	Ciclo de Yang
OsACD#001	Os02g53330	565	PF00291	Enzima dependente de piridoxal fosfato	082 - 388	Degradação de ACC
OsACO#001	Os01g39860	312	PF03171	Superfamília de oxigenases 2OG-Fe(II)	169 - 264	Biossíntese de Etileno
OsACO#002	Os02g53180	344	PF03171	Superfamília de oxigenases 2OG-Fe(II)	184 - 280	Biossíntese de Etileno
OsACO#003	Os05g05680	308	PF03171	Superfamília de oxigenases 2OG-Fe(II)	163 - 259	Biossíntese de Etileno
OsACO#004	Os06g37590	293	PF03171	Superfamília de oxigenases 2OG-Fe(II)	150 - 246	Biossíntese de Etileno
OsACO#005	Os09g27750	322	PF03171	Superfamília de oxigenases 2OG-Fe(II)	164 - 259	Biossíntese de Etileno
OsACO#006	Os09g27820	322	PF03171	Superfamília de oxigenases 2OG-Fe(II)	164 - 259	Biossíntese de Etileno
OsACO#007	Os11g08380	309	PF03171	Superfamília de oxigenases 2OG-Fe(II)	163 - 262	Biossíntese de Etileno
OsACS#001	Os01g09700	510	PF00155	Aminotransferase classe I e II	104 - 488	Biossíntese de Etileno
OsACS#002	Os03g51740	487	PF00155	Aminotransferase classe I e II	051 - 438	Biossíntese de Etileno
OsACS#003	Os04g48850	483	PF00155	Aminotransferase classe I e II	046 - 427	Biossíntese de Etileno
OsACS#004	Os05g10780	437	PF00155	Aminotransferase classe I e II	035 - 423	Biossíntese de Etileno
OsACS#005	Os05g25490	496	PF00155	Aminotransferase classe I e II	056 - 457	Biossíntese de Etileno
OsACS#006	Os06g03990	542	PF00155	Aminotransferase classe I e II	147 - 531	Biossíntese de Etileno <sup>3</sup>
OsACT#001	Os02g57480	453	PF02458	Família de transferases	006 - 444	Síntese de MACC <sup>4</sup>
OsACT#002	Os03g18590	387	PF02458	Família de transferases	086 - 357	Síntese de MACC <sup>4</sup>
OsACT#003	Os07g23150	525	PF02458	Família de transferases	067 - 510	Síntese de MACC <sup>4</sup>
OsARD#001	Os03g06620	198	PF03079	Família ARD/ARD <sup>7</sup>	015 - 168	Ciclo de Yang
			PF07883	Domínio cupin	097 - 153	
OsARD#002	Os04g27830	284	PF03079	Família ARD/ARD <sup>7</sup>	079 - 247	Ciclo de Yang
			PF07883	Domínio cupin	164 - 219	
OsARD#003	Os10g28350	199	PF03079	Família ARD/ARD <sup>7</sup>	015 - 168	Ciclo de Yang
			PF07883	Domínio cupin	096 - 152	
OsARD#004	Os10g28360	184	PF03079	Família ARD/ARD <sup>7</sup>	015 - 155	Ciclo de Yang
			PF07883	Domínio cupin	085 - 143	
OsASP#001	Os01g55540	460	PF00155	Aminotransferase classe I e II	086 - 452	Ciclo de Yang
OsASP#002	Os02g14110	432	PF00155	Aminotransferase classe I e II	060 - 425	Ciclo de Yang
OsASP#003	Os02g55420	458	PF00155	Aminotransferase classe I e II	084 - 450	Ciclo de Yang
OsASP#004	Os06g35540	463	PF00155	Aminotransferase classe I e II	091 - 456	Ciclo de Yang

**Tabela S6.** Lista de Proteínas Relacionadas com a Biossíntese de Etileno em *Oryza sativa* (cont.)

Proteína de <i>Oryza sativa</i>			Domínio Proteico			Via Metabólica <sup>2</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	Código PFAM	Descrição	Posição	
OsDEP#001	Os11g29370	518	PF00596	Domínio N-terminal de aldolases e aducinas classe II	032 - 232	Ciclo de Yang
			PF00702	Hidrolase semelhante à dehalogenase haloácida	279 - 494	
OsGGT#001	Os01g05810	634	PF01019	Gama-glutamiltranspeptidase	098 - 628	Síntese de GACC <sup>4</sup>
OsGGT#002	Os01g05820	563	PF01019	Gama-glutamiltranspeptidase	094 - 457	Síntese de GACC <sup>4</sup>
OsGGT#003	Os04g38450	574	PF01019	Gama-glutamiltranspeptidase	057 - 569	Síntese de GACC <sup>4</sup>
OsMAT#001	Os01g18860	396	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	006 - 104	Biossíntese de Etileno
			PF02772	S-adenosil-metionina sintetase, domínio central	119 - 241	
			PF02773	S-adenosil-metionina sintetase, domínio C-terminal	243 - 384	
OsMAT#002	Os01g22010	394	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	005 - 103	Biossíntese de Etileno
			PF02772	S-adenosil-metionina sintetase, domínio central	118 - 240	
			PF02773	S-adenosil-metionina sintetase, domínio C-terminal	242 - 382	
OsMAT#003	Os05g04510	396	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	006 - 104	Biossíntese de Etileno
			PF02772	S-adenosil-metionina sintetase, domínio central	119 - 241	
			PF02773	S-adenosil-metionina sintetase, domínio C-terminal	243 - 384	
OsMTI#001	Os11g11050	508	PF01008	Família da subunidade do fator de iniciação 2	183 - 484	Ciclo de Yang
OsMTK#001	Os04g57400	430	PF01636	Família da enzima fosfotransferase	047 - 297	Ciclo de Yang
OsMTK#002	Os04g57410	427	PF01636	Família da enzima fosfotransferase	044 - 294	Ciclo de Yang
OsMTN#001	Os06g02220	259	PF01048	Superfamília das fosforilases	020 - 248	Ciclo de Yang

<sup>1</sup> Tamanho em resíduos de aminoácidos.<sup>2</sup> Este parâmetro foi determinado pela homologia com a sequência de aminoácidos de *Arabidopsis thaliana* equivalente.<sup>3</sup> Possível aminoácido transferase (AAT). Não possui atividade de ACC sintase comprovada.<sup>4</sup> A função primária desta proteína não está inicialmente associada a esta rota metabólica.

**Tabela S7.** Lista de Proteínas Relacionadas com a Transdução de Sinal Mediada por Etileno em *Oryza sativa*

Proteína de <i>Oryza sativa</i>			Domínio Proteico			Suposta Função <sup>2</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	Código PFAM	Descrição	Posição	
OsCPK#001	Os11g07040	513	PF00069	Domínio proteína cinase	047 - 305	Ativação de ACS Tipo II
			PF00036	“EF hand”	352 - 413	
					418 - 483	
OsCPK#002	Os12g07230	526	PF00069	Domínio proteína cinase	049 - 307	Ativação de ACS Tipo II
			PF00036	“EF hand”	354 - 415	
					420 - 485	
OsCTR#001	Os02g32610	781	PF07714	Proteína tirosina cinase	167 - 377	Regulador Negativo de Resposta à Etileno
			PF14381	Proteína cinase responsiva à etileno Le-CTR1	512 - 766	
OsCTR#002	Os09g39320	1220	PF07714	Proteína tirosina cinase	121 - 329	Regulador Negativo de Resposta à Etileno
			PF14381	Proteína cinase responsiva à etileno Le-CTR1	481 - 723	
OsEBF#001	Os02g10700	660	PF00646	Domínio F-box	066 - 111	Direciona EIN3 para Proteólise
OsEBF#002	Os06g40360	664	PF00646	Domínio F-box	077 - 113	Direciona EIN3 para Proteólise
OsEIN#001	Os02g28074	822	PF03159	N-terminal da exonuclease 5'-3' XRN	001 - 230	Silenciamento Gênico de EBF
OsEIN#002	Os03g20780	643	PF04873	Insensível à etileno 3	078 - 464	Fator de Transcrição Nuclear
OsEIN#003	Os03g20790	640	PF04873	Insensível à etileno 3	078 - 464	Fator de Transcrição Nuclear
OsEIN#004	Os03g49400	1299	PF01566	Proteína associada à resistência natural em macrófagos	039 - 391	Regulação após CTR1
OsEIN#005	Os03g58060	988	PF03159	N-terminal da exonuclease 5'-3' XRN	001 - 256	Silenciamento Gênico de EBF
			PF00098	Articulação de zinco	263 - 280	
OsEIN#006	Os07g06130	1281	PF01566	Proteína associada à resistência natural em macrófagos	047 - 399	Regulação após CTR1
OsEIN#007	Os07g06190	1367	PF01566	Proteína associada à resistência natural em macrófagos	258 - 584	Regulação após CTR1
OsEIN#008	Os07g06300	1367	PF01566	Proteína associada à resistência natural em macrófagos	258 - 584	Regulação após CTR1
OsETO#001	Os03g18360	958	PF07719	Repetição tetratricopeptídeo	704 - 737	Regulador Negativo de ACS Tipo II
OsETO#002	Os07g08120	886	PF13414	Repetição TPR	707 - 772	Regulador Negativo de ACS Tipo II
OsETO#003	Os11g37520	886	PF13414	Repetição TPR	804 - 869	Regulador Negativo de ACS Tipo II
OsETR#001	Os02g57530	836	PF01590	Domínio GAF	269 - 413	Receptor de Etileno
			PF00512	Domínio histidina cinase A (fosfoceptor)	449 - 514	
			PF00072	Domínio regulador de resposta de receptor	717 - 830	
OsETR#002	Os03g49500	636	PF01590	Domínio GAF	158 - 307	Receptor de Etileno
			PF00512	Domínio histidina cinase A (fosfoceptor)	343 - 408	
			PF02518	Histidina cinase, DNA girase B e semelhante à ATPase	455 - 586	



**Tabela S7.** Lista de Proteínas Relacionadas com a Transdução de Sinal Mediada por Etileno em *Oryza sativa* (cont.)

Proteína de <i>Oryza sativa</i>			Domínio Proteico			Suposta Função <sup>2</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	Código PFAM	Descrição	Posição	
OsETR#003	Os04g08740	763	PF01590	Domínio GAF	190 - 339	Receptor de Etileno
			PF00072	Domínio regulador de resposta de receptor	642 - 757	
OsETR#004	Os05g06320	518	PF01590	Domínio GAF	159 - 308	Receptor de Etileno
			PF00512	Domínio histidina cinase A (fosfoceptor)	344 - 409	
OsETR#005	Os07g15540	777	PF01590	Domínio GAF	184 - 344	Receptor de Etileno
			PF00072	Domínio regulador de resposta de receptor	646 - 770	
OsMKK#001	Os01g32660	355	PF00069	Domínio proteína cinase	071 - 330	Ativação de ACS Tipo I
OsMKK#002	Os02g54600	369	PF00069	Domínio proteína cinase	093 - 350	Ativação de ACS Tipo I
OsMKK#003	Os03g12390	339	PF00069	Domínio proteína cinase	049 - 314	Ativação de EIN3
OsMKK#004	Os03g50550	345	PF00069	Domínio proteína cinase	060 - 326	Ativação de EIN3
OsMKK#005	Os06g09180	342	PF00069	Domínio proteína cinase	070 - 327	Ativação de ACS Tipo I
OsMPK#001	Os03g17700	369	PF00069	Domínio proteína cinase	038 - 322	Ativação de EIN3
OsMPK#002	Os06g06090	398	PF00069	Domínio proteína cinase	070 - 352	Ativação de EIN3 e ACS Tipo I
OsRAN#001	Os05g49890	221	PF00071	Família ras	015 - 173	Transporte de Cobre
OsRAN#002	Os06g39875	226	PF00071	Família ras	019 - 177	Transporte de Cobre
OsRTE#001	Os01g51430	241	PF05608	Proteína sem função conhecida (DUF778)	051 - 186	Regulador Positivo do Receptor de Etileno
OsRTE#002	Os03g58520	204	PF05608	Proteína sem função conhecida (DUF778)	027 - 148	Regulador Positivo do Receptor de Etileno
OsRTE#003	Os05g46240	233	PF05608	Proteína sem função conhecida (DUF778)	043 - 178	Regulador Positivo do Receptor de Etileno

<sup>1</sup> Tamanho em resíduos de aminoácidos.<sup>2</sup> Este parâmetro foi determinado pela homologia com a sequência de aminoácidos de *Arabidopsis thaliana* equivalente.

**Tabela S8.** Lista de Proteínas AP2 de *Oryza sativa*

Proteína de <i>Oryza sativa</i>			Domínio AP2		Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>3</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	Nº <sup>2</sup>	Posição	
OsAP2#001	Os01g67410	695	2	282 - 342 384 - 436	Semelhante à “Aintegumenta” 6 (AIL5); “Plethora” 5 (PLT5); “Chotto” 1 (CHO1)
OsAP2#002	Os02g40070	700	2	302 - 359 401 - 453	“Plethora” 2 (PLT2)
OsAP2#003	Os03g07940	413	1	218 - 270	Proteína da Superfamília de um Tipo de Integrase que se Liga ao DNA
OsAP2#004	Os03g19900	490	1	294 - 346	Semelhante à “Aintegumenta” 6 (AIL6); “Plethora” 3 (PLT3)
OsAP2#005	Os03g56050	652	2	304 - 361 403 - 455	Complementa Mutante da Proteína Cinase C 1 (CKC1); “Aintegumenta” (ANT)
OsAP2#006	Os03g60430	434	2	117 - 167 209 - 260	Mutante Floral 2 (FLO2; AP2; FL1; ATAP2)
OsAP2#007	Os04g42570	658	2	278 - 338 380 - 432	“Plethora” 1 (PLT1)
OsAP2#008	Os04g55560	460	2	110 - 160 202 - 253	Mutante Floral 2 (FLO2; AP2; FL1; ATAP2)
OsAP2#009	Os04g55970	492	2	143 - 200 242 - 294	Semelhante à “Aintegumenta” 6 (AIL5); “Plethora” 5 (PLT5); “Chotto” 1 (CHO1)
OsAP2#010	Os05g03040	512	2	170 - 220 262 - 313	Alvo Marcado de Início de Ativação 1 (TOE1); RAP2.7
OsAP2#011	Os05g32270	314	1	003 - 037	Proteína da Superfamília de um Tipo de Integrase que se Liga ao DNA
OsAP2#012	Os05g45954	431	2	095 - 149 191 - 243	“Wrinkled” 1 (WRI1)
OsAP2#013	Os06g05340	399	2	080 - 140 182 - 234	“Wrinkled” 1 (WRI1)
OsAP2#014	Os06g43220	334	2	001 - 034 076 - 127	Mutante Floral 2 (FLO2; AP2; FL1; ATAP2)
OsAP2#015	Os07g03250	314	1	061 - 113	Complementa Mutante da Proteína Cinase C 1 (CKC1); “Aintegumenta” (ANT)
OsAP2#016	Os07g13170	436	2	121 - 171 213 - 264	Alvo Marcado de Início de Ativação 1 (TOE1); RAP2.7
OsAP2#017	Os11g03540	449	2	082 - 143 185 - 237	“Wrinkled” 1 (WRI1)

**Tabela S8.** Lista de Proteínas AP2 de *Oryza sativa* (cont.)

Proteína de <i>Oryza sativa</i>			Domínio AP2		Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>3</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	Nº <sup>2</sup>	Posição	
OsAP2#018	Os11g19060	559	2	167 - 227	Mutante Floral 2 (FLO2; AP2; FL1; ATAP2)
				269 - 321	
OsAP2#019	Os12g03290	443	2	079 - 137	“Wrinkled” 1 (WRI1)
				179 - 231	

<sup>1</sup> Tamanho em resíduos de aminoácidos.

<sup>2</sup> Número de domínios AP2 (PF00847) presentes na sequência de aminoácidos em questão.

<sup>3</sup> Este parâmetro foi determinado pela homologia com a sequência de aminoácidos de *Arabidopsis thaliana* equivalente.

**Tabela S9.** Lista dos Fatores Transcricionais Responsivos à Etileno (ERF) de *Oryza sativa*<sup>1</sup>

Nome Genérico	Locus ID	C-II <sup>2</sup>	Nome Genérico	Locus ID	C-II <sup>2</sup>
OsERF#001	Os06g40150	V	OsERF#071	Os06g09390	VII
OsERF#002	Os06g08340	V	OsERF#072	Os09g26420	VII
OsERF#003	Os02g10760	V	OsERF#073	Os09g11460	VII
OsERF#004	Os12g39330	V	OsERF#074	Os05g41780	VIII
OsERF#005	Os07g10410	V	OsERF#075	Os01g58420	VIII
OsERF#006	Os07g38750	V	OsERF#076	Os04g57340	VIII
OsERF#007	Os06g07030	II	OsERF#077	Os04g52090	VIII
OsERF#008	Os04g55520	II	OsERF#078	Os07g47330	VIII
OsERF#009	Os03g15660	II	OsERF#079	Os02g38090	VIII
OsERF#010	Os06g09690	II	OsERF#080	Os08g07700	VIII
OsERF#011	Os02g54050	II	OsERF#081	Os02g32040	VIII
OsERF#012	Os08g35240	II	OsERF#082	Os04g32790	VIII
OsERF#013	Os06g11940	II	OsERF#083	Os03g64260	IX
OsERF#014	Os06g09810	II	OsERF#084	Os05g49010	IX
OsERF#015	Os06g09790	II	OsERF#085	Os05g37640	IX
OsERF#016	Os06g09760	II	OsERF#086	Os07g22770	IX
OsERF#017	Os01g66270	II	OsERF#087	Os09g39850	IX
OsERF#018	Os10g38000	III	OsERF#088	Os03g05590	IX
OsERF#019	Os11g13840	III	OsERF#089	Os10g30840	IX
OsERF#020	Os02g45420	III	OsERF#090	Os08g44960	IX
OsERF#021	Os02g35240	III	OsERF#091	Os02g43790	IX
OsERF#022	Os04g36640	III	OsERF#092	Os01g54890	IX
OsERF#023	Os04g48330	III	OsERF#093	Os04g46220	IX
OsERF#024	Os09g35030	III	OsERF#094	Os04g46250	IX
OsERF#025	Os02g45450	III	OsERF#095	Os02g43820	IX
OsERF#026	Os06g03670	III	OsERF#096	Os10g41330	IX
OsERF#027	Os01g73770	III	OsERF#097	Os04g46240	IX
OsERF#028	Os08g43200	III	OsERF#098	Os02g34260	X
OsERF#029	Os08g43210	III	OsERF#099	Os01g64790	X
OsERF#030	Os04g48350	III	OsERF#100	Os04g34970	X
OsERF#031	Os09g35010	III	OsERF#101	Os04g32620	X
OsERF#032	Os02g43940	III	OsERF#102	Os09g28440	X
OsERF#033	Os04g46400	III	OsERF#103	Os02g52670	X
OsERF#034	Os04g46440	III	OsERF#104	Os08g36920	X
OsERF#035	Os02g43970	III	OsERF#105	Os05g36100	X
OsERF#036	Os10g41130	III	OsERF#106	Os08g42550	X
OsERF#037	Os04g46410	III	OsERF#107	Os02g32140	X
OsERF#038	Os02g13710	III	OsERF#108	Os01g04020	VI-L
OsERF#039	Os01g10370	III	OsERF#109	Os09g13940	VI-L
OsERF#040	Os01g07120	IV	OsERF#110	Os12g41030	XI
OsERF#041	Os03g07830	IV	OsERF#111	Os12g41040	XI
OsERF#042	Os05g27930	IV	OsERF#112	Os12g41060	XI
OsERF#043	Os05g39590	IV	OsERF#113	Os06g42990	XII
OsERF#044	Os08g45110	IV	OsERF#114	Os06g42910	XI
OsERF#045	Os04g44670	I	OsERF#115	Os08g41030	XIV
OsERF#046	Os02g42580	I	OsERF#116	AB023482	III
OsERF#047	Os03g09170	I	OsERF#117	Os05g28350	IV
OsERF#048	Os08g31580	I	OsERF#118	Os11g06770	X
OsERF#049	Os02g51670	I	OsERF#119	Os06g10780	II
OsERF#050	Os09g20350	I	OsERF#120	Os06g11860	I
OsERF#051	Os10g22600	I	OsERF#121	Os06g47590	VIII
OsERF#052	Os05g49700	I	OsERF#122	Os06g36000	III
OsERF#053	Os01g12440	VI	OsERF#123	Os09g39810	IX
OsERF#054	Os01g46870	VI	OsERF#124	Os12g07030	X
OsERF#055	Os06g06540	VI	OsERF#125	Os02g34270	X
OsERF#056	Os05g25260	VI	OsERF#126	Os02g52880	II
OsERF#057	Os07g12510	VI	OsERF#127	Os02g55380	V
OsERF#058	Os03g60120	VI	OsERF#128	Os04g18650	IX
OsERF#059	Os10g25170	VII	OsERF#129	Os04g56150	V
OsERF#060	Os03g08460	VII	OsERF#130	Os05g41760	VIII
OsERF#061	Os05g29810	VII	OsERF#131	Os05g34730	II
OsERF#062	Os03g08470	VII	OsERF#132	Os02g06330	VIII
OsERF#063	Os09g11480	VII	OsERF#133	Os09g35020	III
OsERF#064	Os03g08500	VII	OsERF#134	Os02g09650	VIII
OsERF#065	Os07g42510	VII	OsERF#135	BAC99579	XIII
OsERF#066	Os03g22170	VII	OsERF#136	Os07g22730	IX
OsERF#067	Os07g47790	VII	OsERF#137	Os03g02650	III
OsERF#068	Os01g21120	VII	OsERF#138	Os08g27220	VI-L
OsERF#069	Os03g08490	VII	OsERF#139	Os06g09730	II
OsERF#070	Os02g54160	VII	-	-	-

<sup>1</sup> Lista de ERFs de *Oryza sativa* obtida no trabalho de Nakano et al. (2006).<sup>2</sup> Classificação realizada por Nakano et al. (2006).

**Tabela S10.** Lista de Proteínas RAV de *Oryza sativa*

Proteína de <i>Oryza sativa</i>			Domínios AP2/B3		Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>3</sup>
Nome Genérico	Locus ID	R.A. <sup>2</sup>	Código PFAM <sup>3</sup>	Posição	
OsRAV#001	Os01g04750	317	PF00847	065 - 113	Relacionado à ABI3/VP1 2 (RAV2); “Tempranillo” 2 (TEM2)
			PF02362	178 - 273	
OsRAV#002	Os01g04800	365	PF00847	067 - 115	Fator de Interação com DNA em Resposta à Etileno 1 (EDF1); “Tempranillo” 1 (TEM1)
			PF02362	182 - 282	
OsRAV#003	Os01g49830	393	PF00847	080 - 128	Relacionado à ABI3/VP1 2 (RAV2); “Tempranillo” 2 (TEM2)
			PF02362	216 - 319	
OsRAV#004	Os05g47650	394	PF00847	067 - 115	Fator de Interação com DNA em Resposta à Etileno 1 (EDF1); “Tempranillo” 1 (TEM1)
			PF02362	193 - 309	

<sup>1</sup> Tamanho em resíduos de aminoácidos.

<sup>2</sup> PF00847: Domínio AP2; PF02362: Domínio B3 de interação com DNA.

<sup>3</sup> Este parâmetro foi determinado pela homologia com a sequência de aminoácidos de *Arabidopsis thaliana* equivalente.

**Tabela S11.** Sumário das Proteínas de Soja Relacionadas com a Biossíntese de Etileno

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio da Proteína de Soja			Via Metabólica <sup>2</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Código PFAM	Descrição	Posição	
			Nome Genérico	Locus ID	Nome Genérico	Locus ID				
GmAAT#001	Glyma02g01830	401	AtAAT	At1g77670	OsAAT#002	Os10g34350	PF00155	Aminotransferase classe I e II	027 - 388	Ciclo de Yang
GmAAT#002	Glyma04g43080	450	AtAAT	At1g77670	OsAAT#001	Os09g28050	PF00155	Aminotransferase classe I e II	091 - 444	Ciclo de Yang
GmAAT#003	Glyma06g11630	254	AtAAT	At1g77670	OsAAT#001	Os09g28050	PF00155	Aminotransferase classe I e II	001 - 250	Ciclo de Yang
GmAAT#004	Glyma06g11640	439	AtAAT	At1g77670	OsAAT#001	Os09g28050	PF00155	Aminotransferase classe I e II	080 - 433	Ciclo de Yang
GmACD#001	Glyma05g33540	379	AtACD1	At1g48420	OsACD#001	Os02g53330	PF00291	Enzima dependente de piridoxal fosfato	032 - 341	Degradação de ACC
GmACD#002	Glyma08g06170	223	AtACD1	At1g48420	OsACD#001	Os02g53330	PF00291	Enzima dependente de piridoxal fosfato	013 - 185	Degradação de ACC
GmACO#001	Glyma02g43560	315	AtACO4	At1g05010	OsACO#006	Os09g27820	PF03171	Superfamília de oxigenases 2OG-Fe(II)	153 - 253	Biossíntese de Etileno
GmACO#002	Glyma02g43580	307	AtACO4	At1g05010	OsACO#006	Os09g27820	PF03171	Superfamília de oxigenases 2OG-Fe(II)	150 - 250	Biossíntese de Etileno
GmACO#003	Glyma02g43600	291	AtACO4	At1g05010	OsACO#006	Os09g27820	PF03171	Superfamília de oxigenases 2OG-Fe(II)	134 - 234	Biossíntese de Etileno
GmACO#004	Glyma04g42460	308	AtACO#002	At1g77330	OsACO#003	Os05g05680	PF03171	Superfamília de oxigenases 2OG-Fe(II)	154 - 254	Biossíntese de Etileno
GmACO#005	Glyma05g36310	307	AtACO1	At2g19590	OsACO#004	Os06g37590	PF03171	Superfamília de oxigenases 2OG-Fe(II)	152 - 253	Biossíntese de Etileno
GmACO#006	Glyma06g12340	307	AtACO#002	At1g77330	OsACO#003	Os05g05680	PF03171	Superfamília de oxigenases 2OG-Fe(II)	153 - 253	Biossíntese de Etileno
GmACO#007	Glyma07g15480	306	AtACO1	At2g19590	OsACO#004	Os06g37590	PF03171	Superfamília de oxigenases 2OG-Fe(II)	151 - 252	Biossíntese de Etileno
GmACO#008	Glyma07g39420	318	AtACO4	At1g05010	OsACO#005	Os09g27750	PF03171	Superfamília de oxigenases 2OG-Fe(II)	153 - 253	Biossíntese de Etileno
GmACO#009	Glyma08g03310	307	AtACO1	At2g19590	OsACO#004	Os06g37590	PF03171	Superfamília de oxigenases 2OG-Fe(II)	152 - 253	Biossíntese de Etileno
GmACO#010	Glyma08g05500	310	AtACO4	At1g05010	OsACO#006	Os09g27820	PF03171	Superfamília de oxigenases 2OG-Fe(II)	153 - 253	Biossíntese de Etileno
GmACO#011	Glyma09g01110	318	AtACO4	At1g05010	OsACO#005	Os09g27750	PF03171	Superfamília de oxigenases 2OG-Fe(II)	153 - 253	Biossíntese de Etileno
GmACO#012	Glyma14g05350	307	AtACO4	At1g05010	OsACO#006	Os09g27820	PF03171	Superfamília de oxigenases 2OG-Fe(II)	150 - 250	Biossíntese de Etileno
GmACO#013	Glyma14g05360	307	AtACO4	At1g05010	OsACO#006	Os09g27820	PF03171	Superfamília de oxigenases 2OG-Fe(II)	150 - 250	Biossíntese de Etileno
GmACO#014	Glyma14g05390	315	AtACO4	At1g05010	OsACO#002	Os02g53180	PF03171	Superfamília de oxigenases 2OG-Fe(II)	153 - 253	Biossíntese de Etileno
GmACO#015	Glyma15g11930	318	AtACO4	At1g05010	OsACO#005	Os09g27750	PF03171	Superfamília de oxigenases 2OG-Fe(II)	153 - 253	Biossíntese de Etileno
GmACO#016	Glyma17g01330	319	AtACO4	At1g05010	OsACO#005	Os09g27750	PF03171	Superfamília de oxigenases 2OG-Fe(II)	154 - 254	Biossíntese de Etileno
GmACS#001	Glyma01g00700	442	AtACS7	At4g26200	OsACS#004	Os05g10780	PF00155	Aminotransferase classe I e II	044 - 428	Biossíntese de Etileno
GmACS#002	Glyma01g40400	470	AtACS8	At4g37770	OsACS#002	Os03g51740	PF00155	Aminotransferase classe I e II	040 - 423	Biossíntese de Etileno
GmACS#003	Glyma01g42290	502	AtACS10	At1g62960	OsACS#006	Os06g03990	PF00155	Aminotransferase classe I e II	113 - 491	Biossíntese de Etileno <sup>3</sup>
GmACS#004	Glyma04g05150	437	AtACS8	At4g37770	OsACS#002	Os03g51740	PF00155	Aminotransferase classe I e II	040 - 426	Biossíntese de Etileno
GmACS#005	Glyma05g23020	480	AtACS8	At4g37770	OsACS#002	Os03g51740	PF00155	Aminotransferase classe I e II	040 - 427	Biossíntese de Etileno
GmACS#006	Glyma05g36250	440	AtACS7	At4g26200	OsACS#005	Os05g25490	PF00155	Aminotransferase classe I e II	050 - 430	Biossíntese de Etileno
GmACS#007	Glyma05g37410	434	AtACS1	At3g61510	OsACS#003	Os04g48850	PF00155	Aminotransferase classe I e II	001 - 381	Biossíntese de Etileno
GmACS#008	Glyma06g05240	354	AtACS8	At4g37770	OsACS#002	Os03g51740	PF00155	Aminotransferase classe I e II	001 - 341	Biossíntese de Etileno

**Tabela S11.** Sumário das Proteínas de Soja Relacionadas com a Biossíntese de Etileno (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio da Proteína de Soja			Via Metabólica <sup>2</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Código PFAM	Descrição	Posição	
			Nome Genérico	Locus ID	Nome Genérico	Locus ID				
GmACS#009	Glyma07g07160	474	AtACS1	At3g61510	OsACS#003	Os04g48850	PF00155	Aminotransferase classe I e II	043 - 424	Biossíntese de Etileno
GmACS#010	Glyma07g15380	426	AtACS7	At4g26200	OsACS#004	Os05g10780	PF00155	Aminotransferase classe I e II	044 - 425	Biossíntese de Etileno
GmACS#011	Glyma08g02130	484	AtACS6	At4g11280	OsACS#003	Os04g48850	PF00155	Aminotransferase classe I e II	048 - 431	Biossíntese de Etileno
GmACS#012	Glyma08g03400	440	AtACS7	At4g26200	OsACS#004	Os05g10780	PF00155	Aminotransferase classe I e II	050 - 430	Biossíntese de Etileno
GmACS#013	Glyma09g28000	500	AtACS12	At5g51690	OsACS#006	Os06g03990	PF00155	Aminotransferase classe I e II	110 - 489	Biossíntese de Etileno <sup>3</sup>
GmACS#014	Glyma09g39060	485	AtACS1	At3g61510	OsACS#003	Os04g48850	PF00155	Aminotransferase classe I e II	047 - 428	Biossíntese de Etileno
GmACS#015	Glyma11g02390	465	AtACS6	At4g11280	OsACS#003	Os04g48850	PF00155	Aminotransferase classe I e II	038 - 421	Biossíntese de Etileno
GmACS#016	Glyma11g03070	501	AtACS10	At1g62960	OsACS#006	Os06g03990	PF00155	Aminotransferase classe I e II	112 - 490	Biossíntese de Etileno <sup>3</sup>
GmACS#017	Glyma11g04890	471	AtACS8	At4g37770	OsACS#002	Os03g51740	PF00155	Aminotransferase classe I e II	040 - 423	Biossíntese de Etileno
GmACS#018	Glyma16g03600	474	AtACS1	At3g61510	OsACS#003	Os04g48850	PF00155	Aminotransferase classe I e II	043 - 424	Biossíntese de Etileno
GmACS#019	Glyma16g32860	517	AtACS12	At5g51690	OsACS#006	Os06g03990	PF00155	Aminotransferase classe I e II	127 - 506	Biossíntese de Etileno <sup>3</sup>
GmACS#020	Glyma17g16990	475	AtACS8	At4g37770	OsACS#002	Os03g51740	PF00155	Aminotransferase classe I e II	040 - 425	Biossíntese de Etileno
GmACS#021	Glyma18g47280	495	AtACS1	At3g61510	OsACS#003	Os04g48850	PF00155	Aminotransferase classe I e II	047 - 428	Biossíntese de Etileno
GmACT#001	Glyma08g27120	430	AtACT#006	At5g39090	OsACT#001	Os02g57480	PF02458	Família de transferases	002 - 429	Síntese de MACC <sup>4</sup>
GmACT#002	Glyma08g27130	447	AtACT#004	At3g29670	OsACT#003	Os07g23150	PF02458	Família de transferases	242 - 239	Síntese de MACC <sup>4</sup>
GmACT#003	Glyma08g27500	469	AtACT#001	At1g03940	OsACT#001	Os02g57480	PF02458	Família de transferases	005 - 461	Síntese de MACC <sup>4</sup>
GmACT#004	Glyma11g29490	354	AtACT#003	At2g30200	OsACT#002	Os03g18590	PF02458	Família de transferases	053 - 353	Síntese de MACC <sup>4</sup>
GmACT#005	Glyma11g29520	230	AtACT#003	At2g30200	OsACT#002	Os03g18590	PF02458	Família de transferases	002 - 227	Síntese de MACC <sup>4</sup>
GmACT#006	Glyma12g32630	421	AtACT#002	At1g03495	OsACT#001	Os02g57480	PF02458	Família de transferases	002 - 421	Síntese de MACC <sup>4</sup>
GmACT#007	Glyma12g32640	466	AtACT#002	At1g03495	OsACT#003	Os07g23150	PF02458	Família de transferases	008 - 453	Síntese de MACC <sup>4</sup>
GmACT#008	Glyma12g32650	443	AtACT#001	At1g03940	OsACT#001	Os02g57480	PF02458	Família de transferases	213 - 435	Síntese de MACC <sup>4</sup>
GmACT#009	Glyma12g32660	467	AtACT#001	At1g03940	OsACT#001	Os02g57480	PF02458	Família de transferases	008 - 461	Síntese de MACC <sup>4</sup>
GmACT#010	Glyma13g06230	467	AtACT#006	At5g39090	OsACT#001	Os02g57480	PF02458	Família de transferases	005 - 462	Síntese de MACC <sup>4</sup>
GmACT#011	Glyma13g06550	449	AtACT#005	At5g39050	OsACT#001	Os02g57480	PF02458	Família de transferases	015 - 448	Síntese de MACC <sup>4</sup>
GmACT#012	Glyma13g37810	469	AtACT#001	At1g03940	OsACT#001	Os02g57480	PF02458	Família de transferases	002 - 462	Síntese de MACC <sup>4</sup>
GmACT#013	Glyma13g37830	462	AtACT#002	At1g03495	OsACT#001	Os02g57480	PF02458	Família de transferases	008 - 450	Síntese de MACC <sup>4</sup>
GmACT#014	Glyma13g37840	405	AtACT#002	At1g03495	OsACT#001	Os02g57480	PF02458	Família de transferases	001 - 405	Síntese de MACC <sup>4</sup>
GmACT#015	Glyma13g37850	441	AtACT#002	At1g03495	OsACT#001	Os02g57480	PF02458	Família de transferases	010 - 433	Síntese de MACC <sup>4</sup>
GmACT#016	Glyma14g03820	473	AtACT#001	At1g03940	OsACT#003	Os07g23150	PF02458	Família de transferases	005 - 464	Síntese de MACC <sup>4</sup>
GmACT#017	Glyma18g06500	358	AtACT#003	At2g30200	OsACT#002	Os03g18590	PF02458	Família de transferases	056 - 356	Síntese de MACC <sup>4</sup>

**Tabela S11.** Sumário das Proteínas de Soja Relacionadas com a Biossíntese de Etileno (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio da Proteína de Soja			Via Metabólica <sup>2</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Código PFAM	Descrição	Posição	
			Nome Genérico	Locus ID	Nome Genérico	Locus ID				
GmACT#018	Glyma18g49240	511	AtACT#004	At3g29670	OsACT#001	Os02g57480	PF02458	Família de transferases	004 - 507	Síntese de MACC <sup>4</sup>
GmACT#019	Glyma18g50310	479	AtACT#005	At5g39050	OsACT#001	Os02g57480	PF02458	Família de transferases	006 - 475	Síntese de MACC <sup>4</sup>
GmACT#020	Glyma18g50320	476	AtACT#005	At5g39050	OsACT#001	Os02g57480	PF02458	Família de transferases	004 - 473	Síntese de MACC <sup>4</sup>
GmACT#021	Glyma18g50330	452	AtACT#005	At5g39050	OsACT#001	Os02g57480	PF02458	Família de transferases	001 - 452	Síntese de MACC <sup>4</sup>
GmACT#022	Glyma18g50340	450	AtACT#006	At5g39090	OsACT#001	Os02g57480	PF02458	Família de transferases	001 - 444	Síntese de MACC <sup>4</sup>
GmACT#023	Glyma18g50350	450	AtACT#006	At5g39090	OsACT#001	Os02g57480	PF02458	Família de transferases	003 - 444	Síntese de MACC <sup>4</sup>
GmACT#024	Glyma18g50360	389	AtACT#004	At3g29670	OsACT#001	Os02g57480	PF02458	Família de transferases	019 - 154	Síntese de MACC <sup>4</sup>
GmACT#025	Glyma19g03730	460	AtACT#006	At5g39090	OsACT#001	Os02g57480	PF02458	Família de transferases	004 - 456	Síntese de MACC <sup>4</sup>
GmACT#026	Glyma19g03760	476	AtACT#006	At5g39090	OsACT#001	Os02g57480	PF02458	Família de transferases	016 - 472	Síntese de MACC <sup>4</sup>
GmACT#027	Glyma19g03770	464	AtACT#006	At5g39090	OsACT#001	Os02g57480	PF02458	Família de transferases	019 - 461	Síntese de MACC <sup>4</sup>
GmACT#028	Glyma19g05220	457	AtACT#002	At1g03495	OsACT#001	Os02g57480	PF02458	Família de transferases	012 - 412	Síntese de MACC <sup>4</sup>
GmACT#029	Glyma19g05290	477	AtACT#002	At1g03495	OsACT#001	Os02g57480	PF02458	Família de transferases	011 - 471	Síntese de MACC <sup>4</sup>
GmACT#030	Glyma19g11320	451	AtACT#002	At1g03495	OsACT#001	Os02g57480	PF02458	Família de transferases	008 - 304	Síntese de MACC <sup>4</sup>
GmARD#001	Glyma10g38940	200	AtARD2	At4g14710	OsARD#001	Os03g06620	PF03079	Família ARD/ARD'	014 - 168	Ciclo de Yang
							PF07883	Domínio cupin	088 - 160	
GmARD#002	Glyma10g38960	201	AtARD4	At5g43850	OsARD#004	Os10g28360	PF03079	Família ARD/ARD'	018 - 172	Ciclo de Yang
							PF07883	Domínio cupin	092 - 164	
GmARD#003	Glyma20g28850	187	AtARD4	At5g43850	OsARD#001	Os03g06620	PF03079	Família ARD/ARD'	004 - 158	Ciclo de Yang
							PF07883	Domínio cupin	078 - 150	
GmARD#004	Glyma20g28860	200	AtARD2	At4g14710	OsARD#001	Os03g06620	PF03079	Família ARD/ARD'	014 - 168	Ciclo de Yang
							PF07883	Domínio cupin	088 - 160	
GmASP#001	Glyma04g08560	416	AtASP3	At5g11520	OsASP#001	Os01g55540	PF00155	Aminotransferase classe I e II	040 - 408	Ciclo de Yang
GmASP#002	Glyma06g08670	456	AtASP3	At5g11520	OsASP#001	Os01g55540	PF00155	Aminotransferase classe I e II	080 - 448	Ciclo de Yang
GmASP#003	Glyma14g13480	463	AtASP5	At4g31990	OsASP#002	Os02g55420	PF00155	Aminotransferase classe I e II	087 - 455	Ciclo de Yang
GmASP#004	Glyma17g33050	466	AtASP5	At4g31990	OsASP#002	Os02g55420	PF00155	Aminotransferase classe I e II	090 - 458	Ciclo de Yang
GmDEP#001	Glyma06g36620	518	AtDEP1	At5g53850	OsDEP#001	Os11g29370	PF00596	Domínio N-terminal de aldolases e aducinas classe II	027 - 230	Ciclo de Yang
							PF00702	Hidrolase semelhante à dehalogenase haloácida	276 - 489	
GmDEP#002	Glyma12g35390	522	AtDEP1	At5g53850	OsDEP#001	Os11g29370	PF00596	Domínio N-terminal de aldolases e aducinas classe II	030 - 233	Ciclo de Yang
							PF00702	Hidrolase semelhante à dehalogenase haloácida	280 - 493	



**Tabela S11.** Sumário das Proteínas de Soja Relacionadas com a Biossíntese de Etileno (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio da Proteína de Soja			Via Metabólica <sup>2</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Código PFAM	Descrição	Posição	
			Nome Genérico	Locus ID	Nome Genérico	Locus ID				
GmDEP#003	Glyma13g35120	522	AtDEP1	At5g53850	OsDEP#001	Os11g29370	PF00596	Domínio N-terminal de aldolases e aducinas classe II	030 - 233	Ciclo de Yang
							PF00702	Hidrolase semelhante à dehalogenase haloácida	280 - 493	
GmGGT#001	Glyma01g00850	626	AtGGT4	At4g29210	OsGGT#001	Os01g05810	PF01019	Gama-glutamyltranspeptidase	089 - 621	Síntese de GACC <sup>4</sup>
GmGGT#002	Glyma07g15150	618	AtGGT4	At4g29210	OsGGT#001	Os01g05810	PF01019	Gama-glutamyltranspeptidase	081 - 613	Síntese de GACC <sup>4</sup>
GmGGT#003	Glyma11g35950	613	AtGGT1	At4g39640	OsGGT#003	Os04g38450	PF01019	Gama-glutamyltranspeptidase	086 - 608	Síntese de GACC <sup>4</sup>
GmGGT#004	Glyma11g35990	595	AtGGT1	At4g39640	OsGGT#003	Os04g38450	PF01019	Gama-glutamyltranspeptidase	086 - 590	Síntese de GACC <sup>4</sup>
GmGGT#005	Glyma18g02450	624	AtGGT1	At4g39640	OsGGT#003	Os04g38450	PF01019	Gama-glutamyltranspeptidase	098 - 619	Síntese de GACC <sup>4</sup>
GmMAT#001	Glyma02g01170	299	AtMAT2	At4g01850	OsMAT#002	Os01g22010	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	011 - 084	Biossíntese de Etileno
							PF02772	S-adenosil-metionina sintetase, domínio central	099 - 164	
							PF02773	S-adenosil-metionina sintetase, domínio C-terminal	166 - 288	
GmMAT#002	Glyma03g34120	390	AtMAT3	At2g36880	OsMAT#002	Os01g22010	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	002 - 101	Biossíntese de Etileno
							PF02772	S-adenosil-metionina sintetase, domínio central	116 - 238	
							PF02773	S-adenosil-metionina sintetase, domínio C-terminal	240 - 382	
GmMAT#003	Glyma03g38190	394	AtMAT2	At4g01850	OsMAT#002	Os01g22010	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	004 - 103	Biossíntese de Etileno
							PF02772	S-adenosil-metionina sintetase, domínio central	118 - 240	
							PF02773	S-adenosil-metionina sintetase, domínio C-terminal	242 - 383	
GmMAT#004	Glyma07g36150	392	AtMAT4	At3g17390	OsMAT#003	Os05g04510	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	003 - 102	Biossíntese de Etileno
							PF02772	S-adenosil-metionina sintetase, domínio central	117 - 239	
							PF02773	S-adenosil-metionina sintetase, domínio C-terminal	241 - 383	
GmMAT#005	Glyma10g06170	390	AtMAT3	At2g36880	OsMAT#001	Os01g18860	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	002 - 101	Biossíntese de Etileno
							PF02772	S-adenosil-metionina sintetase, domínio central	116 - 238	
							PF02773	S-adenosil-metionina sintetase, domínio C-terminal	240 - 382	
GmMAT#006	Glyma10g28500	389	AtMAT2	At4g01850	OsMAT#001	Os01g18860	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	002 - 101	Biossíntese de Etileno
							PF02772	S-adenosil-metionina sintetase, domínio central	116 - 238	
							PF02773	S-adenosil-metionina sintetase, domínio C-terminal	240 - 382	
GmMAT#007	Glyma13g20480	390	AtMAT3	At2g36880	OsMAT#001	Os01g18860	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	002 - 101	Biossíntese de Etileno
							PF02772	S-adenosil-metionina sintetase, domínio central	116 - 238	
							PF02773	S-adenosil-metionina sintetase, domínio C-terminal	240 - 382	

**Tabela S11.** Sumário das Proteínas de Soja Relacionadas com a Biossíntese de Etileno (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio da Proteína de Soja			Via Metabólica <sup>2</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Código PFAM	Descrição	Posição	
GmMAT#008	Glyma15g21890	395	AtMAT4	At3g17390	OsMAT#002	Os01g22010	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	004 - 103	Biossíntese de Etileno
							PF02772	S-adenosil-metionina sintetase, domínio central	118 - 240	
							PF02773	S-adenosil-metionina sintetase, domínio C-terminal	242 - 384	
GmMAT#009	Glyma17g04330	392	AtMAT4	At3g17390	OsMAT#003	Os05g04510	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	003 - 102	Biossíntese de Etileno
							PF02772	S-adenosil-metionina sintetase, domínio central	117 - 239	
							PF02773	S-adenosil-metionina sintetase, domínio C-terminal	241 - 383	
GmMAT#010	Glyma17g04340	392	AtMAT4	At3g17390	OsMAT#003	Os05g04510	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	003 - 102	Biossíntese de Etileno
							PF02772	S-adenosil-metionina sintetase, domínio central	117 - 239	
							PF02773	S-adenosil-metionina sintetase, domínio C-terminal	241 - 383	
GmMAT#011	Glyma19g40810	394	AtMAT2	At4g01850	OsMAT#002	Os01g22010	PF00438	S-adenosil-metionina sintetase, domínio N-terminal	004 - 103	Biossíntese de Etileno
							PF02772	S-adenosil-metionina sintetase, domínio central	118 - 240	
							PF02773	S-adenosil-metionina sintetase, domínio C-terminal	242 - 383	
GmMTI#001	Glyma09g08190	286	AtMTI1	At2g05830	OsMTI#001	Os11g11050	PF01008	Família da subunidade do fator de iniciação 2	054 - 272	Ciclo de Yang
GmMTI#002	Glyma15g19690	200	AtMTI1	At2g05830	OsMTI#001	Os11g11050	PF01008	Família da subunidade do fator de iniciação 2	001 - 189	Ciclo de Yang
GmMTK#001	Glyma01g45520	425	AtMTK1	At1g49820	OsMTK#001	Os04g57400	PF01636	Família da enzima fosfotransferase	037 - 251	Ciclo de Yang
GmMTK#002	Glyma10g40730	417	AtMTK1	At1g49820	OsMTK#001	Os04g57400	PF01636	Família da enzima fosfotransferase	036 - 296	Ciclo de Yang
GmMTK#003	Glyma20g26580	417	AtMTK1	At1g49820	OsMTK#001	Os04g57400	PF01636	Família da enzima fosfotransferase	036 - 296	Ciclo de Yang
GmMTN#001	Glyma04g00770	265	AtMTN1	At4g38800	OsMTN#001	Os06g02220	PF01048	Superfamília das fosforilases	026 - 255	Ciclo de Yang
GmMTN#002	Glyma06g00790	266	AtMTN1	At4g38800	OsMTN#001	Os06g02220	PF01048	Superfamília das fosforilases	027 - 256	Ciclo de Yang
GmMTN#003	Glyma06g00800	253	AtMTN1	At4g38800	OsMTN#001	Os06g02220	PF01048	Superfamília das fosforilases	014 - 243	Ciclo de Yang

<sup>1</sup> Tamanho em resíduos de aminoácidos.

<sup>2</sup> Este parâmetro foi determinado pela homologia com a sequência de aminoácidos de *Arabidopsis thaliana* equivalente.

<sup>3</sup> Possível aminoácido transferase (AAT). Não possui atividade de ACC sintase comprovada.

<sup>4</sup> A função primária desta proteína não está inicialmente associada a esta rota metabólica.

**Tabela S12.** Sumário das Proteínas de Soja Relacionadas com a Transdução de Sinal Mediada por Etileno

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio da Proteína de Soja			Suposta Função <sup>2</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Código PFAM	Descrição	Posição	
GmCPK#001	Glyma04g38150	496	AtCDPK2	At1g35670	OsCPK#001	Os11g07040	PF00069	Domínio proteína cinase	030 - 288	Ativação de ACS Tipo II
							PF00036	“EF hand”	335 - 396	
									401 - 466	
GmCPK#002	Glyma05g33240	507	AtCDPK2	At1g35670	OsCPK#001	Os11g07040	PF00069	Domínio proteína cinase	033 - 291	Ativação de ACS Tipo II
							PF00036	“EF hand”	338 - 399	
									404 - 469	
GmCPK#003	Glyma06g16920	497	AtCDPK2	At1g35670	OsCPK#001	Os11g07040	PF00069	Domínio proteína cinase	031 - 289	Ativação de ACS Tipo II
							PF00036	“EF hand”	336 - 397	
									402 - 467	
GmCPK#004	Glyma08g00840	508	AtCDPK2	At1g35670	OsCPK#001	Os11g07040	PF00069	Domínio proteína cinase	034 - 292	Ativação de ACS Tipo II
							PF00036	“EF hand”	339 - 400	
									405 - 470	
GmCPK#005	Glyma10g36090	482	AtCDPK2	At1g35670	OsCPK#002	Os12g07230	PF00069	Domínio proteína cinase	021 - 279	Ativação de ACS Tipo II
							PF00036	“EF hand”	326 - 387	
									392 - 457	
GmCTR#001	Glyma02g27680	660	AtCTR1	At5g03730	OsCTR#001	Os02g32610	PF14381	Proteína cinase responsiva à etileno Le-CTR1	064 - 273	Regulador Negativo de Resposta à Etileno
							PF07714	Proteína tirosina cinase	397 - 651	
GmCTR#002	Glyma03g34890	803	AtCTR1	At5g03730	OsCTR#001	Os02g32610	PF14381	Proteína cinase responsiva à etileno Le-CTR1	172 - 371	Regulador Negativo de Resposta à Etileno
							PF07714	Proteína tirosina cinase	529 - 783	
GmCTR#003	Glyma10g07610	793	AtCTR1	At5g03730	OsCTR#001	Os02g32610	PF14381	Proteína cinase responsiva à etileno Le-CTR1	197 - 396	Regulador Negativo de Resposta à Etileno
							PF07714	Proteína tirosina cinase	505 - 760	
GmCTR#004	Glyma13g21480	836	AtCTR1	At5g03730	OsCTR#001	Os02g32610	PF14381	Proteína cinase responsiva à etileno Le-CTR1	176 - 386	Regulador Negativo de Resposta à Etileno
							PF07714	Proteína tirosina cinase	562 - 816	
GmCTR#005	Glyma19g37570	803	AtCTR1	At5g03730	OsCTR#001	Os02g32610	PF14381	Proteína cinase responsiva à etileno Le-CTR1	167 - 376	Regulador Negativo de Resposta à Etileno
							PF07714	Proteína tirosina cinase	529 - 783	
GmEBF#001	Glyma04g07110	636	AtEBF1	At2g25490	OsEBF#001	Os02g10700	PF00646	Domínio F-box	065 - 113	Direciona EIN3 para Proteólise
GmEBF#002	Glyma04g20330	650	AtEBF1	At2g25490	OsEBF#001	Os02g10700	PF00646	Domínio F-box	068 - 106	Direciona EIN3 para Proteólise
GmEBF#003	Glyma06g07200	638	AtEBF1	At2g25490	OsEBF#001	Os02g10700	PF00646	Domínio F-box	067 - 114	Direciona EIN3 para Proteólise
GmEBF#004	Glyma13g23510	639	AtEBF1	At2g25490	OsEBF#001	Os02g10700	PF00646	Domínio F-box	065 - 103	Direciona EIN3 para Proteólise
GmEBF#005	Glyma14g14410	644	AtEBF1	At2g25490	OsEBF#001	Os02g10700	PF00646	Domínio F-box	065 - 111	Direciona EIN3 para Proteólise

**Tabela S12.** Sumário das Proteínas de Soja Relacionadas com a Transdução de Sinal Mediada por Etileno (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio da Proteína de Soja			Suposta Função <sup>2</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Código PFAM	Descrição	Posição	
GmEBF#006	Glyma17g12270	639	AtEBF1	At2g25490	OsEBF#001	Os02g10700	PF00646	Domínio F-box	065 - 103	Direciona EIN3 para Proteólise
GmEBF#007	Glyma17g31940	610	AtEBF1	At2g25490	OsEBF#001	Os02g10700	PF00646	Domínio F-box	067 - 106	Direciona EIN3 para Proteólise
GmEIN#001	Glyma02g44220	614	AtEIN3	At3g20770	OsEIN#002	Os03g20780	PF04873	Insensível à etileno 3	048 - 426	Fator de Transcrição Nuclear
GmEIN#002	Glyma03g33850	1281	AtEIN2	At5g03280	OsEIN#006	Os07g06130	PF01566	Proteína associada à resistência natural em macrófagos	038 - 392	Regulação após CTR1
GmEIN#003	Glyma03g40500	966	AtEIN5	At1g54490	OsEIN#005	Os03g58060	PF03159	N-terminal da exonuclease 5'-3' XRN	001 - 257	Silenciamento Gênico de EBF
							PF00098	Articulação de zinco	264 - 281	
GmEIN#004	Glyma10g06610	1298	AtEIN2	At5g03280	OsEIN#006	Os07g06130	PF01566	Proteína associada à resistência natural em macrófagos	038 - 392	Regulação após CTR1
GmEIN#005	Glyma13g03660	618	AtEIN3	At3g20770	OsEIN#003	Os03g20790	PF04873	Insensível à etileno 3	049 - 426	Fator de Transcrição Nuclear
GmEIN#006	Glyma13g03700	621	AtEIN3	At3g20770	OsEIN#002	Os03g20780	PF04873	Insensível à etileno 3	049 - 427	Fator de Transcrição Nuclear
GmEIN#007	Glyma13g20810	1334	AtEIN2	At5g03280	OsEIN#006	Os07g06130	PF01566	Proteína associada à resistência natural em macrófagos	038 - 392	Regulação após CTR1
GmEIN#008	Glyma14g04550	610	AtEIN3	At3g20770	OsEIN#003	Os03g20790	PF04873	Insensível à etileno 3	048 - 426	Fator de Transcrição Nuclear
GmEIN#009	Glyma19g43160	957	AtEIN5	At1g54490	OsEIN#005	Os03g58060	PF03159	N-terminal da exonuclease 5'-3' XRN	001 - 255	Silenciamento Gênico de EBF
							PF00098	Articulação de zinco	262 - 279	
GmEIN#010	Glyma20g12250	624	AtEIN3	At3g20770	OsEIN#002	Os03g20780	PF04873	Insensível à etileno 3	048 - 423	Fator de Transcrição Nuclear
GmEIN#011	Glyma20g37260	932	AtEIN5	At1g54490	OsEIN#005	Os03g58060	PF03159	N-terminal da exonuclease 5'-3' XRN	001 - 256	Silenciamento Gênico de EBF
							PF00098	Articulação de zinco	263 - 280	
GmETO#001	Glyma01g31170	954	AtETO1	At3g51770	OsETO#001	Os03g18360	PF13414	Repetição TPR	871 - 934	Regulador Negativo de ACS Tipo II
GmETO#002	Glyma02g16310	886	AtEOL1	At4g02680	OsETO#003	Os11g37520	PF00515	Repetição tetratricopeptídeo	709 - 742	Regulador Negativo de ACS Tipo II
GmETO#003	Glyma02g39460	937	AtETO1	At3g51770	OsETO#001	Os03g18360	PF07719	Repetição tetratricopeptídeo	560 - 593	Regulador Negativo de ACS Tipo II
GmETO#004	Glyma03g06600	960	AtETO1	At3g51770	OsETO#001	Os03g18360	PF13414	Repetição TPR	875 - 940	Regulador Negativo de ACS Tipo II
GmETO#005	Glyma03g31370	888	AtEOL1	At4g02680	OsETO#002	Os07g08120	PF13414	Repetição TPR	709 - 777	Regulador Negativo de ACS Tipo II
GmETO#006	Glyma10g03520	887	AtEOL1	At4g02680	OsETO#003	Os11g37520	PF00515	Repetição tetratricopeptídeo	710 - 743	Regulador Negativo de ACS Tipo II
GmETO#007	Glyma11g29000	935	AtETO1	At3g51770	OsETO#001	Os03g18360	PF13414	Repetição TPR	853 - 918	Regulador Negativo de ACS Tipo II
GmETO#008	Glyma14g37640	924	AtETO1	At3g51770	OsETO#001	Os03g18360	PF13414	Repetição TPR	745 - 810	Regulador Negativo de ACS Tipo II
GmETO#009	Glyma18g06680	873	AtETO1	At3g51770	OsETO#001	Os03g18360	PF13414	Repetição TPR	742 - 807	Regulador Negativo de ACS Tipo II
GmETO#010	Glyma19g34200	886	AtEOL1	At4g02680	OsETO#003	Os11g37520	PF13414	Repetição TPR	707 - 772	Regulador Negativo de ACS Tipo II
GmETR#001	Glyma03g37470	636	AtERS1	At2g40940	OsETR#002	Os03g49500	PF01590	Domínio GAF	159 - 308	Receptor de Etileno
							PF00512	Domínio histidina cinase A (fosfoceptor)	344 - 409	
							PF02518	Histidina cinase, DNA girase B e semelhante à ATPase	456 - 587	

**Tabela S12.** Sumário das Proteínas de Soja Relacionadas com a Transdução de Sinal Mediada por Etileno (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio da Proteína de Soja			Suposta Função <sup>2</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Código PFAM	Descrição	Posição	
GmETR#002	Glyma03g41220	760	AtEIN4	At3g04580	OsETR#003	Os04g08740	PF01590	Domínio GAF	183 - 334	Receptor de Etileno
							PF00512	Domínio histidina cinase A (fosfoceptor)	370 - 435	
							PF00072	Domínio regulador de resposta de receptor	636 - 750	
GmETR#003	Glyma09g00490	740	AtETR1	At1g66340	OsETR#002	Os03g49500	PF01590	Domínio GAF	158 - 307	Receptor de Etileno
							PF00512	Domínio histidina cinase A (fosfoceptor)	343 - 408	
							PF02518	Histidina cinase <sup>-</sup> , DNA girase B <sup>-</sup> e semelhante à ATPase	455 - 586	
							PF00072	Domínio regulador de resposta de receptor	616 - 728	
GmETR#004	Glyma10g31040	767	AtEIN4	At3g04580	OsETR#003	Os04g08740	PF01590	Domínio GAF	192 - 341	Receptor de Etileno
							PF00512	Domínio histidina cinase A (fosfoceptor)	377 - 442	
							PF00072	Domínio regulador de resposta de receptor	643 - 757	
GmETR#005	Glyma10g33240	751	AtETR2	At3g23150	OsETR#003	Os04g08740	PF01590	Domínio GAF	182 - 331	Receptor de Etileno
							PF00512	Domínio histidina cinase A (fosfoceptor)	367 - 432	
							PF00072	Domínio regulador de resposta de receptor	627 - 741	
GmETR#006	Glyma12g37050	739	AtETR1	At1g66340	OsETR#002	Os03g49500	PF01590	Domínio GAF	158 - 307	Receptor de Etileno
							PF00512	Domínio histidina cinase A (fosfoceptor)	343 - 408	
							PF02518	Histidina cinase <sup>-</sup> , DNA girase B <sup>-</sup> e semelhante à ATPase	455 - 585	
							PF00072	Domínio regulador de resposta de receptor	615 - 727	
GmETR#007	Glyma19g40090	636	AtERS1	At2g40940	OsETR#002	Os03g49500	PF01590	Domínio GAF	159 - 308	Receptor de Etileno
							PF00512	Domínio histidina cinase A (fosfoceptor)	344 - 409	
							PF02518	Histidina cinase <sup>-</sup> , DNA girase B <sup>-</sup> e semelhante à ATPase	456 - 587	
GmETR#008	Glyma19g43840	731	AtEIN4	At3g04580	OsETR#003	Os04g08740	PF01590	Domínio GAF	183 - 333	Receptor de Etileno
							PF00512	Domínio histidina cinase A (fosfoceptor)	369 - 434	
							PF00072	Domínio regulador de resposta de receptor	635 - 713	
GmETR#009	Glyma20g21780	682	AtETR2	At3g23150	OsETR#003	Os04g08740	PF01590	Domínio GAF	186 - 332	Receptor de Etileno
							PF00512	Domínio histidina cinase A (fosfoceptor)	368 - 433	
							PF00072	Domínio regulador de resposta de receptor	559 - 672	
GmETR#010	Glyma20g34420	798	AtETR2	At3g23150	OsETR#003	Os04g08740	PF01590	Domínio GAF	182 - 331	Receptor de Etileno
							PF00512	Domínio histidina cinase A (fosfoceptor)	367 - 432	
							PF00072	Domínio regulador de resposta de receptor	674 - 788	

**Tabela S12.** Sumário das Proteínas de Soja Relacionadas com a Transdução de Sinal Mediada por Etileno (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio da Proteína de Soja			Suposta Função <sup>2</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Código PFAM	Descrição	Posição	
GmETR#011	Glyma20g36440	734	AtEIN4	At3g04580	OsETR#003	Os04g08740	PF01590	Domínio GAF	159 - 308	Receptor de Etileno
							PF00512	Domínio histidina cinase A (fosfoceptor)	344 - 409	
							PF00072	Domínio regulador de resposta de receptor	610 - 724	
GmMKK#001	Glyma07g00520	351	AtMKK4	At1g51660	OsMKK#005	Os06g09180	PF00069	Domínio proteína cinase	069 - 324	Ativação de ACS Tipo I
GmMKK#002	Glyma07g11910	318	AtMKK9	At1g73500	OsMKK#003	Os03g12390	PF00069	Domínio proteína cinase	049 - 311	Ativação de EIN3
GmMKK#003	Glyma08g23900	364	AtMKK4	At1g51660	OsMKK#005	Os06g09180	PF00069	Domínio proteína cinase	082 - 337	Ativação de ACS Tipo I
GmMKK#004	Glyma09g30300	319	AtMKK9	At1g73500	OsMKK#003	Os03g12390	PF00069	Domínio proteína cinase	050 - 312	Ativação de EIN3
GmMPK#001	Glyma02g15690	391	AtMPK6	At2g43790	OsMPK#002	Os06g06090	PF00069	Domínio proteína cinase	061 - 344	Ativação de EIN3 e ACS Tipo I
GmMPK#002	Glyma07g32750	433	AtMPK6	At2g43790	OsMPK#002	Os06g06090	PF00069	Domínio proteína cinase	103 - 386	Ativação de EIN3 e ACS Tipo I
GmMPK#003	Glyma11g15700	371	AtMPK3	At3g45640	OsMPK#001	Os03g17700	PF00069	Domínio proteína cinase	041 - 325	Ativação de EIN3
GmMPK#004	Glyma12g07770	371	AtMPK3	At3g45640	OsMPK#001	Os03g17700	PF00069	Domínio proteína cinase	041 - 325	Ativação de EIN3
GmRAN#001	Glyma04g07350	221	AtRAN3	At5g55190	OsRAN#001	Os05g49890	PF00071	Família ras	015 - 173	Transporte de Cobre
GmRAN#002	Glyma04g07360	221	AtRAN3	At5g55190	OsRAN#001	Os05g49890	PF00071	Família ras	015 - 173	Transporte de Cobre
GmRAN#003	Glyma04g07370	221	AtRAN3	At5g55190	OsRAN#001	Os05g49890	PF00071	Família ras	015 - 173	Transporte de Cobre
GmRAN#004	Glyma05g08260	221	AtRAN3	At5g55190	OsRAN#001	Os05g49890	PF00071	Família ras	015 - 173	Transporte de Cobre
GmRAN#005	Glyma06g07400	221	AtRAN3	At5g55190	OsRAN#001	Os05g49890	PF00071	Família ras	015 - 173	Transporte de Cobre
GmRAN#006	Glyma06g07410	221	AtRAN3	At5g55190	OsRAN#001	Os05g49890	PF00071	Família ras	015 - 173	Transporte de Cobre
GmRAN#007	Glyma06g07420	221	AtRAN3	At5g55190	OsRAN#001	Os05g49890	PF00071	Família ras	015 - 173	Transporte de Cobre
GmRTE#001	Glyma01g37560	231	AtRTH1	At3g51040	OsRTE#002	Os03g58520	PF05608	Proteína sem função conhecida (DUF778)	040 - 175	Regulador Positivo do Receptor de Etileno
GmRTE#002	Glyma11g07740	230	AtRTH1	At3g51040	OsRTE#002	Os03g58520	PF05608	Proteína sem função conhecida (DUF778)	039 - 174	Regulador Positivo do Receptor de Etileno
GmRTE#003	Glyma14g10770	235	AtRTE1	At2g26070	OsRTE#001	Os01g51430	PF05608	Proteína sem função conhecida (DUF778)	044 - 179	Regulador Positivo do Receptor de Etileno
GmRTE#004	Glyma17g34740	235	AtRTE1	At2g26070	OsRTE#001	Os01g51430	PF05608	Proteína sem função conhecida (DUF778)	044 - 179	Regulador Positivo do Receptor de Etileno

<sup>1</sup> Tamanho em resíduos de aminoácidos.

<sup>2</sup> Este parâmetro foi determinado pela homologia com a sequência de aminoácidos de *Arabidopsis thaliana* equivalente.

**Tabela S13.** Sumário das Proteínas AP2 de Soja

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio AP2 de Soja		Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>3</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Nº <sup>2</sup>	Posição	
GmAP2#001	Glyma01g02760	507	AtAP2#013	At5g10510	OsAP2#009	Os04g55970	2	223 - 280 322 - 374	Semelhante à “Aintegumenta” 6 (AIL6); “Plethora” 3 (PLT3)
GmAP2#002	Glyma01g39520	500	AtAP2#011	At4g36920	OsAP2#008	Os04g55560	2	149 - 199 241 - 292	Mutante Floral 2 (FLO2; AP2; FL1; ATAP2)
GmAP2#003	Glyma01g40380	507	AtAP2#012	At4g37750	OsAP2#005	Os03g56050	2	166 - 226 268 - 323	Complementa Mutante da Proteína Cinase C 1 (CKC1); “Aintegumenta” (ANT)
GmAP2#004	Glyma02g09600	416	AtAP2#005	At2g28550	OsAP2#006	Os03g60430	1	161 - 211	Alvo Marcado de Início de Ativação 1 (TOE1); RAP2.7
GmAP2#005	Glyma02g33090	447	AtAP2#007	At2g41710	OsAP2#011	Os05g32270	1	090 - 150	Proteína da Superfamília de um Tipo de Integrase que se Liga ao DNA
GmAP2#006	Glyma02g36880	372	AtAP2#001	At1g16060	OsAP2#019	Os12g03290	2	048 - 105 147 - 185	Proteína com Duplo Domínio AP2 que Interage com ARIA (ADAP); “Wrinkled” 3 (WRI3)
GmAP2#007	Glyma03g29240	420	AtAP2#007	At2g41710	OsAP2#011	Os05g32270	1	061 - 118	Proteína da Superfamília de um Tipo de Integrase que se Liga ao DNA
GmAP2#008	Glyma03g33470	459	AtAP2#005	At2g28550	OsAP2#016	Os07g13170	2	124 - 174 216 - 267	Alvo Marcado de Início de Ativação 1 (TOE1); RAP2.7
GmAP2#009	Glyma04g05080	518	AtAP2#012	At4g37750	OsAP2#005	Os03g56050	2	155 - 212 254 - 306	Complementa Mutante da Proteína Cinase C 1 (CKC1); “Aintegumenta” (ANT)
GmAP2#010	Glyma04g24010	253	AtAP2#004	At1g79700	OsAP2#013	Os06g05340	1	064 - 116	“Wrinkled” 4 (WRI4)
GmAP2#011	Glyma05g18170	368	AtAP2#011	At4g36920	OsAP2#008	Os04g55560	2	013 - 063 105 - 156	Mutante Floral 2 (FLO2; AP2; FL1; ATAP2)
GmAP2#012	Glyma05g22970	612	AtAP2#012	At4g37750	OsAP2#005	Os03g56050	2	265 - 325 367 - 420	Complementa Mutante da Proteína Cinase C 1 (CKC1); “Aintegumenta” (ANT)
GmAP2#013	Glyma06g05170	546	AtAP2#012	At4g37750	OsAP2#005	Os03g56050	2	266 - 323 365 - 417	Complementa Mutante da Proteína Cinase C 1 (CKC1); “Aintegumenta” (ANT)
GmAP2#014	Glyma06g30840	286	AtAP2#004	At1g79700	OsAP2#013	Os06g05340	1	087 - 139	“Wrinkled” 4 (WRI4)
GmAP2#015	Glyma07g02380	287	AtAP2#001	At1g16060	OsAP2#013	Os06g05340	2	030 - 087 129 - 181	Proteína com Duplo Domínio AP2 que Interage com ARIA (ADAP); “Wrinkled” 3 (WRI3)
GmAP2#016	Glyma07g04260	344	AtAP2#003	At1g72570	OsAP2#005	Os03g56050	1	234 - 291	Proteína da Superfamília de um Tipo de Integrase que se Liga ao DNA
GmAP2#017	Glyma08g23630	296	AtAP2#001	At1g16060	OsAP2#013	Os06g05340	2	036 - 093 135 - 187	Proteína com Duplo Domínio AP2 que Interage com ARIA (ADAP); “Wrinkled” 3 (WRI3)
GmAP2#018	Glyma08g24420	423	AtAP2#009	At3g54320	OsAP2#019	Os12g03290	2	055 - 112 154 - 206	“Wrinkled” 1 (WRI1)

**Tabela S13.** Sumário das Proteínas AP2 de Soja (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio AP2 de Soja		Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>3</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Nº <sup>2</sup>	Posição	
GmAP2#019	Glyma08g38190	400	AtAP2#013	At5g10510	OsAP2#002	Os02g40070	2	163 - 211 218 - 270	Semelhante à “Aintegumenta” 6 (AIL6); “Plethora” 3 (PLT3)
GmAP2#020	Glyma08g40830	190	AtAP2#009	At3g54320	OsAP2#013	Os06g05340	2	039 - 094 136 - 188	“Wrinkled” 1 (WRI1)
GmAP2#021	Glyma09g37540	408	AtAP2#009	At3g54320	OsAP2#013	Os06g05340	2	049 - 106 148 - 200	“Wrinkled” 1 (WRI1)
GmAP2#022	Glyma09g38370	536	AtAP2#014	At5g17430	OsAP2#002	Os02g40070	2	151 - 189 208 - 248	“Baby Boom” (BBM)
GmAP2#023	Glyma10g22390	398	AtAP2#005	At2g28550	OsAP2#006	Os03g60430	1	156 - 206	Alvo Marcado de Início de Ativação 1 (TOE1); RAP2.7
GmAP2#024	Glyma10g31440	514	AtAP2#014	At5g17430	OsAP2#002	Os02g40070	2	169 - 226 269 - 321	“Baby Boom” (BBM)
GmAP2#025	Glyma11g04910	515	AtAP2#012	At4g37750	OsAP2#005	Os03g56050	2	240 - 297 339 - 392	Complementa Mutante da Proteína Cinase C 1 (CKC1); “Aintegumenta” (ANT)
GmAP2#026	Glyma11g05720	494	AtAP2#011	At4g36920	OsAP2#008	Os04g55560	1	236 - 287	Mutante Floral 2 (FLO2; AP2; FL1; ATAP2)
GmAP2#027	Glyma11g14040	562	AtAP2#008	At3g20840	OsAP2#002	Os02g40070	2	168 - 228 270 - 322	“Plethora” 1 (PLT1)
GmAP2#028	Glyma11g15650	372	AtAP2#005	At2g28550	OsAP2#010	Os05g03040	2	160 - 210 252 - 303	Alvo Marcado de Início de Ativação 1 (TOE1); RAP2.7
GmAP2#029	Glyma12g06010	553	AtAP2#002	At1g51190	OsAP2#007	Os04g42570	2	163 - 223 265 - 317	“Plethora” 2 (PLT2)
GmAP2#030	Glyma12g07800	478	AtAP2#005	At2g28550	OsAP2#010	Os05g03040	1	156 - 206	Alvo Marcado de Início de Ativação 1 (TOE1); RAP2.7
GmAP2#031	Glyma13g00950	528	AtAP2#015	At5g57390	OsAP2#009	Os04g55970	2	146 - 206 248 - 300	Semelhante à “Aintegumenta” 6 (AIL5); “Plethora” 5 (PLT5); “Chotto” 1 (CHO1)
GmAP2#032	Glyma13g40470	476	AtAP2#005	At2g28550	OsAP2#010	Os05g03040	2	136 - 186 228 - 279	Alvo Marcado de Início de Ativação 1 (TOE1); RAP2.7
GmAP2#033	Glyma14g10130	610	AtAP2#012	At4g37750	OsAP2#005	Os03g56050	2	317 - 374 416 - 468	Complementa Mutante da Proteína Cinase C 1 (CKC1); “Aintegumenta” (ANT)
GmAP2#034	Glyma15g04930	485	AtAP2#005	At2g28550	OsAP2#010	Os05g03040	2	136 - 186 228 - 279	Alvo Marcado de Início de Ativação 1 (TOE1); RAP2.7



**Tabela S13.** Sumário das Proteínas AP2 de Soja (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio AP2 de Soja		Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>3</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Nº <sup>2</sup>	Posição	
GmAP2#035	Glyma15g34770	409	AtAP2#009	At3g54320	OsAP2#019	Os12g03290	2	053 - 110 152 - 204	“Wrinkled” 1 (WRI1)
GmAP2#036	Glyma16g00950	512	AtAP2#003	At1g72570	OsAP2#005	Os03g56050	2	204 - 261 303 - 355	Proteína da Superfamília de um Tipo de Integrase que se Liga ao DNA
GmAP2#037	Glyma17g07010	530	AtAP2#015	At5g57390	OsAP2#001	Os01g67410	2	150 - 210 252 - 304	Semelhante à “Aintegumenta” 6 (AIL5); “Plethora” 5 (PLT5); “Chotto” 1 (CHO1)
GmAP2#038	Glyma17g07860	350	AtAP2#004	At1g79700	OsAP2#019	Os12g03290	2	049 - 106 148 - 200	“Wrinkled” 4 (WRI4)
GmAP2#039	Glyma17g17010	641	AtAP2#012	At4g37750	OsAP2#005	Os03g56050	2	293 - 350 392 - 445	Complementa Mutante da Proteína Cinase C 1 (CKC1); “Aintegumenta” (ANT)
GmAP2#040	Glyma17g18640	531	AtAP2#011	At4g36920	OsAP2#008	Os04g55560	2	176 - 226 268 - 319	Mutante Floral 2 (FLO2; AP2; FL1; ATAP2)
GmAP2#041	Glyma18g16240	371	AtAP2#009	At3g54320	OsAP2#013	Os06g05340	2	075 - 126 168 - 220	“Wrinkled” 1 (WRI1)
GmAP2#042	Glyma18g29400	245	AtAP2#017	At5g65510	OsAP2#001	Os01g67410	1	038 - 073	Semelhante à “Aintegumenta” 6 (AIL7); “Plethora” 7 (PLT7)
GmAP2#043	Glyma18g47980	616	AtAP2#014	At5g17430	OsAP2#002	Os02g40070	2	214 - 271 313 - 365	“Baby Boom” (BBM)
GmAP2#044	Glyma18g49100	392	AtAP2#009	At3g54320	OsAP2#013	Os06g05340	2	071 - 131 173 - 225	“Wrinkled” 1 (WRI1)
GmAP2#045	Glyma19g31960	413	AtAP2#007	At2g41710	OsAP2#011	Os05g32270	1	060 - 117	Proteína da Superfamília de um Tipo de Integrase que se Liga ao DNA
GmAP2#046	Glyma19g36200	458	AtAP2#005	At2g28550	OsAP2#016	Os07g13170	2	123 - 173 215 - 266	Alvo Marcado de Início de Ativação 1 (TOE1); RAP2.7

<sup>1</sup> Tamanho em resíduos de aminoácidos.

<sup>2</sup> Número de domínios AP2 (PF00847) presentes na sequência de aminoácidos em questão.

<sup>3</sup> Este parâmetro foi determinado pela homologia com a sequência de aminoácidos de *Arabidopsis thaliana* equivalente.

**Tabela S14.** Sumário dos Fatores Transcricionais Responsivos à Etileno (ERF) de Soja

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio AP2 de Soja <sup>2</sup>			Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>5</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Posição	C-I <sup>3</sup>	C-II <sup>4</sup>	
GmERF#001	Glyma01g03110	401	AtERF#090	At1g24590	OsERF#079	Os02g38090	050 - 107	B1	VIII	Regeneração de “Tiros” em Cultura de Tecidos (DRNL)
GmERF#002	Glyma01g13410	263	AtERF#034	At2g44940	OsERF#036	Os10g41130	072 - 130	A4	III	Subfamília DREB A4
GmERF#003	Glyma01g35010	214	AtERF#041	At5g11590	OsERF#038	Os02g13710	061 - 119	A4	III	Proteína com Elemento de Interação Responsivo à Desidratação (TINY2)
GmERF#004	Glyma01g39540	168	AtERF#010	At5g67190	OsERF#008	Os04g55520	026 - 084	A5	II	Proteína com Motivos DREB e EAR 2 (DEAR2)
GmERF#005	Glyma01g41520	274	AtERF#100	At4g17500	OsERF#091	Os02g43790	142 - 200	B3	IX	Subfamília ERF B3 (ERF1A)
GmERF#006	Glyma01g42500	226	AtERF#033	At1g12610	OsERF#025	Os02g45450	062 - 119	A1	III	Floração Anã e Atrasada (DDF1)
GmERF#007	Glyma01g42510	183	AtERF#025	At5g52020	OsERF#020	Os02g45420	014 - 070	A4	III	Subfamília DREB A4
GmERF#008	Glyma01g43350	274	AtERF#066	At4g27950	OsERF#055	Os06g06540	106 - 163	B5	VI	Fator de Resposta à Citocinina 4 (CRF4)
GmERF#009	Glyma01g43450	314	AtERF#061	At1g64380	OsERF#052	Os05g49700	116 - 173	A6	I	Subfamília DREB A6
GmERF#010	Glyma01g44130	213	AtERF#022	At1g33760	OsERF#019	Os11g13840	026 - 084	A4	III	Subfamília DREB A4
GmERF#011	Glyma01g44140	170	AtERF#021	At1g71450	OsERF#019	Os11g13840	001 - 059	A4	III	Subfamília DREB A4
GmERF#012	Glyma01g44230	178	AtERF#020	At1g71520	OsERF#017	Os01g66270	009 - 068	A5	II	Subfamília DREB A5
GmERF#013	Glyma02g00870	203	AtERF#092	At3g23240	OsERF#083	Os03g64260	062 - 120	B3	IX	Subfamília ERF B3 (ERF1)
GmERF#014	Glyma02g00890	157	AtERF#098	At3g23230	OsERF#090	Os08g44960	021 - 079	B3	IX	Regulador Transcricional de Resposta à Defesa 1 (TDR1)
GmERF#015	Glyma02g01960	300	AtERF#072	At3g16770	OsERF#068	Os01g21120	074 - 131	B2	VII	Subfamília ERF B2 (RAP2.3)
GmERF#016	Glyma02g04460	326	AtERF#090	At1g24590	OsERF#079	Os02g38090	051 - 108	B1	VIII	Regeneração de “Tiros” em Cultura de Tecidos (DRNL)
GmERF#017	Glyma02g07310	228	AtERF#114	At5g61890	OsERF#101	Os04g32620	043 - 100	B4	X	Subfamília ERF B4
GmERF#018	Glyma02g07460	262	AtERF#085	At5g13910	OsERF#082	Os04g32790	031 - 088	B1	VIII	Peciolo Frondoso (LEP)
GmERF#019	Glyma02g08020	309	AtERF#002	At5g19790	OsERF#006	Os07g38750	032 - 089	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#020	Glyma02g08840	370	AtERF#110	At5g50080	OsERF#099	Os01g64790	201 - 258	B4	X	Subfamília ERF B4 (ERF110)
GmERF#021	Glyma02g14940	215	AtERF#112	At2g33710	OsERF#099	Os01g64790	088 - 145	B4	X	Subfamília ERF B4
GmERF#022	Glyma02g31350	283	AtERF#051	At3g57600	OsERF#041	Os03g07830	023 - 081	A2	IV	Subfamília DREB A2
GmERF#023	Glyma02g40320	282	AtERF#066	At4g27950	OsERF#055	Os06g06540	126 - 183	B5	VI	Fator de Resposta à Citocinina 4 (CRF4)
GmERF#024	Glyma02g42960	392	AtERF#048	At2g40340	OsERF#042	Os05g27930	078 - 136	A2	IV	Proteína com Elemento de Interação Responsivo à Desidratação 2C (DREB2C)
GmERF#025	Glyma02g43240	273	AtERF#052	At2g40220	OsERF#117	Os05g28350	037 - 095	A3	IV	Insensível à ABA 4 (ABI4)
GmERF#026	Glyma02g43500	215	AtERF#082	At1g50640	OsERF#134	Os02g09650	026 - 083	B1	VIII	Subfamília ERF B1 (ERF3)
GmERF#027	Glyma02g46340	222	AtERF#083	At3g20310	OsERF#121	Os06g47590	025 - 082	B1	VIII	Subfamília ERF B1 (ERF7)
GmERF#028	Glyma03g23330	283	AtERF#066	At4g27950	OsERF#053	Os01g12440	099 - 156	B5	VI	Fator de Resposta à Citocinina 4 (CRF4)
GmERF#029	Glyma03g26310	195	AtERF#099	At2g44840	OsERF#091	Os02g43790	120 - 178	B3	IX	Subfamília ERF B3 (ERF13)
GmERF#030	Glyma03g26390	158	AtERF#099	At2g44840	OsERF#095	Os02g43820	033 - 091	B3	IX	Subfamília ERF B3 (ERF13)

**Tabela S14.** Sumário dos Fatores Transcricionais Responsivos à Etileno (ERF) de Soja (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio AP2 de Soja <sup>2</sup>			Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>5</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Posição	C-I <sup>3</sup>	C-II <sup>4</sup>	
GmERF#031	Glyma03g26450	152	AtERF#099	At2g44840	OsERF#093	Os04g46220	041 - 099	B3	IX	Subfamília ERF B3 (ERF13)
GmERF#032	Glyma03g26480	182	AtERF#099	At2g44840	OsERF#093	Os04g46220	063 - 121	B3	IX	Subfamília ERF B3 (ERF13)
GmERF#033	Glyma03g26520	223	AtERF#099	At2g44840	OsERF#091	Os02g43790	085 - 143	B3	IX	Subfamília ERF B3 (ERF13)
GmERF#034	Glyma03g26530	151	AtERF#099	At2g44840	OsERF#093	Os04g46220	080 - 137	B3	IX	Subfamília ERF B3 (ERF13)
GmERF#035	Glyma03g27050	287	AtERF#034	At2g44940	OsERF#034	Os04g46440	113 - 171	A4	III	Subfamília DREB A4
GmERF#036	Glyma03g29530	284	AtERF#051	At3g57600	OsERF#041	Os03g07830	028 - 086	A2	IV	Subfamília DREB A2
GmERF#037	Glyma03g31640	172	AtERF#082	At1g50640	OsERF#077	Os04g52090	021 - 078	B1	VIII	Subfamília ERF B1 (ERF3)
GmERF#038	Glyma03g31920	231	AtERF#093	At2g31230	OsERF#083	Os03g64260	086 - 144	B3	IX	Subfamília ERF B3 (ERF15)
GmERF#039	Glyma03g31930	153	AtERF#098	At3g23230	OsERF#090	Os08g44960	018 - 076	B3	IX	Regulador Transcricional de Resposta à Defesa 1 (TDR1)
GmERF#040	Glyma03g31940	242	AtERF#093	At2g31230	OsERF#087	Os09g39850	106 - 164	B3	IX	Subfamília ERF B3 (ERF15)
GmERF#041	Glyma03g34970	188	AtERF#024	At2g36450	OsERF#018	Os10g38000	022 - 079	A4	III	Tolerância à Sal e Seca (HARDY)
GmERF#042	Glyma03g41640	300	AtERF#067	At3g61630	OsERF#053	Os01g12440	100 - 157	B5	VI	Fator de Resposta à Citocinina 6 (CRF6)
GmERF#043	Glyma03g41910	184	AtERF#023	At1g01250	OsERF#033	Os04g46400	026 - 083	A4	III	Subfamília DREB A4
GmERF#044	Glyma03g42450	345	AtERF#074	At1g53910	OsERF#070	Os02g54160	099 - 156	B2	VII	Subfamília ERF B2 (RAP2.12)
GmERF#045	Glyma04g03070	214	AtERF#049	At1g75490	OsERF#044	Os08g45110	038 - 096	A2	IV	Subfamília DREB A2
GmERF#046	Glyma04g04350	160	AtERF#010	At5g67190	OsERF#008	Os04g55520	032 - 090	A5	II	Proteína com Motivos DREB e EAR 2 (DEAR2)
GmERF#047	Glyma04g06100	183	AtERF#018	At1g74930	OsERF#011	Os02g54050	001 - 058	A5	II	Subfamília DREB A5 (ORA47)
GmERF#048	Glyma04g06690	193	AtERF#003	At5g25190	OsERF#127	Os02g55380	007 - 064	B6	V	Induzível por Etileno e Sal 3 (ESE3)
GmERF#049	Glyma04g07140	173	AtERF#004	At5g11190	OsERF#001	Os06g40150	006 - 063	B6	V	Resposta de Tolerância à Seca (SHN2)
GmERF#050	Glyma04g08900	188	AtERF#041	At5g11590	OsERF#038	Os02g13710	027 - 085	A4	III	Proteína com Elemento de Interação Responsivo à Desidratação (TINY2)
GmERF#051	Glyma04g11290	314	AtERF#059	At1g78080	OsERF#049	Os02g51670	138 - 195	A6	I	Desdiferenciação Induzida por Ferimento 1 (WIND1); RAP2.4
GmERF#052	Glyma04g16700	102	AtERF#002	At5g19790	OsERF#005	Os07g10410	024 - 081	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#053	Glyma04g19650	218	AtERF#001	At1g15360	OsERF#003	Os02g10760	006 - 063	B6	V	Resposta de Tolerância à Seca (SHN1)
GmERF#054	Glyma04g21710	248	AtERF#066	At4g27950	OsERF#053	Os01g12440	038 - 095	B5	VI	Fator de Resposta à Citocinina 4 (CRF4)
GmERF#055	Glyma04g37870	174	AtERF#114	At5g61890	OsERF#101	Os04g32620	013 - 070	B4	X	Subfamília ERF B4
GmERF#056	Glyma04g37890	262	AtERF#114	At5g61890	OsERF#101	Os04g32620	086 - 143	B4	X	Subfamília ERF B4
GmERF#057	Glyma04g39510	281	AtERF#108	At1g43160	OsERF#118	Os11g06770	135 - 192	B4	X	Subfamília ERF B4 (RAP2.6)
GmERF#058	Glyma04g41740	324	AtERF#066	At4g27950	OsERF#054	Os01g46870	102 - 159	B5	VI	Fator de Resposta à Citocinina 4 (CRF4)
GmERF#059	Glyma04g43040	232	AtERF#012	At1g21910	OsERF#119	Os06g10780	036 - 093	A5	II	Proteína com Elemento de Interação Responsivo à Desidratação 26 (DREB26)
GmERF#060	Glyma05g03540	156	AtERF#025	At5g52020	OsERF#020	Os02g45420	015 - 071	A4	III	Subfamília DREB A4

**Tabela S14.** Sumário dos Fatores Transcricionais Responsivos à Etileno (ERF) de Soja (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio AP2 de Soja <sup>2</sup>			Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>5</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Posição	C-I <sup>3</sup>	C-II <sup>4</sup>	
GmERF#061	Glyma05g03560	211	AtERF#033	At1g12610	OsERF#025	Os02g45450	020 - 078	A1	III	Floração Anã e Atrasada (DDF1)
GmERF#062	Glyma05g04920	230	AtERF#034	At2g44940	OsERF#035	Os02g43970	057 - 115	A4	III	Subfamília DREB A4
GmERF#063	Glyma05g05130	278	AtERF#102	At5g47230	OsERF#097	Os04g46240	129 - 187	B3	IX	Subfamília ERF B3 (ERF5)
GmERF#064	Glyma05g05180	255	AtERF#100	At4g17500	OsERF#091	Os02g43790	137 - 195	B3	IX	Subfamília ERF B3 (ERF1A)
GmERF#065	Glyma05g07690	204	AtERF#002	At5g19790	OsERF#004	Os12g39330	049 - 106	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#066	Glyma05g19050	150	AtERF#010	At5g67190	OsERF#008	Os04g55520	022 - 080	A5	II	Proteína com Motivos DREB e EAR 2 (DEAR2)
GmERF#067	Glyma05g29010	141	AtERF#002	At5g19790	OsERF#005	Os07g10410	004 - 061	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#068	Glyma05g31370	312	AtERF#060	At4g39780	OsERF#120	Os06g11860	116 - 173	A6	I	Subfamília DREB A6
GmERF#069	Glyma05g32040	345	AtERF#111	At5g64750	OsERF#099	Os01g64790	165 - 222	B4	X	Repressor ABA 1 (ABR1)
GmERF#070	Glyma05g33440	125	AtERF#114	At5g61890	OsERF#101	Os04g32620	015 - 073	B4	X	Subfamília ERF B4
GmERF#071	Glyma05g35740	147	AtERF#036	At3g16280	OsERF#039	Os01g10370	021 - 079	A4	III	Subfamília DREB A4
GmERF#072	Glyma05g37120	334	AtERF#064	At4g23750	OsERF#054	Os01g46870	108 - 165	B5	VI	Alvo de Monópteros 3 (TMO3)
GmERF#073	Glyma06g03110	249	AtERF#049	At1g75490	OsERF#040	Os01g07120	038 - 096	A2	IV	Subfamília DREB A2
GmERF#074	Glyma06g04490	159	AtERF#010	At5g67190	OsERF#008	Os04g55520	032 - 090	A5	II	Proteína com Motivos DREB e EAR 2 (DEAR2)
GmERF#075	Glyma06g06100	234	AtERF#018	At1g74930	OsERF#012	Os08g35240	024 - 081	A5	II	Subfamília DREB A5 (ORA47)
GmERF#076	Glyma06g06780	194	AtERF#003	At5g25190	OsERF#127	Os02g55380	007 - 064	B6	V	Induzível por Etileno e Sal 3 (ESE3)
GmERF#077	Glyma06g07240	185	AtERF#004	At5g11190	OsERF#003	Os02g10760	006 - 063	B6	V	Resposta de Tolerância à Seca (SHN2)
GmERF#078	Glyma06g08990	194	AtERF#041	At5g11590	OsERF#122	Os06g36000	033 - 091	A4	III	Proteína com Elemento de Interação Responsivo à Desidratação (TINY2)
GmERF#079	Glyma06g11010	302	AtERF#059	At1g78080	OsERF#049	Os02g51670	127 - 184	A6	I	Desdiferenciação Induzida por Ferimento 1 (WIND1); RAP2.4
GmERF#080	Glyma06g11700	231	AtERF#012	At1g21910	OsERF#119	Os06g10780	033 - 090	A5	II	Proteína com Elemento de Interação Responsivo à Desidratação 26 (DREB26)
GmERF#081	Glyma06g13040	300	AtERF#066	At4g27950	OsERF#054	Os01g46870	101 - 158	B5	VI	Fator de Resposta à Citocinina 4 (CRF4)
GmERF#082	Glyma06g17180	239	AtERF#114	At5g61890	OsERF#105	Os05g36100	086 - 143	B4	X	Subfamília ERF B4
GmERF#083	Glyma06g35710	183	AtERF#078	At3g15210	OsERF#075	Os01g58420	026 - 083	B1	VIII	Subfamília ERF B1 (RAP2.5)
GmERF#084	Glyma06g40010	071	AtERF#080	At5g44210	OsERF#074	Os05g41780	002 - 059	B1	VIII	Subfamília ERF B1 (ERF9)
GmERF#085	Glyma06g44430	208	AtERF#080	At5g44210	OsERF#077	Os04g52090	023 - 080	B1	VIII	Subfamília ERF B1 (ERF9)
GmERF#086	Glyma06g45010	355	AtERF#053	At2g20880	OsERF#046	Os02g42580	207 - 264	A6	I	Fator de Transcrição Induzido por Seca (ERF53)
GmERF#087	Glyma06g45680	214	AtERF#045	At5g05410	OsERF#042	Os05g27930	065 - 123	A2	IV	Proteína com Elemento de Interação Responsivo à Desidratação 2A (DREB2A)
GmERF#088	Glyma07g02000	259	AtERF#041	At5g11590	OsERF#034	Os04g46440	033 - 091	A4	III	Proteína com Elemento de Interação Responsivo à Desidratação (TINY2)
GmERF#089	Glyma07g02930	194	AtERF#003	At5g25190	OsERF#127	Os02g55380	007 - 064	B6	V	Induzível por Etileno e Sal 3 (ESE3)
GmERF#090	Glyma07g03040	127	AtERF#084	At1g80580	OsERF#074	Os05g41780	017 - 072	B1	VIII	Subfamília ERF B1

**Tabela S14.** Sumário dos Fatores Transcricionais Responsivos à Etileno (ERF) de Soja (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio AP2 de Soja <sup>2</sup>			Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>5</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Posição	C-I <sup>3</sup>	C-II <sup>4</sup>	
GmERF#091	Glyma07g03500	189	AtERF#001	At1g15360	OsERF#003	Os02g10760	006 - 063	B6	V	Resposta de Tolerância à Seca (SHN1)
GmERF#092	Glyma07g04950	392	AtERF#074	At1g53910	OsERF#070	Os02g54160	118 - 175	B2	VII	Subfamília ERF B2 (RAP2.12)
GmERF#093	Glyma07g06080	191	AtERF#023	At1g01250	OsERF#033	Os04g46400	037 - 094	A4	III	Subfamília DREB A4
GmERF#094	Glyma07g13980	231	AtERF#099	At2g44840	OsERF#091	Os02g43790	089 - 147	B3	IX	Subfamília ERF B3 (ERF13)
GmERF#095	Glyma07g14060	205	AtERF#101	At5g47220	OsERF#091	Os02g43790	085 - 143	B3	IX	Subfamília ERF B3 (ERF2)
GmERF#096	Glyma07g14070	144	AtERF#099	At2g44840	OsERF#091	Os02g43790	081 - 139	B3	IX	Subfamília ERF B3 (ERF13)
GmERF#097	Glyma07g14560	258	AtERF#034	At2g44940	OsERF#033	Os04g46400	092 - 150	A4	III	Subfamília DREB A4
GmERF#098	Glyma07g23240	142	AtERF#080	At5g44210	OsERF#074	Os05g41780	023 - 080	B1	VIII	Subfamília ERF B1 (ERF9)
GmERF#099	Glyma07g33510	230	AtERF#109	At4g34410	OsERF#103	Os02g52670	096 - 153	B4	X	Fator de Transcrição de Resposta Redox 1 (RRTF1)
GmERF#100	Glyma07g37410	102	AtERF#072	At3g16770	OsERF#070	Os02g54160	017 - 074	B2	VII	Subfamília ERF B2 (RAP2.3)
GmERF#101	Glyma07g37990	297	AtERF#086	At5g18560	OsERF#078	Os07g47330	048 - 105	B1	VIII	Gene Induzido por Auxina (PUCHI)
GmERF#102	Glyma08g02460	293	AtERF#064	At4g23750	OsERF#054	Os01g46870	108 - 165	B5	VI	Alvo de Monópteros 3 (TMO3)
GmERF#103	Glyma08g03910	242	AtERF#041	At5g11590	OsERF#122	Os06g36000	045 - 103	A4	III	Proteína com Elemento de Interação Responsivo à Desidratação (TINY2)
GmERF#104	Glyma08g12130	239	AtERF#002	At5g19790	OsERF#005	Os07g10410	008 - 065	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#105	Glyma08g14600	312	AtERF#060	At4g39780	OsERF#049	Os02g51670	120 - 177	A6	I	Subfamília DREB A6
GmERF#106	Glyma08g15350	296	AtERF#114	At5g61890	OsERF#099	Os01g64790	159 - 216	B4	X	Subfamília ERF B4
GmERF#107	Glyma08g21650	251	AtERF#041	At5g11590	OsERF#034	Os04g46440	075 - 133	A4	III	Proteína com Elemento de Interação Responsivo à Desidratação (TINY2)
GmERF#108	Glyma08g22590	200	AtERF#001	At1g15360	OsERF#003	Os02g10760	006 - 063	B6	V	Resposta de Tolerância à Seca (SHN1)
GmERF#109	Glyma08g23070	130	AtERF#084	At1g80580	OsERF#074	Os05g41780	027 - 082	B1	VIII	Subfamília ERF B1
GmERF#110	Glyma08g23160	195	AtERF#003	At5g25190	OsERF#127	Os02g55380	007 - 064	B6	V	Induzível por Etileno e Sal 3 (ESE3)
GmERF#111	Glyma08g28820	190	AtERF#113	At5g13330	OsERF#107	Os02g32140	006 - 063	B4	X	Subfamília ERF B4 (RAP2.6L)
GmERF#112	Glyma08g38800	251	AtERF#090	At1g24590	OsERF#079	Os02g38090	047 - 104	B1	VIII	Regeneração de "Tiros" em Cultura de Tecidos (DRNL)
GmERF#113	Glyma08g43300	210	AtERF#082	At1g50640	OsERF#134	Os02g09650	027 - 084	B1	VIII	Subfamília ERF B1 (ERF3)
GmERF#114	Glyma09g04630	237	AtERF#072	At3g16770	OsERF#061	Os05g29810	084 - 141	B2	VII	Subfamília ERF B2 (RAP2.3)
GmERF#115	Glyma09g05840	132	AtERF#098	At3g23230	OsERF#089	Os10g30840	019 - 077	B3	IX	Regulador Transcricional de Resposta à Defesa 1 (TDR1)
GmERF#116	Glyma09g05850	122	AtERF#096	At5g43410	OsERF#128	Os04g18650	011 - 069	B3	IX	Subfamília ERF B3
GmERF#117	Glyma09g05860	137	AtERF#098	At3g23230	OsERF#089	Os10g30840	019 - 077	B3	IX	Regulador Transcricional de Resposta à Defesa 1 (TDR1)
GmERF#118	Glyma09g08330	214	AtERF#017	At1g19210	OsERF#012	Os08g35240	021 - 078	A5	II	Subfamília DREB A5
GmERF#119	Glyma09g27180	234	AtERF#028	At5g51990	OsERF#025	Os02g45450	064 - 121	A1	III	Proteína com Elemento de Interação Responsivo à Desidratação 1D (DREB1D)
GmERF#120	Glyma09g32730	227	AtERF#041	At5g11590	OsERF#122	Os06g36000	052 - 110	A4	III	Proteína com Elemento de Interação Responsivo à Desidratação (TINY2)

**Tabela S14.** Sumário dos Fatores Transcricionais Responsivos à Etileno (ERF) de Soja (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio AP2 de Soja <sup>2</sup>			Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>5</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Posição	C-I <sup>3</sup>	C-II <sup>4</sup>	
GmERF#121	Glyma09g36840	164	AtERF#041	At5g11590	OsERF#038	Os02g13710	015 - 072	A4	III	Proteína com Elemento de Interação Responsivo à Desidratação (TINY2)
GmERF#122	Glyma09g37780	203	AtERF#099	At2g44840	OsERF#091	Os02g43790	077 - 135	B3	IX	Subfamília ERF B3 (ERF13)
GmERF#123	Glyma10g00980	167	AtERF#092	At3g23240	OsERF#086	Os07g22770	046 - 104	B3	IX	Subfamília ERF B3 (ERF1)
GmERF#124	Glyma10g00990	124	AtERF#098	At3g23230	OsERF#123	Os09g39810	008 - 066	B3	IX	Regulador Transcricional de Resposta à Defesa 1 (TDR1)
GmERF#125	Glyma10g02080	304	AtERF#072	At3g16770	OsERF#068	Os01g21120	074 - 131	B2	VII	Subfamília ERF B2 (RAP2.3)
GmERF#126	Glyma10g04160	137	AtERF#098	At3g23230	OsERF#089	Os10g30840	016 - 074	B3	IX	Regulador Transcricional de Resposta à Defesa 1 (TDR1)
GmERF#127	Glyma10g04170	188	AtERF#092	At3g23240	OsERF#087	Os09g39850	056 - 114	B3	IX	Subfamília ERF B3 (ERF1)
GmERF#128	Glyma10g04190	158	AtERF#098	At3g23230	OsERF#096	Os10g41330	016 - 074	B3	IX	Regulador Transcricional de Resposta à Defesa 1 (TDR1)
GmERF#129	Glyma10g04210	270	AtERF#092	At3g23240	OsERF#083	Os03g64260	119 - 177	B3	IX	Subfamília ERF B3 (ERF1)
GmERF#130	Glyma10g07740	159	AtERF#024	At2g36450	OsERF#018	Os10g38000	001 - 057	A4	III	Tolerância à Sal e Seca (HARDY)
GmERF#131	Glyma10g21850	291	AtERF#051	At3g57600	OsERF#041	Os03g07830	023 - 081	A2	IV	Subfamília DREB A2
GmERF#132	Glyma10g23440	281	AtERF#105	At5g51190	OsERF#094	Os04g46250	125 - 183	B3	IX	Subfamília ERF B3 (ERF105)
GmERF#133	Glyma10g23460	220	AtERF#106	At5g07580	OsERF#096	Os10g41330	131 - 189	B3	IX	Subfamília ERF B3
GmERF#134	Glyma10g33060	219	AtERF#092	At3g23240	OsERF#083	Os03g64260	073 - 131	B3	IX	Subfamília ERF B3 (ERF1)
GmERF#135	Glyma10g33070	141	AtERF#098	At3g23230	OsERF#090	Os08g44960	016 - 074	B3	IX	Regulador Transcricional de Resposta à Defesa 1 (TDR1)
GmERF#136	Glyma10g33080	147	AtERF#095	At3g23220	OsERF#089	Os10g30840	018 - 076	B3	IX	Induzível por Etileno e Sal 1 (ESE1)
GmERF#137	Glyma10g33700	387	AtERF#062	At4g13620	OsERF#048	Os08g31580	221 - 278	A6	I	Subfamília DREB A6
GmERF#138	Glyma10g33810	201	AtERF#106	At5g07580	OsERF#095	Os02g43820	110 - 168	B3	IX	Subfamília ERF B3
GmERF#139	Glyma10g36300	135	AtERF#088	At1g12890	OsERF#080	Os08g07700	022 - 079	B1	VIII	Subfamília ERF B1
GmERF#140	Glyma10g36760	247	AtERF#110	At5g50080	OsERF#118	Os11g06770	048 - 105	B4	X	Subfamília ERF B4 (ERF110)
GmERF#141	Glyma10g38420	205	AtERF#025	At5g52020	OsERF#020	Os02g45420	043 - 100	A4	III	Subfamília DREB A4
GmERF#142	Glyma10g38440	185	AtERF#028	At5g51990	OsERF#025	Os02g45450	036 - 093	A1	III	Proteína com Elemento de Interação Responsivo à Desidratação 1D (DREB1D)
GmERF#143	Glyma10g42130	355	AtERF#118	At1g68550	OsERF#108	Os01g04020	120 - 178	B6	VI-L	Fator de Resposta à Citocinina 10 (CRF10)
GmERF#144	Glyma11g01640	169	AtERF#021	At1g71450	OsERF#019	Os11g13840	001 - 059	A4	III	Subfamília DREB A4
GmERF#145	Glyma11g01700	179	AtERF#020	At1g71520	OsERF#017	Os01g66270	016 - 075	A5	II	Subfamília DREB A5
GmERF#146	Glyma11g02050	325	AtERF#061	At1g64380	OsERF#052	Os05g49700	136 - 193	A6	I	Subfamília DREB A6
GmERF#147	Glyma11g02140	289	AtERF#064	At4g23750	OsERF#053	Os01g12440	112 - 169	B5	VI	Alvo de Monópteros 3 (TMO3)
GmERF#148	Glyma11g03790	184	AtERF#034	At2g44940	OsERF#034	Os04g46440	029 - 087	A4	III	Subfamília DREB A4
GmERF#149	Glyma11g03900	276	AtERF#102	At5g47230	OsERF#097	Os04g46240	138 - 196	B3	IX	Subfamília ERF B3 (ERF5)
GmERF#150	Glyma11g03910	240	AtERF#100	At4g17500	OsERF#091	Os02g43790	142 - 200	B3	IX	Subfamília ERF B3 (ERF1A)

**Tabela S14.** Sumário dos Fatores Transcricionais Responsivos à Etileno (ERF) de Soja (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio AP2 de Soja <sup>2</sup>			Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>5</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Posição	C-I <sup>3</sup>	C-II <sup>4</sup>	
GmERF#151	Glyma11g05700	153	AtERF#008	At2g23340	OsERF#008	Os04g55520	031 - 089	A5	II	Proteína com Motivos DREB e EAR 3 (DEAR3)
GmERF#152	Glyma11g31400	280	AtERF#066	At4g27950	OsERF#056	Os05g25260	130 - 187	B5	VI	Fator de Resposta à Citocinina 4 (CRF4)
GmERF#153	Glyma12g09130	216	AtERF#029	At4g25490	OsERF#027	Os01g73770	052 - 109	A1	III	Proteína com Elemento de Interação Responsivo à Desidratação 1B (DREB1B)
GmERF#154	Glyma12g11150	211	AtERF#045	At5g05410	OsERF#042	Os05g27930	065 - 123	A2	IV	Proteína com Elemento de Interação Responsivo à Desidratação 2A (DREB2A)
GmERF#155	Glyma12g13320	141	AtERF#080	At5g44210	OsERF#077	Os04g52090	022 - 079	B1	VIII	Subfamília ERF B1 (ERF9)
GmERF#156	Glyma12g30710	239	AtERF#030	At4g25470	OsERF#027	Os01g73770	063 - 119	A1	III	Repetição C/Fator de Interação DRE 2 (CBF2)
GmERF#157	Glyma12g30740	189	AtERF#030	At4g25470	OsERF#027	Os01g73770	015 - 072	A1	III	Repetição C/Fator de Interação DRE 2 (CBF2)
GmERF#158	Glyma12g32400	197	AtERF#045	At5g05410	OsERF#040	Os01g07120	063 - 121	A2	IV	Proteína com Elemento de Interação Responsivo à Desidratação 2A (DREB2A)
GmERF#159	Glyma12g33020	406	AtERF#053	At2g20880	OsERF#046	Os02g42580	207 - 264	A6	I	Fator de Transcrição Induzido por Seca (ERF53)
GmERF#160	Glyma12g35550	193	AtERF#078	At3g15210	OsERF#075	Os01g58420	026 - 083	B1	VIII	Subfamília ERF B1 (RAP2.5)
GmERF#161	Glyma13g01930	311	AtERF#059	At1g78080	OsERF#049	Os02g51670	137 - 194	A6	I	Desdiferenciação Induzida por Ferimento 1 (WIND1); RAP2.4
GmERF#162	Glyma13g05690	362	AtERF#002	At5g19790	OsERF#004	Os12g39330	017 - 074	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#163	Glyma13g08490	335	AtERF#066	At4g27950	OsERF#054	Os01g46870	107 - 164	B5	VI	Fator de Resposta à Citocinina 4 (CRF4)
GmERF#164	Glyma13g17250	199	AtERF#017	At1g19210	OsERF#011	Os02g54050	018 - 075	A5	II	Subfamília DREB A5
GmERF#165	Glyma13g18330	138	AtERF#098	At3g23230	OsERF#089	Os10g30840	017 - 075	B3	IX	Regulador Transcricional de Resposta à Defesa 1 (TDR1)
GmERF#166	Glyma13g18340	213	AtERF#092	At3g23240	OsERF#087	Os09g39850	075 - 133	B3	IX	Subfamília ERF B3 (ERF1)
GmERF#167	Glyma13g18350	213	AtERF#092	At3g23240	OsERF#087	Os09g39850	077 - 135	B3	IX	Subfamília ERF B3 (ERF1)
GmERF#168	Glyma13g18370	160	AtERF#092	At3g23240	OsERF#087	Os09g39850	031 - 089	B3	IX	Subfamília ERF B3 (ERF1)
GmERF#169	Glyma13g18390	172	AtERF#092	At3g23240	OsERF#087	Os09g39850	027 - 085	B3	IX	Subfamília ERF B3 (ERF1)
GmERF#170	Glyma13g18400	153	AtERF#098	At3g23230	OsERF#128	Os04g18650	016 - 074	B3	IX	Regulador Transcricional de Resposta à Defesa 1 (TDR1)
GmERF#171	Glyma13g18410	259	AtERF#092	At3g23240	OsERF#083	Os03g64260	112 - 170	B3	IX	Subfamília ERF B3 (ERF1)
GmERF#172	Glyma13g21560	159	AtERF#024	At2g36450	OsERF#018	Os10g38000	001 - 057	A4	III	Tolerância à Sal e Seca (HARDY)
GmERF#173	Glyma13g21570	190	AtERF#030	At4g25470	OsERF#027	Os01g73770	016 - 073	A1	III	Repetição C/Fator de Interação DRE 2 (CBF2)
GmERF#174	Glyma13g23570	238	AtERF#001	At1g15360	OsERF#003	Os02g10760	006 - 063	B6	V	Resposta de Tolerância à Seca (SHN1)
GmERF#175	Glyma13g28810	284	AtERF#086	At5g18560	OsERF#078	Os07g47330	047 - 104	B1	VIII	Gene Induzido por Auxina (PUCHI)
GmERF#176	Glyma13g29920	373	AtERF#002	At5g19790	OsERF#006	Os07g38750	029 - 086	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#177	Glyma13g30710	255	AtERF#091	At4g18450	OsERF#084	Os05g49010	113 - 171	B3	IX	Subfamília ERF B3
GmERF#178	Glyma13g30720	171	AtERF#101	At5g47220	OsERF#091	Os02g43790	053 - 111	B3	IX	Subfamília ERF B3 (ERF2)
GmERF#179	Glyma13g30990	222	AtERF#080	At5g44210	OsERF#075	Os01g58420	025 - 082	B1	VIII	Subfamília ERF B1 (ERF9)
GmERF#180	Glyma13g31010	163	AtERF#081	At1g28360	OsERF#074	Os05g41780	011 - 068	B1	VIII	Subfamília ERF B1 (ERF12)

**Tabela S14.** Sumário dos Fatores Transcricionais Responsivos à Etileno (ERF) de Soja (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio AP2 de Soja <sup>2</sup>			Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>5</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Posição	C-I <sup>3</sup>	C-II <sup>4</sup>	
GmERF#181	Glyma13g34920	193	AtERF#080	At5g44210	OsERF#075	Os01g58420	026 - 083	B1	VIII	Subfamília ERF B1 (ERF9)
GmERF#182	Glyma13g37450	277	AtERF#053	At2g20880	OsERF#046	Os02g42580	148 - 205	A6	I	Fator de Transcrição Induzido por Seca (ERF53)
GmERF#183	Glyma13g38030	198	AtERF#045	At5g05410	OsERF#040	Os01g07120	063 - 121	A2	IV	Proteína com Elemento de Interação Responsivo à Desidratação 2A (DREB2A)
GmERF#184	Glyma13g39540	193	AtERF#030	At4g25470	OsERF#027	Os01g73770	027 - 084	A1	III	Repetição C/Fator de Interação DRE 2 (CBF2)
GmERF#185	Glyma13g43210	211	AtERF#041	At5g11590	OsERF#034	Os04g46440	044 - 102	A4	III	Proteína com Elemento de Interação Responsivo à Desidratação (TINY2)
GmERF#186	Glyma13g44660	179	AtERF#003	At5g25190	OsERF#127	Os02g55380	007 - 064	B6	V	Induzível por Etileno e Sal 3 (ESE3)
GmERF#187	Glyma14g02360	222	AtERF#082	At1g50640	OsERF#076	Os04g57340	026 - 083	B1	VIII	Subfamília ERF B1 (ERF3)
GmERF#188	Glyma14g05470	212	AtERF#082	At1g50640	OsERF#134	Os02g09650	022 - 079	B1	VIII	Subfamília ERF B1 (ERF3)
GmERF#189	Glyma14g06080	393	AtERF#048	At2g40340	OsERF#042	Os05g27930	078 - 136	A2	IV	Proteína com Elemento de Interação Responsivo à Desidratação 2C (DREB2C)
GmERF#190	Glyma14g06290	309	AtERF#052	At2g40220	OsERF#117	Os05g28350	027 - 085	A3	IV	Insensível à ABA 4 (ABI4)
GmERF#191	Glyma14g07620	283	AtERF#049	At1g75490	OsERF#042	Os05g27930	033 - 091	A2	IV	Subfamília DREB A2
GmERF#192	Glyma14g09320	174	AtERF#006	At1g46768	OsERF#008	Os04g55520	043 - 101	A5	II	Subfamília DREB A5 (RAP2.1)
GmERF#193	Glyma14g13470	199	AtERF#003	At5g25190	OsERF#127	Os02g55380	007 - 064	B6	V	Induzível por Etileno e Sal 3 (ESE3)
GmERF#194	Glyma14g22740	244	AtERF#041	At5g11590	OsERF#122	Os06g36000	048 - 106	A4	III	Proteína com Elemento de Interação Responsivo à Desidratação (TINY2)
GmERF#195	Glyma14g22970	220	AtERF#013	At1g77640	OsERF#009	Os03g15660	009 - 066	A5	II	Subfamília DREB A5
GmERF#196	Glyma14g29040	321	AtERF#066	At4g27950	OsERF#054	Os01g46870	102 - 159	B5	VI	Fator de Resposta à Citocinina 4 (CRF4)
GmERF#197	Glyma14g32210	259	AtERF#051	At3g57600	OsERF#041	Os03g07830	017 - 075	A2	IV	Subfamília DREB A2
GmERF#198	Glyma14g34590	312	AtERF#059	At1g78080	OsERF#049	Os02g51670	150 - 207	A6	I	Desdiferenciação Induzida por Ferimento 1 (WIND1); RAP2.4
GmERF#199	Glyma14g38610	282	AtERF#066	At4g27950	OsERF#055	Os06g06540	125 - 182	B5	VI	Fator de Resposta à Citocinina 4 (CRF4)
GmERF#200	Glyma15g00660	194	AtERF#003	At5g25190	OsERF#127	Os02g55380	023 - 080	B6	V	Induzível por Etileno e Sal 3 (ESE3)
GmERF#201	Glyma15g01140	176	AtERF#001	At1g15360	OsERF#003	Os02g10760	006 - 063	B6	V	Resposta de Tolerância à Seca (SHN1)
GmERF#202	Glyma15g02130	215	AtERF#041	At5g11590	OsERF#034	Os04g46440	047 - 105	A4	III	Proteína com Elemento de Interação Responsivo à Desidratação (TINY2)
GmERF#203	Glyma15g02900	188	AtERF#024	At2g36450	OsERF#018	Os10g38000	022 - 079	A4	III	Tolerância à Sal e Seca (HARDY)
GmERF#204	Glyma15g08360	172	AtERF#081	At1g28360	OsERF#075	Os01g58420	014 - 071	B1	VIII	Subfamília ERF B1 (ERF12)
GmERF#205	Glyma15g08370	219	AtERF#080	At5g44210	OsERF#075	Os01g58420	022 - 079	B1	VIII	Subfamília ERF B1 (ERF9)
GmERF#206	Glyma15g08560	183	AtERF#101	At5g47220	OsERF#091	Os02g43790	074 - 132	B3	IX	Subfamília ERF B3 (ERF2)
GmERF#207	Glyma15g08580	253	AtERF#091	At4g18450	OsERF#084	Os05g49010	111 - 169	B3	IX	Subfamília ERF B3
GmERF#208	Glyma15g09190	362	AtERF#002	At5g19790	OsERF#006	Os07g38750	027 - 084	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#209	Glyma15g10250	233	AtERF#086	At5g18560	OsERF#078	Os07g47330	047 - 104	B1	VIII	Gene Induzido por Auxina (PUCHI)
GmERF#210	Glyma15g16260	223	AtERF#072	At3g16770	OsERF#070	Os02g54160	080 - 137	B2	VII	Subfamília ERF B2 (RAP2.3)



**Tabela S14.** Sumário dos Fatores Transcricionais Responsivos à Etileno (ERF) de Soja (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio AP2 de Soja <sup>2</sup>			Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>5</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Posição	C-I <sup>3</sup>	C-II <sup>4</sup>	
GmERF#211	Glyma15g17090	132	AtERF#098	At3g23230	OsERF#089	Os10g30840	019 - 077	B3	IX	Regulador Transcricional de Resposta à Defesa 1 (TDR1)
GmERF#212	Glyma15g17100	121	AtERF#098	At3g23230	OsERF#128	Os04g18650	011 - 069	B3	IX	Regulador Transcricional de Resposta à Defesa 1 (TDR1)
GmERF#213	Glyma15g19910	205	AtERF#016	At5g21960	OsERF#012	Os08g35240	013 - 070	A5	II	Subfamília DREB A5
GmERF#214	Glyma16g01500	382	AtERF#074	At1g53910	OsERF#072	Os09g26420	114 - 171	B2	VII	Subfamília ERF B2 (RAP2.12)
GmERF#215	Glyma16g02680	194	AtERF#023	At1g01250	OsERF#033	Os04g46400	037 - 094	A4	III	Subfamília DREB A4
GmERF#216	Glyma16g04410	273	AtERF#002	At5g19790	OsERF#005	Os07g10410	024 - 081	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#217	Glyma16g05070	191	AtERF#085	At5g13910	OsERF#082	Os04g32790	014 - 071	B1	VIII	Pecíolo Frondoso (LEP)
GmERF#218	Glyma16g05190	260	AtERF#114	At5g61890	OsERF#107	Os02g32140	048 - 105	B4	X	Subfamília ERF B4
GmERF#219	Glyma16g26320	239	AtERF#114	At5g61890	OsERF#101	Os04g32620	044 - 101	B4	X	Subfamília ERF B4
GmERF#220	Glyma16g26460	274	AtERF#085	At5g13910	OsERF#082	Os04g32790	031 - 088	B1	VIII	Pecíolo Frondoso (LEP)
GmERF#221	Glyma16g27040	315	AtERF#002	At5g19790	OsERF#005	Os07g10410	034 - 091	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#222	Glyma16g27950	414	AtERF#110	At5g50080	OsERF#105	Os05g36100	210 - 267	B4	X	Subfamília ERF B4 (ERF110)
GmERF#223	Glyma16g32330	231	AtERF#028	At5g51990	OsERF#025	Os02g45450	064 - 121	A1	III	Proteína com Elemento de Interação Responsivo à Desidratação 1D (DREB1D)
GmERF#224	Glyma17g02710	216	AtERF#086	At5g18560	OsERF#078	Os07g47330	048 - 105	B1	VIII	Gene Induzido por Auxina (PUCHI)
GmERF#225	Glyma17g05240	198	AtERF#017	At1g19210	OsERF#011	Os02g54050	017 - 074	A5	II	Subfamília DREB A5
GmERF#226	Glyma17g12330	239	AtERF#001	At1g15360	OsERF#003	Os02g10760	006 - 063	B6	V	Resposta de Tolerância à Seca (SHN1)
GmERF#227	Glyma17g13320	210	AtERF#002	At5g19790	OsERF#005	Os07g10410	051 - 108	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#228	Glyma17g14100	165	AtERF#025	At5g52020	OsERF#020	Os02g45420	012 - 068	A4	III	Subfamília DREB A4
GmERF#229	Glyma17g14110	170	AtERF#031	At4g25480	OsERF#025	Os02g45450	026 - 084	A1	III	Proteína com Elemento de Interação Responsivo à Desidratação 1A (DREB1A)
GmERF#230	Glyma17g15310	232	AtERF#034	At2g44940	OsERF#035	Os02g43970	060 - 118	A4	III	Subfamília DREB A4
GmERF#231	Glyma17g15460	275	AtERF#102	At5g47230	OsERF#097	Os04g46240	124 - 182	B3	IX	Subfamília ERF B3 (ERF5)
GmERF#232	Glyma17g15480	251	AtERF#100	At4g17500	OsERF#091	Os02g43790	133 - 191	B3	IX	Subfamília ERF B3 (ERF1A)
GmERF#233	Glyma17g16080	105	AtERF#002	At5g19790	OsERF#005	Os07g10410	004 - 061	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#234	Glyma17g18580	147	AtERF#008	At2g23340	OsERF#008	Os04g55520	023 - 080	A5	II	Proteína com Motivos DREB e EAR 3 (DEAR3)
GmERF#235	Glyma17g27520	209	AtERF#041	At5g11590	OsERF#038	Os02g13710	012 - 070	A4	III	Proteína com Elemento de Interação Responsivo à Desidratação (TINY2)
GmERF#236	Glyma17g31900	199	AtERF#004	At5g11190	OsERF#003	Os02g10760	006 - 063	B6	V	Resposta de Tolerância à Seca (SHN2)
GmERF#237	Glyma17g33060	148	AtERF#003	At5g25190	OsERF#127	Os02g55380	007 - 064	B6	V	Induzível por Etileno e Sal 3 (ESE3)
GmERF#238	Glyma17g33530	160	AtERF#017	At1g19210	OsERF#012	Os08g35240	001 - 058	A5	II	Subfamília DREB A5
GmERF#239	Glyma17g35860	174	AtERF#006	At1g46768	OsERF#008	Os04g55520	044 - 102	A5	II	Subfamília DREB A5 (RAP2.1)
GmERF#240	Glyma17g37350	240	AtERF#049	At1g75490	OsERF#042	Os05g27930	033 - 091	A2	IV	Subfamília DREB A2

**Tabela S14.** Sumário dos Fatores Transcricionais Responsivos à Etileno (ERF) de Soja (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio AP2 de Soja <sup>2</sup>			Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>5</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Posição	C-I <sup>3</sup>	C-II <sup>4</sup>	
GmERF#241	Glyma18g02170	309	AtERF#060	At4g39780	OsERF#047	Os03g09170	122 - 179	A6	I	Subfamília DREB A6
GmERF#242	Glyma18g10290	212	AtERF#082	At1g50640	OsERF#076	Os04g57340	027 - 084	B1	VIII	Subfamília ERF B1 (ERF3)
GmERF#243	Glyma18g20960	196	AtERF#090	At1g24590	OsERF#079	Os02g38090	002 - 059	B1	VIII	Regeneração de “Tiros” em Cultura de Tecidos (DRNL)
GmERF#244	Glyma18g48720	112	AtERF#099	At2g44840	OsERF#093	Os04g46220	052 - 109	B3	IX	Subfamília ERF B3 (ERF13)
GmERF#245	Glyma18g48730	202	AtERF#099	At2g44840	OsERF#091	Os02g43790	076 - 134	B3	IX	Subfamília ERF B3 (ERF13)
GmERF#246	Glyma18g48740	179	AtERF#099	At2g44840	OsERF#091	Os02g43790	033 - 091	B3	IX	Subfamília ERF B3 (ERF13)
GmERF#247	Glyma18g49760	272	AtERF#002	At5g19790	OsERF#004	Os12g39330	022 - 079	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#248	Glyma18g51680	242	AtERF#114	At5g61890	OsERF#101	Os04g32620	056 - 113	B4	X	Subfamília ERF B4
GmERF#249	Glyma19g03120	317	AtERF#002	At5g19790	OsERF#004	Os12g39330	017 - 074	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#250	Glyma19g03170	187	AtERF#002	At5g19790	OsERF#004	Os12g39330	017 - 074	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#251	Glyma19g27790	252	AtERF#114	At5g61890	OsERF#101	Os04g32620	060 - 117	B4	X	Subfamília ERF B4
GmERF#252	Glyma19g29000	253	AtERF#002	At5g19790	OsERF#005	Os07g10410	025 - 082	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#253	Glyma19g32380	282	AtERF#051	At3g57600	OsERF#041	Os03g07830	028 - 086	A2	IV	Subfamília DREB A2
GmERF#254	Glyma19g34650	112	AtERF#095	At3g23220	OsERF#090	Os08g44960	012 - 070	B3	IX	Induzível por Etileno e Sal 1 (ESE1)
GmERF#255	Glyma19g34670	237	AtERF#092	At3g23240	OsERF#087	Os09g39850	070 - 128	B3	IX	Subfamília ERF B3 (ERF1)
GmERF#256	Glyma19g34690	174	AtERF#092	At3g23240	OsERF#083	Os03g64260	038 - 096	B3	IX	Subfamília ERF B3 (ERF1)
GmERF#257	Glyma19g37670	188	AtERF#024	At2g36450	OsERF#032	Os02g43940	022 - 079	A4	III	Tolerância à Sal e Seca (HARDY)
GmERF#258	Glyma19g40070	194	AtERF#071	At2g47520	OsERF#068	Os01g21120	047 - 104	B2	VII	Subfamília ERF B2 (ERF71)
GmERF#259	Glyma19g43820	162	AtERF#092	At3g23240	OsERF#083	Os03g64260	031 - 089	B3	IX	Subfamília ERF B3 (ERF1)
GmERF#260	Glyma19g44240	287	AtERF#066	At4g27950	OsERF#055	Os06g06540	101 - 158	B5	VI	Fator de Resposta à Citocinina 4 (CRF4)
GmERF#261	Glyma19g44580	185	AtERF#023	At1g01250	OsERF#033	Os04g46400	027 - 084	A4	III	Subfamília DREB A4
GmERF#262	Glyma19g45200	259	AtERF#075	At3g14230	OsERF#072	Os09g26420	041 - 098	B2	VII	Subfamília ERF B2 (RAP2.2)
GmERF#263	Glyma20g03890	257	AtERF#002	At5g19790	OsERF#004	Os12g39330	014 - 071	B6	V	Subfamília ERF B6 (RAP2.11)
GmERF#264	Glyma20g16910	267	AtERF#105	At5g51190	OsERF#097	Os04g46240	118 - 176	B3	IX	Subfamília ERF B3 (ERF105)
GmERF#265	Glyma20g16920	209	AtERF#106	At5g07580	OsERF#097	Os04g46240	113 - 171	B3	IX	Subfamília ERF B3
GmERF#266	Glyma20g29410	207	AtERF#031	At4g25480	OsERF#025	Os02g45450	053 - 110	A1	III	Proteína com Elemento de Interação Responsivo à Desidratação 1A (DREB1A)
GmERF#267	Glyma20g29440	190	AtERF#025	At5g52020	OsERF#020	Os02g45420	027 - 084	A4	III	Subfamília DREB A4
GmERF#268	Glyma20g31300	087	AtERF#088	At1g12890	OsERF#080	Os08g07700	001 - 058	B1	VIII	Subfamília ERF B1
GmERF#269	Glyma20g33800	199	AtERF#106	At5g07580	OsERF#097	Os04g46240	108 - 166	B3	IX	Subfamília ERF B3
GmERF#270	Glyma20g33840	155	AtERF#101	At5g47220	OsERF#093	Os04g46220	068 - 126	B3	IX	Subfamília ERF B3 (ERF2)

**Tabela S14.** Sumário dos Fatores Transcricionais Responsivos à Etileno (ERF) de Soja (cont.)

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínio AP2 de Soja <sup>2</sup>			Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>5</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Posição	C-I <sup>3</sup>	C-II <sup>4</sup>	
			Nome Genérico	Locus ID	Nome Genérico	Locus ID				
GmERF#271	Glyma20g33890	386	AtERF#062	At4g13620	OsERF#048	Os08g31580	220 - 277	A6	I	Subfamília DREB A6
GmERF#272	Glyma20g34550	147	AtERF#096	At5g43410	OsERF#089	Os10g30840	018 - 076	B3	IX	Subfamília ERF B3
GmERF#273	Glyma20g34560	134	AtERF#098	At3g23230	OsERF#090	Os08g44960	017 - 075	B3	IX	Regulador Transcricional de Resposta à Defesa 1 (TDR1)
GmERF#274	Glyma20g34570	214	AtERF#092	At3g23240	OsERF#083	Os03g64260	073 - 131	B3	IX	Subfamília ERF B3 (ERF1)
GmERF#275	Glyma20g35820	193	AtERF#064	At4g23750	OsERF#053	Os01g12440	080 - 138	B5	VI	Alvo de Monópteros 3 (TMO3)

<sup>1</sup> Tamanho em resíduos de aminoácidos.

<sup>2</sup> Análise do único domínio AP2 (PF00847) presente na sequência de aminoácidos em questão.

<sup>3</sup> Classificação de acordo com Sakuma et al. (2002).

<sup>4</sup> Classificação de acordo com Nakano et al. (2006).

<sup>3,4</sup> Ambas as classificações foram determinadas pela homologia da sequência de aminoácidos e distância filogenética com seus respectivos homólogos em *A. thaliana* e *O. sativa*.

<sup>5</sup> Este parâmetro foi determinado pela homologia com a sequência de aminoácidos de *Arabidopsis thaliana* equivalente.

**Tabela S15.** Sumário das Proteínas RAV de Soja

Proteína de <i>Glycine max</i>			Melhor Homologia da Sequência de Aminoácidos				Domínios AP2/B3 de Soja		Suposta Função e/ou Nome(s) Completo(s) da Proteína <sup>3</sup>
Nome Genérico	Locus ID	R.A. <sup>1</sup>	<i>Arabidopsis thaliana</i>		<i>Oryza sativa</i>		Código PFAM <sup>2</sup>	Posição	
GmRAV#001	Glyma01g22260	384	AtRAV#001	At1g25560	OsRAV#003	Os01g49830	PF00847	075 - 123	Fator de Interação com DNA em Resposta à Etileno 1 (EDF1); “Tempranillo” 1 (TEM1)
							PF02362	205 - 313	
GmRAV#002	Glyma02g11060	401	AtRAV2	At1g68840	OsRAV#004	Os05g47650	PF00847	081 - 129	Relacionado à ABI3/VP1 2 (RAV2); “Tempranillo” 2 (TEM2)
							PF02362	210 - 325	
GmRAV#003	Glyma10g34760	351	AtRAV#001	At1g25560	OsRAV#004	Os05g47650	PF00847	052 - 100	Fator de Interação com DNA em Resposta à Etileno 1 (EDF1); “Tempranillo” 1 (TEM1)
							PF02362	172 - 281	
GmRAV#004	Glyma20g32730	362	AtRAV#001	At1g25560	OsRAV#004	Os05g47650	PF00847	058 - 106	Fator de Interação com DNA em Resposta à Etileno 1 (EDF1); “Tempranillo” 1 (TEM1)
							PF02362	178 - 276	
GmRAV#005	Glyma20g39140	256	AtRAV#002	At1g50680	OsRAV#002	Os01g04800	PF00847	001 - 050	Proteína Relacionada à ABI3/VP1 (RAV)
							PF02362	122 - 233	

<sup>1</sup> Tamanho em resíduos de aminoácidos.

<sup>2</sup> PF00847: Domínio AP2; PF02362: Domínio B3 de interação com DNA.

<sup>3</sup> Este parâmetro foi determinado pela homologia com a sequência de aminoácidos de *Arabidopsis thaliana* equivalente.

**Tabela S16.** Experimento de BBH - Biossíntese de Etileno em Soja<sup>1,2</sup>

Proteína de <i>Glycine max</i>		Proteína de <i>Arabidopsis thaliana</i>		Proteína de <i>Oryza sativa</i>	
Nome Genérico	Locus ID	Nome Genérico	Locus ID	Nome Genérico	Locus ID
GmAAT#001	Glyma02g01830	–	–	OsAAT#002	Os10g34350
GmAAT#004	Glyma06g11640	AtAAT	At1g77670	OsAAT#001	Os09g28050
GmACD#001	Glyma05g33540	AtACD1	At1g48420	OsACD#001	Os02g53330
GmACO#004	Glyma04g42460	AtACO#002	At1g77330	OsACO#003	Os05g05680
GmACO#006	Glyma06g12340	AtACO#002	At1g77330	OsACO#003	Os05g05680
GmACO#007	Glyma07g15480	AtACO1	At2g19590	–	–
GmACO#008	Glyma07g39420	–	–	OsACO#005	Os09g27750
GmACO#009	Glyma08g03310	–	–	OsACO#004	Os06g37590
GmACO#014	Glyma14g05390	AtACO4	At1g05010	–	–
GmACS#006	Glyma05g36250	AtACS7	At4g26200	–	–
GmACS#011	Glyma08g02130	AtACS6	At4g11280	OsACS#003	Os04g48850
GmACS#012	Glyma08g03400	–	–	OsACS#004	Os05g10780
GmACS#013	Glyma09g28000	AtACS12	At5g51690	OsACS#006	Os06g03990
GmACS#014	Glyma09g39060	AtACS1	At3g61510	–	–
GmACS#016	Glyma11g03070	AtACS10	At1g62960	–	–
GmACS#017	Glyma11g04890	AtACS8	At4g37770	–	–
GmACS#020	Glyma17g16990	–	–	OsACS#002	Os03g51740
GmACT#003	Glyma08g27500	AtACT#001	At1g03940	–	–
GmACT#006	Glyma12g32630	AtACT#002	At1g03495	–	–
GmACT#017	Glyma18g06500	AtACT#003	At2g30200	OsACT#002	Os03g18590
GmACT#020	Glyma18g50320	AtACT#005	At5g39050	–	–
GmACT#023	Glyma18g50350	–	–	OsACT#001	Os02g57480
GmARD#001	Glyma10g38940	–	–	OsARD#001	Os03g06620
GmARD#002	Glyma10g38960	AtARD4	At5g43850	OsARD#004	Os10g28360
GmARD#004	Glyma20g28860	AtARD2	At4g14710	–	–
GmASP#002	Glyma06g08670	AtASP3	At5g11520	OsASP#001	Os01g55540
GmASP#004	Glyma17g33050	AtASP5	At4g31990	OsASP#003	Os02g55420
GmDEP#001	Glyma06g36620	AtDEP1	At5g53850	–	–
GmDEP#002	Glyma12g35390	–	–	OsDEP#001	Os11g29370
GmGGT#001	Glyma01g00850	AtGGT4	At4g29210	OsGGT#001	Os01g05810
GmGGT#003	Glyma11g35950	AtGGT1	At4g39640	OsGGT#003	Os04g38450
GmMAT#007	Glyma13g20480	AtMAT3	At2g36880	–	–
GmMAT#008	Glyma15g21890	AtMAT4	At3g17390	–	–
GmMAT#009	Glyma17g04330	–	–	OsMAT#003	Os05g04510
GmMAT#011	Glyma19g40810	AtMAT2	At4g01850	OsMAT#002	Os01g22010
GmMTI#001	Glyma09g08190	AtMTI1	At2g05830	OsMTI#001	Os11g11050
GmMTI#002	Glyma15g19690	AtMTI1	At2g05830	–	–
GmMTK#003	Glyma20g26580	AtMTK1	At1g49820	OsMTK#001	Os04g57400
GmMTN#001	Glyma04g00770	AtMTN1	At4g38800	–	–
GmMTN#002	Glyma06g00790	–	–	OsMTN#001	Os06g02220

<sup>1</sup> Proteínas de soja relacionadas à biossíntese de etileno que possuem um homólogo BBH positivo com *Arabidopsis thaliana* e/ou *Oryza sativa*.

<sup>2</sup> BBH: Melhor “Hit” Bidirecional.

**Tabela S17.** Experimento de BBH - Transdução de Sinal Mediada por Etileno em Soja<sup>1,2</sup>

Proteína de <i>Glycine max</i>		Proteína de <i>Arabidopsis thaliana</i>		Proteína de <i>Oryza sativa</i>	
Nome Genérico	Locus ID	Nome Genérico	Locus ID	Nome Genérico	Locus ID
GmCPK#004	Glyma08g00840	–	–	OsCPK#001	Os11g07040
GmCTR#004	Glyma13g21480	AtCTR1	At5g03730	OsCTR#001	Os02g32610
GmEBF#005	Glyma14g14410	AtEBF1	At2g25490	OsEBF#001	Os02g10700
GmEIN#007	Glyma13g20810	AtEIN2	At5g03280	OsEIN#006	Os07g06130
GmEIN#008	Glyma14g04550	–	–	OsEIN#003	Os03g20790
GmEIN#009	Glyma19g43160	AtEIN5	At1g54490	–	–
GmEIN#010	Glyma20g12250	AtEIN3	At3g20770	OsEIN#002	Os03g20780
GmEIN#011	Glyma20g37260	–	–	OsEIN#005	Os03g58060
GmETO#001	Glyma01g31170	AtETO1	At3g51770	OsETO#001	Os03g18360
GmETO#002	Glyma02g16310	–	–	OsETO#003	Os11g37520
GmETO#010	Glyma19g34200	AtEOL1	At4g02680	–	–
GmETR#002	Glyma03g41220	AtEIN4	At3g04580	OsETR#003	Os04g08740
GmETR#003	Glyma09g00490	–	–	OsETR#002	Os03g49500
GmETR#006	Glyma12g37050	AtETR1	At1g66340	–	–
GmETR#007	Glyma19g40090	AtERS1	At2g40940	–	–
GmETR#010	Glyma20g34420	AtETR2	At3g23150	–	–
GmMKK#001	Glyma07g00520	AtMKK4	At1g51660	OsMKK#005	Os06g09180
GmMKK#002	Glyma07g11910	AtMKK9	At1g73500	OsMKK#001	Os03g12390
GmMPK#001	Glyma02g15690	AtMPK6	At2g43790	OsMPK#002	Os06g06090
GmMPK#004	Glyma12g07770	AtMPK3	At3g45640	OsMPK#001	Os03g17700
GmRAN#002	Glyma04g07360	–	–	OsRAN#001	Os05g49890
GmRAN#003	Glyma04g07370	–	–	OsRAN#001	Os05g49890
GmRAN#005	Glyma06g07400	AtRAN3	At5g55190	–	–
GmRAN#006	Glyma06g07410	–	–	OsRAN#001	Os05g49890
GmRTE#002	Glyma11g07740	AtRTH1	At3g51040	OsRTE#002	Os03g58520
GmRTE#003	Glyma14g10770	AtRTE1	At2g26070	OsRTE#001	Os01g51430

<sup>1</sup> Proteínas de soja relacionadas à transdução de sinal induzida por etileno que possuem um homólogo BBH positivo com *Arabidopsis thaliana* e/ou *Oryza sativa*.

<sup>2</sup> BBH: Melhor “Hit” Bidirecional.

**Tabela S18.** Experimento de BBH - Proteínas AP2 de Soja<sup>1,2</sup>

Proteína de <i>Glycine max</i>		Proteína de <i>Arabidopsis thaliana</i>		Proteína de <i>Oryza sativa</i>	
Nome Genérico	Locus ID	Nome Genérico	Locus ID	Nome Genérico	Locus ID
GmAP2#001	Glyma01g02760	AtAP2#013	At5g10510	–	–
GmAP2#002	Glyma01g39520	AtAP2#011	At4g36920	–	–
GmAP2#007	Glyma03g29240	AtAP2#007	At2g41710	OsAP2#011	Os05g32270
GmAP2#008	Glyma03g33470	–	–	OsAP2#016	Os07g13170
GmAP2#009	Glyma04g05080	–	–	OsAP2#005	Os03g56050
GmAP2#013	Glyma06g05170	AtAP2#012	At4g37750	–	–
GmAP2#015	Glyma07g02380	AtAP2#001	At1g16060	–	–
GmAP2#018	Glyma08g24420	AtAP2#009	At3g54320	–	–
GmAP2#028	Glyma11g15650	AtAP2#005	At2g28550	–	–
GmAP2#029	Glyma12g06010	AtAP2#002	At1g51190	–	–
GmAP2#030	Glyma12g07800	AtAP2#005	At2g28550	–	–
GmAP2#031	Glyma13g00950	–	–	OsAP2#009	Os04g55970
GmAP2#032	Glyma13g40470	–	–	OsAP2#010	Os05g03040
GmAP2#035	Glyma15g34770	–	–	OsAP2#019	Os12g03290
GmAP2#036	Glyma16g00950	AtAP2#003	At1g72570	–	–
GmAP2#037	Glyma17g07010	AtAP2#015	At5g57390	OsAP2#001	Os01g67410
GmAP2#040	Glyma17g18640	–	–	OsAP2#008	Os04g55560
GmAP2#043	Glyma18g47980	AtAP2#014	At5g17430	OsAP2#002	Os02g40070
GmAP2#044	Glyma18g49100	–	–	OsAP2#013	Os06g05340
GmAP2#046	Glyma19g36200	–	–	OsAP2#016	Os07g13170

<sup>1</sup> Proteínas AP2 de soja que possuem um homólogo BBH positivo com *Arabidopsis thaliana* e/ou *Oryza sativa*.

<sup>2</sup> BBH: Melhor “Hit” Bidirecional.

**Tabela S19.** Experimento de BBH - Fatores Transcricionais Responsivos à Etileno de Soja<sup>1,2</sup>

Proteína de <i>Glycine max</i>		Proteína de <i>Arabidopsis thaliana</i>		Proteína de <i>Oryza sativa</i>	
Nome Genérico	Locus ID	Nome Genérico	Locus ID	Nome Genérico	Locus ID
GmERF#001	Glyma01g03110	AtERF#090	At1g24590	OsERF#079	Os02g38090
GmERF#005	Glyma01g41520	AtERF#100	At4g17500	–	–
GmERF#006	Glyma01g42500	AtERF#033	At1g12610	–	–
GmERF#007	Glyma01g42510	–	–	OsERF#020	Os02g45420
GmERF#009	Glyma01g43450	AtERF#061	At1g64380	OsERF#052	Os05g49700
GmERF#010	Glyma01g44130	AtERF#022	At1g33760	OsERF#019	Os11g13840
GmERF#012	Glyma01g44230	AtERF#020	At1g71520	–	–
GmERF#016	Glyma02g04460	–	–	OsERF#079	Os02g38090
GmERF#020	Glyma02g08840	–	–	OsERF#099	Os01g64790
GmERF#024	Glyma02g42960	AtERF#048	At2g40340	OsERF#042	Os05g27930
GmERF#025	Glyma02g43240	AtERF#052	At2g40220	OsERF#117	Os05g28350
GmERF#029	Glyma03g26310	AtERF#099	At2g44840	–	–
GmERF#035	Glyma03g27050	AtERF#034	At2g44940	–	–
GmERF#036	Glyma03g29530	AtERF#051	At3g57600	–	–
GmERF#040	Glyma03g31940	AtERF#093	At2g31230	–	–
GmERF#041	Glyma03g34970	AtERF#024	At2g36450	OsERF#018	Os10g38000
GmERF#042	Glyma03g41640	AtERF#067	At3g61630	–	–
GmERF#043	Glyma03g41910	AtERF#023	At1g01250	–	–
GmERF#045	Glyma04g03070	AtERF#049	At1g75490	–	–
GmERF#048	Glyma04g06690	AtERF#003	At5g25190	–	–
GmERF#049	Glyma04g07140	AtERF#004	At5g11190	–	–
GmERF#050	Glyma04g08900	AtERF#041	At5g11590	OsERF#038	Os02g13710
GmERF#053	Glyma04g19650	–	–	OsERF#003	Os02g10760
GmERF#056	Glyma04g37890	–	–	OsERF#101	Os04g32620
GmERF#059	Glyma04g43040	AtERF#012	At1g21910	OsERF#119	Os06g10780
GmERF#060	Glyma05g03540	–	–	OsERF#020	Os02g45420
GmERF#063	Glyma05g05130	AtERF#102	At5g47230	OsERF#097	Os04g46240
GmERF#066	Glyma05g19050	AtERF#010	At5g67190	–	–
GmERF#074	Glyma06g04490	–	–	OsERF#008	Os04g55520
GmERF#076	Glyma06g06780	AtERF#003	At5g25190	–	–
GmERF#079	Glyma06g11010	AtERF#059	At1g78080	OsERF#049	Os02g51670
GmERF#080	Glyma06g11700	–	–	OsERF#119	Os06g10780
GmERF#082	Glyma06g17180	AtERF#114	At5g61890	–	–
GmERF#085	Glyma06g44430	AtERF#080	At5g44210	–	–
GmERF#087	Glyma06g45680	AtERF#045	At5g05410	–	–
GmERF#091	Glyma07g03500	AtERF#001	At1g15360	–	–
GmERF#092	Glyma07g04950	AtERF#074	At1g53910	–	–
GmERF#099	Glyma07g33510	AtERF#109	At4g34410	–	–
GmERF#101	Glyma07g37990	AtERF#086	At5g18560	OsERF#078	Os07g47330
GmERF#102	Glyma08g02460	AtERF#064	At4g23750	–	–
GmERF#105	Glyma08g14600	AtERF#060	At4g39780	–	–
GmERF#107	Glyma08g21650	–	–	OsERF#034	Os04g46440
GmERF#109	Glyma08g23070	AtERF#084	At1g80580	–	–
GmERF#113	Glyma08g43300	AtERF#082	At1g50640	OsERF#134	Os02g09650
GmERF#119	Glyma09g27180	AtERF#028	At5g51990	–	–
GmERF#124	Glyma10g00990	AtERF#098	At3g23230	–	–
GmERF#134	Glyma10g33060	AtERF#092	At3g23240	–	–

**Tabela S19.** Experimento de BBH - Fatores Transcricionais Responsivos à Etileno de Soja (cont.)<sup>1,2</sup>

Proteína de <i>Glycine max</i>		Proteína de <i>Arabidopsis thaliana</i>		Proteína de <i>Oryza sativa</i>	
Nome Genérico	Locus ID	Nome Genérico	Locus ID	Nome Genérico	Locus ID
GmERF#135	Glyma10g33070	AtERF#098	At3g23230	–	–
GmERF#137	Glyma10g33700	AtERF#062	At4g13620	–	–
GmERF#139	Glyma10g36300	–	–	OsERF#080	Os08g07700
GmERF#140	Glyma10g36760	AtERF#110	At5g50080	–	–
GmERF#143	Glyma10g42130	AtERF#118	At1g68550	–	–
GmERF#144	Glyma11g01640	AtERF#021	At1g71450	–	–
GmERF#145	Glyma11g01700	AtERF#020	At1g71520	OsERF#017	Os01g66270
GmERF#150	Glyma11g03910	–	–	OsERF#091	Os02g43790
GmERF#159	Glyma12g33020	AtERF#053	At2g20880	OsERF#046	Os02g42580
GmERF#160	Glyma12g35550	AtERF#078	At3g15210	–	–
GmERF#164	Glyma13g17250	–	–	OsERF#011	Os02g54050
GmERF#177	Glyma13g30710	AtERF#091	At4g18450	OsERF#084	Os05g49010
GmERF#180	Glyma13g31010	AtERF#081	At1g28360	–	–
GmERF#184	Glyma13g39540	–	–	OsERF#027	Os01g73770
GmERF#190	Glyma14g06290	–	–	OsERF#117	Os05g28350
GmERF#192	Glyma14g09320	AtERF#006	At1g46768	–	–
GmERF#195	Glyma14g22970	AtERF#013	At1g77640	OsERF#009	Os03g15660
GmERF#196	Glyma14g29040	AtERF#066	At4g27950	–	–
GmERF#200	Glyma15g00660	–	–	OsERF#127	Os02g55380
GmERF#203	Glyma15g02900	–	–	OsERF#018	Os10g38000
GmERF#205	Glyma15g08370	–	–	OsERF#075	Os01g58420
GmERF#210	Glyma15g16260	AtERF#072	At3g16770	–	–
GmERF#212	Glyma15g17100	–	–	OsERF#128	Os04g18650
GmERF#216	Glyma16g04410	AtERF#002	At5g19790	OsERF#005	Os07g10410
GmERF#220	Glyma16g26460	–	–	OsERF#082	Os04g32790
GmERF#224	Glyma17g02710	AtERF#086	At5g18560	OsERF#078	Os07g47330
GmERF#225	Glyma17g05240	AtERF#017	At1g19210	OsERF#011	Os02g54050
GmERF#234	Glyma17g18580	AtERF#008	At2g23340	–	–
GmERF#243	Glyma18g20960	–	–	OsERF#079	Os02g38090
GmERF#249	Glyma19g03120	–	–	OsERF#004	Os12g39330
GmERF#253	Glyma19g32380	–	–	OsERF#041	Os03g07830
GmERF#254	Glyma19g34650	AtERF#095	At3g23220	–	–
GmERF#255	Glyma19g34670	–	–	OsERF#087	Os09g39850
GmERF#258	Glyma19g40070	AtERF#071	At2g47520	OsERF#068	Os01g21120
GmERF#259	Glyma19g43820	–	–	OsERF#083	Os03g64260
GmERF#264	Glyma20g16910	AtERF#105	At5g51190	–	–
GmERF#265	Glyma20g16920	AtERF#106	At5g07580	–	–
GmERF#266	Glyma20g29410	–	–	OsERF#025	Os02g45450
GmERF#267	Glyma20g29440	AtERF#025	At5g52020	–	–
GmERF#273	Glyma20g34560	–	–	OsERF#090	Os08g44960

<sup>1</sup> Fatores transcricionais responsivos à etileno de soja que possuem um homólogo BBH positivo com *Arabidopsis thaliana* e/ou *Oryza sativa*.

<sup>2</sup> BBH: Melhor “Hit” Bidirecional.

**Tabela S20.** Experimento de BBH - Proteínas RAV de Soja<sup>1,2</sup>

Proteína de <i>Glycine max</i>		Proteína de <i>Arabidopsis thaliana</i>		Proteína de <i>Oryza sativa</i>	
Nome Genérico	Locus ID	Nome Genérico	Locus ID	Nome Genérico	Locus ID
GmRAV#001	Glyma01g22260	AtRAV#001	At1g25560	–	–
GmRAV#002	Glyma02g11060	–	–	OsRAV#004	Os05g47650
GmRAV#005	Glyma20g39140	AtRAV#002	At1g50680	–	–

<sup>1</sup> Proteínas RAV de soja que possuem um homólogo BBH positivo com *Arabidopsis thaliana* e/ou *Oryza sativa*.

<sup>2</sup> BBH: Melhor “Hit” Bidirecional.

As Tabelas Suplementares S21 - S83 apresentam os motivos proteicos comuns presentes nos membros dos grupos de fatores de transcrição de soja responsáveis à etileno (ERFs). Na nomenclatura, o numeral romano (I - X) representa o grupo ERF e o arábico corresponde ao número do motivo. O grupo VI-L não foi representado por apresentar apenas um membro.

Tabela S21. Motivo I-1

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#009	293	2.91e <sup>-12</sup>	CQTEETELLF <b>EGCSLARLPSFDPELI</b> WEVLTV
GmERF#051	297	1.94e <sup>-19</sup>	VSEPQWEGDS <b>DNFNLQKYPSEIDWD</b> SL
GmERF#068	295	2.11e <sup>-17</sup>	SDSNNQWDEM <b>ENFGLKFPSEIDWA</b> AI
GmERF#079	285	4.39e <sup>-18</sup>	VSEAQWEGDS <b>DNYNLQKYPSEIDWD</b> SL
GmERF#105	295	1.64e <sup>-17</sup>	SDSNNQWDEM <b>ENFGLKFPSEIDWE</b> AI
GmERF#137	363	5.32e <sup>-14</sup>	KSTHHTSDV <b>EGVQLSRMPSLMDII</b> WDELFSNS
GmERF#146	304	2.91e <sup>-12</sup>	CQSMETEELF <b>EGCSLARLPSFDPELI</b> WEVLAV
GmERF#161	294	7.69e <sup>-18</sup>	VTEPQWEDAS <b>AFNQLKFPSEIDWD</b> SL
GmERF#198	295	4.42e <sup>-20</sup>	LTFDVTEPQW <b>EFNQLKFPSEIDWD</b> SL
GmERF#241	289	9.06e <sup>-15</sup>	FSDSCETMDN <b>LGLDLEKYPSEIDWA</b> ALSDS
GmERF#271	362	5.32e <sup>-14</sup>	KSKHHTSDV <b>EGVQLSRMPSLMDII</b> WDELFSNS
<b>Consenso</b>			<b>E</b> [GN] <b>FNL</b> [QE] <b>[KR]</b> [FY] <b>PS</b> [YV] <b>[ED]</b> IDW[DI]

Tabela S22. Motivo I-2

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#051	001	2.78e <sup>-15</sup>	<b>MAAMDFYSSSTEFQ</b> LHSDPFRGEL
GmERF#068	001	2.24e <sup>-10</sup>	<b>MGTAIMYSNNNIVP</b> DFLDPYSEEL
GmERF#079	001	1.92e <sup>-13</sup>	<b>MAALMDFYSSSPEFQ</b> LHSDPFRGEL
GmERF#086	179	3.28e <sup>-09</sup>	YWSDAINLSP <b>RGRMMMMNRTEGRQ</b> MLRPQAQPLN
GmERF#105	001	1.92e <sup>-13</sup>	<b>MGTAIMYNSNNIVA</b> DFLDPYSEEL
GmERF#137	160	8.75e <sup>-12</sup>	SSSNIKGGEV <b>MSATSDPSYFVSQFV</b> QNQHQAGSGD
GmERF#161	001	1.04e <sup>-14</sup>	<b>MAATMNFYNETSQQV</b> QSDPFRGELM
GmERF#198	001	1.47e <sup>-14</sup>	<b>MAATMNFYNGTSQEQ</b> VESDPFRGEL
GmERF#241	001	6.63e <sup>-16</sup>	<b>MATVMDMYNTSSIMI</b> PDFLDQCSQDQ
GmERF#271	159	3.36e <sup>-11</sup>	SSSNIKGGEV <b>MSATSDPSYFVSQFG</b> QNQHQAGSGD
<b>Consenso</b>			<b>M</b> [AGS] <b>[AT]</b> [TAM] <b>[MIS]</b> [DN] <b>[FMP]</b> [YS] <b>[NSY]</b> [SH] <b>[STV]</b> [SN] <b>[QIE]</b> [FV] <b>[QV]</b>

Tabela S23. Motivo I-3

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#009	182	2.79e <sup>-30</sup>	<b>FPNLKDPTKL</b> <b>GFGDSARLNALKSSVD</b> AKIQ <b>KVKRE</b> AKK <b>NAAKL</b> HN STDPKRGQSE
GmERF#051	204	7.88e <sup>-46</sup>	<b>FPNLRHQGSS</b> <b>VGGDFGEYKPLHS</b> AVDAKL <b>QAICEGLAELQKQKTEKPPRK</b> TRSKLASPPE
GmERF#068	182	5.99e <sup>-39</sup>	<b>FPHLRHHGAF</b> <b>VFGEFGDYREL</b> PSSVDS <b>KLQAICESLAKQEEKPCCSVEDVK</b> PVIHAAELAE
GmERF#079	193	7.88e <sup>-46</sup>	<b>FPNLRHQGSS</b> <b>VGGDFGEYKPLHS</b> AVDAKL <b>QAICEGLAELQKQKTEKPPRK</b> SRSKLAEKVV
GmERF#105	186	3.09e <sup>-40</sup>	<b>FPHLRHHGAF</b> <b>VFGEFGDYKPL</b> PSSVDS <b>KLQAICESLAKQEEKPCCSVEDVK</b> PVIHAAELAE
GmERF#146	202	7.48e <sup>-31</sup>	<b>FPNLKDPTKL</b> <b>GFGDSARLNALKSSVD</b> AKIQ <b>KVKKER</b> AKK <b>NATKKLH</b> DSSGSTDPNG
GmERF#161	198	5.27e <sup>-35</sup>	<b>LARLNFPNLK</b> <b>GSCPGEEYKPMQA</b> AVDAKL <b>DAICANLAEMQKQKNEKGARS</b> GKSKSQGPNL
GmERF#198	211	1.58e <sup>-37</sup>	<b>FARLNFPNLK</b> <b>GSCPGEEYKPMQA</b> AVDAKL <b>DAICANLAEMQKQKTEKGARS</b> AKSKSQGPNQ
GmERF#241	188	2.09e <sup>-33</sup>	<b>FPHLRHHGAR</b> <b>AYGEFGNYKEL</b> PSAVDAKL <b>QAICQSLGTNSQKLTQNP</b> IVL DTHKAETETS
<b>Consenso</b>			[GV] <b>[FGS]</b> [GC] <b>[DEP]</b> [FGS] <b>[GAE]</b> [EDR] <b>[YL]</b> [KN] <b>[PA]</b> L <b>[HPK]</b> S <b>[AS]</b> VD <b>[AS]</b> K [LI] <b>[QD]</b> AIC <b>[EQA]</b> [SGKN] <b>[LV]</b> [AK] <b>[EK]</b> [ELMQ] <b>[QER]</b> [KAE] <b>[KQ]</b> [GKP] <b>[KCN]</b> [TAC] <b>[ES]</b> [KV] <b>[PEGK]</b> [ADP] <b>[RV]</b> [KS]

Tabela S24. Motivo I-4

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#051	019	1.57e <sup>-24</sup>	SSSTEFQLHS <b>DPFRGELMEVLEPFMKSEFST</b> PSPNSCFLS
GmERF#068	019	1.62e <sup>-22</sup>	SNNNIVPDFL <b>DPYSEELMKALMPFMKSDYFS</b> ASSSSSSRES
GmERF#079	019	2.26e <sup>-22</sup>	SSSPEFQLHS <b>DPFRGELMEVLEPFMKSPSPN</b> YFPSSPSLPN
GmERF#105	019	4.30e <sup>-22</sup>	NSSNIVADFL <b>DPYSEELMKALMPFMKSDYFS</b> ASSSSSLESQ
GmERF#161	018	3.31e <sup>-25</sup>	YNETSQQVQS <b>DPFRGELMEVLEPFMKTCFSS</b> TPSILSSDSP
GmERF#198	019	1.40e <sup>-23</sup>	NGTSQEQVES <b>DPFRGELMEVLEPFMKTSFSS</b> TTPSIILSSD
GmERF#241	021	3.50e <sup>-18</sup>	SSSIMIPDFLD <b>QCSQELMKALEPFMKSDSSS</b> ITYPSFSSFS
<b>Consenso</b>			DP <b>[FY]</b> [RS] <b>[GE]</b> ELM <b>[EK]</b> [VA] <b>LEPFMK</b> [ST] <b>[DP]</b> [PSY] [SF]S



Tabela S25. Motivo I-5

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#051	086	2.51e <sup>-29</sup>	QNLIGFGQGQ <b>PTSLVGLNHLTPSQISQIQAIQIQNHS</b> NTLFSLGPKP
GmERF#068	062	2.59e <sup>-25</sup>	SPLIPSNLSP <b>SSNQIRLNQLTQDQILQIQAIHIQQQH</b> VAQQQAHLGP
GmERF#079	070	3.78e <sup>-32</sup>	QSLIGFGQAQ <b>PTSLVGLNHLTPSQISQIQAIQIQIQAQQ</b> HQNRSNTLSF
GmERF#105	064	2.00e <sup>-28</sup>	SSNSLPTSYP <b>SSNQIKLNQLTQDQIVQIQAIHIQQQH</b> QHVAQTQTHL
GmERF#161	086	9.68e <sup>-24</sup>	QQPLIGFEQQ <b>PSSLLGLNHLSTSQISQIQAAQAQNSL</b> SLNFLGPKPV
GmERF#198	094	4.04e <sup>-27</sup>	QQSLIGFEQP <b>PSSLLGLNHLSPSQISQIQAIQAQQSQ</b> NQNPHSLNFL
GmERF#241	072	4.07e <sup>-22</sup>	YEQKVSITPN <b>HASSIKLNQLTQSQMFIQARIQVPRGQ</b> FLSPKPIPMK
<b>Consenso</b>			[PS] [ST] [SN] [LQ] [ILV] [GK] LN [HQ] L [TS] P [SD] QISQIQAI [QH] [IA] Q [QN] [QS] Q

Tabela S26. Motivo I-6

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#051	115	3.99e <sup>-26</sup>	AQIQIQNHSN <b>TLFSLGPKPIPMKHVGMPPKP</b> TKLYRGVRQR
GmERF#068	093	2.12e <sup>-20</sup>	IHIQQQHVAQ <b>GQAHLGPKRVPMKHAGTAAKA</b> AKLYRGVRQR
GmERF#079	104	2.15e <sup>-25</sup>	QAQQHQNRSN <b>TLFSLGPKPIPMKHAGMPPKP</b> TKLYRGVRQR
GmERF#105	097	7.85e <sup>-23</sup>	IQQQQQHVAQ <b>TQTHLGPKRVPMKHAGTAAKA</b> TKLYRGVRQR
GmERF#161	114	1.13e <sup>-24</sup>	QAQAQAQNSL <b>SLNFLGPKPVPMKHVGGPAKP</b> TKLYRGVRQR
GmERF#198	127	4.66e <sup>-25</sup>	AQQSQNQNP <b>SLNFLGPKPVPMKHVGGPPKP</b> TKLYRGVRQR
<b>Consenso</b>			[TS] [LQ] [NS] [FH] LGPK [PR] [VI] PMKH [AV] G [GMT] [PA] [AP] KP

Tabela S27. Motivo I-7

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#086	103	5.36e <sup>-15</sup>	IVFPFAFEGS <b>QHMPFPHPFGTT</b> NLPLFRPNLH
GmERF#137	281	4.88e <sup>-13</sup>	EYAQLNFPDL <b>KHVIQANSLNGTT</b> AALVEAKLQA
GmERF#159	113	3.83e <sup>-18</sup>	VFPPAFDGSQ <b>QHMQFPHQFGTT</b> NPPFPHPSLQ
GmERF#182	068	3.83e <sup>-18</sup>	IVFPFAFDGS <b>QHMQFPHQFGTT</b> TNSPFPHPSP
GmERF#271	280	4.88e <sup>-13</sup>	EYAQLNFPDL <b>KHVIQANSLNGTT</b> AALVEAKLQA
<b>Consenso</b>			[QK] H [PV] [MI] [QP] [FA] [PN] [HS] [LQP] [FN] GTT

Tabela S28. Motivo II-1

Nome Genérico	Início	Valor $p$	Sítio
GmERF#004	096	1.72e <sup>-21</sup>	LLVREGPAAL VAGCDMSAASIRKKATEVGARVDAL QATLHHHYVP
GmERF#012	076	8.13e <sup>-17</sup>	NFPETLSSFS LQLRDMSPRSVQKVASDIGMDVDAR NILAGKTSTV
GmERF#046	102	5.24e <sup>-24</sup>	LLSQDDDVST QQQGNMSADSIKKATQVGARVDAL QTALQQSSSS
GmERF#047	069	2.77e <sup>-16</sup>	NTPCNMDTTN APHQSLTPQEIQEVAAANFANQAPLL QKPQEPLEYK
GmERF#059	102	1.71e <sup>-18</sup>	FPLSSSQYI PGEAVMSPKSIQRVAAAAANSFIDN ANTPPSPPPA
GmERF#066	090	1.20e <sup>-19</sup>	PELLAAEGPA ASDAVMSAASIRKKATEVGARVDAL HRQDPHAPPA
GmERF#074	102	9.76e <sup>-23</sup>	LLSQDDDVST QQQGKMSADSIKKATQVGARVDAL QTALQQSSST
GmERF#075	094	3.99e <sup>-16</sup>	PCNMDTATNA PENQSLTPQEIQEVAAKFANQQVDR GDMMDWTFIN
GmERF#080	099	1.21e <sup>-18</sup>	FPLSSSQYI PGDAVMSPKSIQRVAAAAANSFIDN ATTPPSPPPA
GmERF#118	085	2.48e <sup>-17</sup>	FNFPSPDPNI AGGGNMTSSQIQIAAARFANSEPRN ERSDQPVESL
GmERF#145	083	6.42e <sup>-18</sup>	NFPETLSSYS VQLRDMSPRSVQKVASDVGMVDAR NIVAGKTSTV
GmERF#151	101	1.41e <sup>-13</sup>	LLIGEGAAAL TGGCDMSAASIRKKASELLSASC GG GSGDFAVRVV
GmERF#164	082	3.34e <sup>-18</sup>	FNFPDPNPPDI AGGESMTSPSIQVAAAQFANAGPHE GHSGRPEHSP
GmERF#192	119	1.21e <sup>-20</sup>	DQEGSDSVQH GAAGNMSADSIKKATQVGARVDAL QTALHHHAPS
GmERF#195	079	3.04e <sup>-17</sup>	SSSSQHYIIP QDTAMMSPKSIQRVAAAAANNFLDN NAIAINNATT
GmERF#213	077	3.34e <sup>-18</sup>	FNFPSPDRPNI AGGRNMTSPSIQIAAARFANSEPRK ECSGKPVESL
GmERF#225	081	4.72e <sup>-19</sup>	FNFPDPNPPDI AGGTSMTSPSIQIAAAQFANAGPHE GHSGRPEHPP
GmERF#234	087	3.21e <sup>-20</sup>	LNFPPELVAAE GFAADMSAASIRKKATEVGARVDAL HRQHPHALPA
GmERF#238	074	8.70e <sup>-15</sup>	LETNMVLSRD KQPQLSHQEIQEVAVKFANNDHP LEEEEWDWDT
GmERF#239	121	2.35e <sup>-21</sup>	EEGNSVQHG AAAGNMSADSIKKATQVGARVDAL QTALHHHASS
<b>Consenso</b>			A[GQ]G[AG][DNS]M[ST][PA][ADS][SQ]I[QR][KR][KVA]A [AT][QE][VF][AG][NA][RS]V[DP][AD][LN]

Tabela S29. Motivo II-2 (Motivo EAR)

Nome Genérico	Início	Valor $p$	Sítio
GmERF#004	149	3.96e <sup>-13</sup>	GGGSGDFPVR VVDLNKMPEPE SSDCEWDVN
GmERF#046	149	1.20e <sup>-13</sup>	ANSSHVSSD KPDLNEYPKPE D
GmERF#066	130	1.38e <sup>-12</sup>	HAPPAGEFAD RVDLNKIPEPE NSDCDYWGDD
GmERF#074	148	1.20e <sup>-13</sup>	SISSSHVSYE KPDLNEYPKPE D
GmERF#151	134	3.96e <sup>-13</sup>	GGGSGDFAVR VVDLNKMPEPE SSDCEWNVN
GmERF#192	160	1.50e <sup>-12</sup>	HAPSTNSLNL KPDLNEFPKLE ELQD
GmERF#234	127	4.40e <sup>-13</sup>	HALPAGEFAD RVDLNKMPEPE NSDCDYWDRD
GmERF#239	160	1.50e <sup>-12</sup>	HHASSTNSL KPDLNEFPKLE SLQD
<b>Consenso</b>			[KRV][PV]DLN[EK][MFY]P[EK][PL]E

Tabela S30. Motivo III-1

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#002	136	4.96e <sup>-20</sup>	FLNFPNLAQD LPRPTTTS PKDIQAAAAKAAT TTFDDEIVN
GmERF#003	125	2.55e <sup>-20</sup>	ILNFPELAAS LPRPASNS PRDVQAAAAKAAS MEAPPRTPPS
GmERF#006	125	4.26e <sup>-13</sup>	CLNFADSRWR LTVPATNAEEI RRAAGEAAE AFAVADGDDV
GmERF#007	076	5.36e <sup>-13</sup>	VLNLPHSVSK YQMLTNS PADIRSAASAAAA MIKAETETEA
GmERF#010	090	1.03e <sup>-14</sup>	RLNFPPELVET LPRPTSSK PEDVQAAQAAV MFRISPTMSP
GmERF#011	065	2.41e <sup>-15</sup>	RLNFPPELAST LPRPVSNNADH IRMAAQAAAL RLRTNPAAPD
GmERF#035	177	1.27e <sup>-19</sup>	YLNFPELAQE LPRPASTS PKDIQAAASKAAN TAFEVVKHCQ
GmERF#041	085	3.31e <sup>-17</sup>	ELNFPNSASS LPVPTSSAARDIQMAAASAAA AVGAANDALE
GmERF#043	089	1.86e <sup>-16</sup>	QLNFPDDVDS LPLPSSRTARDIQ TAAQAAAR MMKASGNDEK
GmERF#050	091	1.06e <sup>-20</sup>	ILNFPHLANS LPRPASLAP RDVQAAAAKAAS H MDPSSLSSLV
GmERF#060	077	2.65e <sup>-15</sup>	VLNFPDSVGG YPVPAKSK PADIRTAATAAAE LMKPEASHNN
GmERF#061	084	1.03e <sup>-14</sup>	CLNFADSARW LPLPASTNAKE IRRVAAAAAV AIAAEDSCGE
GmERF#062	121	3.96e <sup>-22</sup>	YLNFPELAQE LPRPASTS PKDIQAAAAKAAA LDFGHRSHDA
GmERF#071	085	1.48e <sup>-19</sup>	ILNFPPELASS LPRPDSNS PRDVQAAAAKAAA MESSWSLSLA
GmERF#078	097	1.06e <sup>-20</sup>	ILNFPHFANS LPRPASLAP RDVQAAAAKAAS H MDPSSLSSLV
GmERF#088	097	3.07e <sup>-18</sup>	ILNFPPEIADL LPRPVTCS PRDIQTAAATAAS MVKFDPTQS
GmERF#093	101	1.61e <sup>-18</sup>	LNFPDEVHRL PPLPSSCTARDIQ AAAAAKAAS H MMMVQAASAD
GmERF#097	156	4.15e <sup>-21</sup>	YLNFPELAQE LPRPVSTSPKDIQ AAAAAKAAN TAFEVVRHCQ
GmERF#103	109	1.48e <sup>-19</sup>	ILNFPELAAS LPRPDSNS PRDVQAAAAKAAA MEVNDVPTTP
GmERF#107	139	1.24e <sup>-18</sup>	ILNFPPEIADL LPRPVTCA PRDIQAAATAAS MVKFDPTQS
GmERF#119	127	6.45e <sup>-18</sup>	CLNFADSARW LPVPATAEAKDIQ KAAAAEAAQ AFRPDQTLKN
GmERF#120	116	2.55e <sup>-20</sup>	ILNFPELAAT LPRPASNS PRDVQAAAAKAAS MEAPPRTPPS
GmERF#121	078	1.86e <sup>-16</sup>	TLNFPDLVHS LPRPLSSRRDIQS AAAEAAAL HSPPPKEEQE
GmERF#130	063	1.48e <sup>-19</sup>	GLNFPDSASS LPVPASLSARDIQ VAAAAAAA AAGAANDAMR
GmERF#141	106	4.66e <sup>-16</sup>	PVNFPSILS YPIPASLSSTDIRAAAAAAAQ ARIVRAPQES
GmERF#142	099	5.71e <sup>-18</sup>	CLNFADSASR LPVPATAEARDIQ KAAAAEAAE AFRPGKDDGA
GmERF#144	065	3.19e <sup>-15</sup>	RLNFPPELAST LPRPVSNNADH IRMAAEAAAL RLRTNPAAPN
GmERF#148	093	3.58e <sup>-20</sup>	FLNFPPELASH LPRPLTTS PKDIQAAAAKAAA LNFEAQVEAQ
GmERF#153	115	1.08e <sup>-10</sup>	NFNFPDSVSL LPLAKRSR SAVDIREAAKATTT AEAFIPNTVF
GmERF#156	126	1.69e <sup>-10</sup>	LNFPNSVSL PIANSLSSRKDIRAAAVEAAE SLRPISNSHH
GmERF#157	078	3.96e <sup>-12</sup>	LFNFPDSVSL LPVAKSSAADVRVAASKVST IFCPSYSSVT
GmERF#172	063	1.48e <sup>-19</sup>	GLNFPDSASS LPVPASLSARDIQ VAAAAAAA AAGAAKDSMR
GmERF#173	081	3.56e <sup>-10</sup>	NFPHAATSLP RLNSRTCSVRAIQFAATQAAE KHFSCAESQQ
GmERF#184	090	7.49e <sup>-12</sup>	VFNFPDSVSL LPVANSSAADIRLAASKVSS VFGPSSSSSS
GmERF#185	108	5.01e <sup>-21</sup>	VLNFPNIVNM LPRPVTCA PRDIQAAAAEAAA MVEFDRESSE
GmERF#194	112	2.15e <sup>-20</sup>	ILNFPELAAS LPRPVSLAP RDVQAAAAKAAS H MDLPSTSSSI
GmERF#202	111	5.01e <sup>-21</sup>	ILNFPNIVNM LPRPVTCA PRDIQAAAAEAAA MVEFDREGSE
GmERF#203	085	5.05e <sup>-18</sup>	ELNFPNSASS LPVPASSAARDIQMAAASAAA AVGAANDALE
GmERF#215	101	2.00e <sup>-15</sup>	LNFPDEVHRL PLLPSACTARDIQ AAAAAKAAS H MMIVQAASVT
GmERF#223	127	4.46e <sup>-18</sup>	CLNFADSTWR LPIPATANAKDIQ KAAAAEAAE AFRPSQTLEN
GmERF#228	074	1.50e <sup>-15</sup>	VLNFPDSVGR YVPASNSPS DIRFAATAAAE LMKPEANNNN
GmERF#229	090	1.03e <sup>-14</sup>	CLNFADSARW LPLPASTNAKE IRRVAAAAAV AIAAEDSRGK
GmERF#230	124	6.84e <sup>-20</sup>	YLNFPELAAG LPRPASTS PKDIQAAAAKASA LDFGHQSHEA
GmERF#235	076	1.41e <sup>-18</sup>	ILNFPELAAL LPRPVSLAS RDVQAAAAKAAS H MDLPSSSSSI
GmERF#257	085	1.68e <sup>-16</sup>	ELNFPNSASS LPIPASSAA HDIQMAAALAAT AVGAANDALE
GmERF#261	090	6.45e <sup>-18</sup>	HLNFPDEVDS LPLPSSRSARDIQ AAAAAQAAAR MMKASGNDEK
GmERF#266	116	5.71e <sup>-18</sup>	CLNFADSASR LPVPATAEARDIQ KAAAAEAAE AFRPGKDDDA
GmERF#267	090	3.70e <sup>-17</sup>	PLNFPNSILS YPIPASLSATDIRAAAAAAAQ ARIVRAPQEC
Consenso			LP [RV] PA [ST] x [SA] [AP] [RK] D [IV] [QR] AAAA [KAE] A AA

Tabela S31. Motivo III-2 (Motivo LWSY)

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#002	248	4.25e <sup>-10</sup>	SYSSCWHLCA VDSGFRLEEQLWEN F
GmERF#006	210	1.04e <sup>-10</sup>	PCARDGRDWN DVDIFDDDEISLWNF SI
GmERF#007	168	1.30e <sup>-08</sup>	SPPRMSPPPS DNSLEKQNGETLWSY F
GmERF#011	156	1.56e <sup>-08</sup>	MFASNGYSFE DNEWEHMQNDLWDP
GmERF#035	272	2.22e <sup>-08</sup>	CSFSSWHLCA VDSGFRLEESFFWEN Y
GmERF#041	171	1.03e <sup>-12</sup>	SPPRFDNFAA TDYEMYDEDPNLWGF PNY
GmERF#050	174	1.02e <sup>-09</sup>	AWMYQPPFGF DTEQDTGFEGLMWNY
GmERF#060	142	4.07e <sup>-08</sup>	LSPPRMSPPS DDSTQSYVGETLWNF
GmERF#061	195	5.99e <sup>-12</sup>	FVTYVNRDV WNFVEDDAEVSLWSF TI
GmERF#062	215	1.56e <sup>-08</sup>	SAWLVAGAEH IELGFRLEDPFLWES Y
GmERF#078	180	1.02e <sup>-09</sup>	AWMYQPPFGF DTEQDTGFEGLMWNY
GmERF#103	228	9.15e <sup>-10</sup>	EGYFIGDSET ESMILCGFEGSLWQH
GmERF#119	218	1.33e <sup>-10</sup>	THCLGYEYED ADLDAQDAEVSLWNF SI
GmERF#141	189	5.93e <sup>-10</sup>	SPLRITSYSS DDSPGNSDGDNLWSY TL
GmERF#142	169	2.50e <sup>-11</sup>	HCFGSDHEYGS ADVEFDAAEVRLWSY SI
GmERF#144	155	3.14e <sup>-09</sup>	MMFANGYSFE DNEWEDMHNDLWDP
GmERF#153	202	4.64e <sup>-09</sup>	TPPSMKRAMD WDDIGCVIDLTLWTN
GmERF#156	225	1.86e <sup>-08</sup>	FPLLPSMDYY WDDHAYCTDFNLWTD
GmERF#184	179	2.57e <sup>-09</sup>	TPPSMKRAFD WDEVDCETDLTLWTD
GmERF#185	188	1.19e <sup>-08</sup>	PPPMTVGSIC DMMMLGEHECGLFQF GTDHQTCVS
GmERF#194	230	8.99e <sup>-09</sup>	GSDAEFVVP ESGVVTSEFSLWDY
GmERF#202	191	1.19e <sup>-08</sup>	PPPMVGSIG DMMMLGEHECGLFQF ATDHQKTCVS
GmERF#203	171	1.03e <sup>-12</sup>	SPPRFDNFAA TDYEMYDEDPNLWGF PNY
GmERF#223	215	2.70e <sup>-10</sup>	THCIGYEYED ADLDAQDAEVSLWSF SI
GmERF#229	154	5.99e <sup>-12</sup>	KLQCQMRDV WNFVEDDAEVSLWSF TV
GmERF#230	217	7.46e <sup>-09</sup>	SAWLVAGAEH IELGFRLEEPFLWES Y
GmERF#235	195	2.10e <sup>-09</sup>	GSDANEFAVP ESRVVTNFEGLWDY
GmERF#257	171	1.03e <sup>-12</sup>	SPPRFDNFAA TDYEMYDEDPNLWGF PNY
GmERF#266	191	2.01e <sup>-12</sup>	HCFGSDHEYGS ADVEFDAAEVSLWSY SI
GmERF#267	174	5.93e <sup>-10</sup>	SPLRITSYSS DDSPGNSDGDNLWSY TL
<b>Consenso</b>			DDSxRDD [EA] [ED] V [SFN] LWS [FY]

Tabela S32. Motivo III-3

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#002	215	1.05e <sup>-16</sup>	SSTQESSSSH STADEETLFDLPDLFPDGNLFSYSSCW HLCVDSGFR
GmERF#006	173	1.47e <sup>-14</sup>	VMATNDDEVQ EPLQQEEVQDLHDLISIANEPLMSPPPC ARDGRDNDV
GmERF#007	134	4.64e <sup>-28</sup>	YTASTSWFET DFIDEVIFGMPDLLVDMAGMLLSPPRM SPPPSDNSLE
GmERF#035	240	1.49e <sup>-16</sup>	DNTQEPSSSH STNDDTLFDLPDLFPDGNLCSFSSWH LCAVDSGFR
GmERF#041	137	3.08e <sup>-33</sup>	TEEFSGGNLN HFVDEDLIFDMPNVLNMAEGMLLSPPRF DNFAATDYEY
GmERF#060	109	6.20e <sup>-28</sup>	MKPEASHNNN NFMDEEAFSMPSLLVDMAGMLLSPPRM SPPSDSTQS
GmERF#061	158	2.48e <sup>-13</sup>	LRVFCDLDEI TMADAPVFEDMREWLQSMADPELRSPTFV TYVNRDVWN
GmERF#097	218	2.90e <sup>-17</sup>	DNPQEPSSSH STIDDDTLFDLPDLFPDGNLCSFSSSW HLCVDSGFR
GmERF#119	185	1.62e <sup>-17</sup>	TTTATAQGVF YMEEEEQVLDMPPELLRNVMVLSPTHCLGY EYEDADLDAQ
GmERF#130	114	2.09e <sup>-30</sup>	AQENQTGMGN QFVDEDLIFDMPNVLNMAQGMLLSPPPF DIGLEPNSPE
GmERF#141	155	2.53e <sup>-23</sup>	GGQGLSERRE EYIDEDELLNMPNLLDMARGMQVSPRLI TSYSSDDSPG
GmERF#153	168	7.22e <sup>-27</sup>	NIGMHDDGSE MFFDEEAVYNMPGLLDSMAEGLLITPPSM KRAMDWDIG
GmERF#156	191	9.08e <sup>-19</sup>	TKNNLDLSS TFFDEEALFNTPGLLDRMAESLDIFPLLP SMDYYWDDHA
GmERF#157	142	9.80e <sup>-26</sup>	TENNVDEAKS VFFDEEAFYNMPLLLDSMAEGLLITPPSM KRVDWDQVD
GmERF#172	114	2.09e <sup>-30</sup>	GQENPTGMGN QFVDEDLIFDMPNVLNMAQGMLLSPPPF DIGLEPNSPE
GmERF#173	145	1.68e <sup>-21</sup>	FSSKEGGSER FFWDEEEVFNMPPELLNSMAEALIITPPAL ERGFNWVGGE
GmERF#184	145	1.11e <sup>-22</sup>	TENNVDEVKT VFFDEEAFYNMPLVFLDSMAEALLITPPSM KRAFDWDEV
GmERF#203	137	3.08e <sup>-33</sup>	TEEFSGGNLN HFVDEDLIFDMPNVLNMAEGMLLSPPRF DNFAATDYEY
GmERF#223	182	2.90e <sup>-17</sup>	TITEQKRGMF YTEEEEQVLDMPPELLRNVMVLSPTHCLGY EYEDADLDAQ
GmERF#228	130	1.91e <sup>-28</sup>	SNDVLPFSET EFMDEEAFSMPSLLVDMAGMLLSPPRM SPPSDNE
GmERF#257	137	3.08e <sup>-33</sup>	AEEFSGGNLN HFVDEDLIFDMPNVLNMAEGMLLSPPRF DNFAATDYEY
GmERF#267	140	2.08e <sup>-24</sup>	GGQGMSEERE EYIDEDELLNMPNLLDMARGMQVSPRLI TSYSSDDSPG
<b>Consenso</b>			xFVDE [ED] [AL] [ILV] F [DN] MPNLL [VD] [DNS] MAEG [ML] L [LI] SP PR [MF]



Tabela S33. Motivo III-4

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#002	018	3.46e <sup>-13</sup>	NLFGTFSSST <b>TSIT</b> ASSSSSPTSSTTSSNS CTKHDSKISP
GmERF#003	024	2.62e <sup>-12</sup>	DSTRNTNTNT <b>PS</b> SSSSSLSPSPSSSSSK KRARDNNNSR
GmERF#035	043	1.79e <sup>-13</sup>	FATSTPSIFF <b>D</b> TNTSSTSSSSSSSSSSSTS PTSNKDSC TK
GmERF#062	014	1.01e <sup>-16</sup>	EESNPLDNEV <b>Q</b> PTSSSSFTITTPSTSSSS IEEATNTTKE
GmERF#088	04	2.76e <sup>-08</sup>	MTQ <b>Q</b> TTSSSS <b>HS</b> GPRAQAQKQSK RPRDCSKHPV
GmERF#097	033	7.55e <sup>-12</sup>	TTPSIFFDN <b>TS</b> STSSSSTSPS <b>NS</b> ESCT <b>KE</b> HSKTPQNDST
GmERF#103	138	8.05e <sup>-12</sup>	AAMEVNDVPT <b>TP</b> SSSSASLSLSQSSWSS LAVSSSSGDG
GmERF#107	027	1.06e <sup>-14</sup>	FTHLQITMTQ <b>Q</b> ATTSS <b>TT</b> SSSSSSSSSSS TFSPSQHSGP
GmERF#120	012	2.62e <sup>-12</sup>	VEPPAAAATS <b>ES</b> DTT <b>GN</b> TNT <b>PS</b> SSSSSSSS SKKRARDNNN
GmERF#141	020	1.52e <sup>-12</sup>	PPPPIQVPDN <b>L</b> AISSSSPTPTTSS <b>PS</b> PT <b>PS</b> GGFYRGRTRC
GmERF#153	013	1.54e <sup>-08</sup>	FEIINESPDS <b>TS</b> LSS <b>PS</b> SLSPCT <b>PP</b> NS <b>N</b> LS DSPNPQPSRK
GmERF#157	099	4.55e <sup>-10</sup>	VRVAASKVST <b>I</b> FC <b>PS</b> YSSVTLSSSS <b>PS</b> NS <b>RV</b> KAKPCLIDDF
GmERF#185	001	1.42e <sup>-12</sup>	<b>M</b> TE <b>T</b> AR <b>CH</b> ET <b>PS</b> SST <b>S</b> NS <b>S</b> TS KPLKTETPKQ
GmERF#194	138	9.37e <sup>-13</sup>	AKAAHMDLPS <b>T</b> SSSISSSLVPTTQ <b>S</b> ASSSSS SLSSLTSADL
GmERF#202	001	2.23e <sup>-13</sup>	<b>M</b> TE <b>T</b> TR <b>CH</b> ET <b>PS</b> SST <b>S</b> TS <b>S</b> NS KSKPPKRKNT
GmERF#223	005	1.26e <sup>-08</sup>	MYTL <b>N</b> SSYLYHV <b>SP</b> ELSSSLDSSS PASEGSRGVA
GmERF#230	014	1.01e <sup>-16</sup>	EESNPLDNEV <b>Q</b> PTSSSSFTITTPSTSSSS IEEATNTTKE
GmERF#235	105	1.21e <sup>-10</sup>	AHMDLPSSSS <b>S</b> ISLS <b>LG</b> STTPSSSSSSLSL ASTADLLLLS
GmERF#267	004	1.52e <sup>-12</sup>	PDN <b>L</b> AISSSSPTPTTSS <b>PS</b> PT <b>PS</b> GGVRYRGRTRY
Consenso			[TQ] [ST] [ST] [TS] SSSSx [TS] P [ST] [ST] SSSSSSS

Tabela S34. Motivo III-5

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#003	149	5.05e <sup>-16</sup>	AAAKAASMEA <b>PER</b> T <b>PP</b> SP <b>AAT</b> PE <b>DDL</b> GE <b>IV</b> EL <b>P</b> PL <b>G</b> TS <b>F</b> DSPDRSSEFV
GmERF#043	119	7.05e <sup>-22</sup>	RMMKASGNDE <b>K</b> SG <b>IA</b> S <b>DD</b> GD <b>SG</b> CD <b>DF</b> W <b>GE</b> IE <b>L</b> PE <b>LM</b> D <b>GE</b> CCWGPCAGAS
GmERF#050	118	2.05e <sup>-21</sup>	KAHMDPSSL <b>S</b> SLV <b>S</b> AM <b>L</b> SS <b>AS</b> DE <b>L</b> SQ <b>IE</b> LP <b>S</b> LE <b>ST</b> D DGSVDLKKEF
GmERF#071	115	4.18e <sup>-18</sup>	AMESSSWSSL <b>AV</b> SSSS <b>G</b> ED <b>P</b> ST <b>PE</b> EL <b>GE</b> IV <b>EL</b> PT <b>IL</b> RF <b>W</b> ETGN
GmERF#078	124	2.05e <sup>-21</sup>	KAHMDPSSL <b>S</b> SLV <b>S</b> AM <b>L</b> SS <b>AS</b> DE <b>L</b> SQ <b>IE</b> LP <b>S</b> LE <b>ST</b> D DGSVVLEKEF
GmERF#088	123	1.94e <sup>-23</sup>	TAAAMVKFD <b>F</b> VT <b>Q</b> SS <b>D</b> SE <b>T</b> PE <b>SS</b> EL <b>SE</b> IV <b>EL</b> PN <b>IE</b> D <b>SS</b> VDSTPEFVLV
GmERF#093	136	3.41e <sup>-21</sup>	QAASADSPEK <b>S</b> SS <b>IT</b> SD <b>CD</b> GG <b>DD</b> FW <b>GE</b> IE <b>L</b> PE <b>LL</b> NG <b>K</b> WWASDRHVTP
GmERF#103	160	3.41e <sup>-21</sup>	LSQSSSWSSL <b>AV</b> SSSS <b>G</b> DP <b>ST</b> PE <b>EL</b> GE <b>IV</b> EL <b>P</b> SL <b>G</b> TS <b>F</b> EIPDPGELV
GmERF#107	165	1.94e <sup>-23</sup>	TAAAMVKFD <b>F</b> VT <b>Q</b> SS <b>D</b> SE <b>T</b> PE <b>SS</b> EL <b>SE</b> IV <b>EL</b> PN <b>IE</b> D <b>SS</b> FDSVDSRPEF
GmERF#120	141	8.55e <sup>-18</sup>	AAKAASMEAP <b>P</b> RT <b>PP</b> SP <b>AST</b> PA <b>ED</b> DL <b>GE</b> IV <b>EL</b> P <b>P</b> L <b>G</b> TT <b>T</b> STTWFOCKSQ
GmERF#185	129	6.43e <sup>-20</sup>	IQAAAAEAAA <b>M</b> VE <b>F</b> D <b>R</b> ES <b>SE</b> A <b>ES</b> T <b>EL</b> SE <b>IV</b> EL <b>PN</b> IE <b>S</b> F DSVESRSEFA
GmERF#194	161	1.87e <sup>-19</sup>	QSASSSSSSL <b>S</b> SL <b>T</b> S <b>AD</b> LL <b>P</b> T <b>AS</b> E <b>EL</b> SE <b>II</b> KL <b>P</b> R <b>LE</b> TT <b>S</b> YELGNEFVFM
GmERF#202	132	2.63e <sup>-20</sup>	IQAAAAEAAA <b>M</b> VE <b>F</b> D <b>R</b> EG <b>SE</b> A <b>ES</b> SE <b>EL</b> SE <b>IV</b> EL <b>PN</b> IE <b>S</b> F ESVESRSEFT
GmERF#215	139	8.45e <sup>-19</sup>	SVTADSPEKS <b>S</b> S <b>IT</b> SD <b>CG</b> GG <b>DD</b> FW <b>GE</b> IE <b>L</b> PE <b>LL</b> NA <b>K</b> WWASDHDVTP
GmERF#235	126	2.55e <sup>-19</sup>	SSSSSSLSSL <b>A</b> ST <b>AD</b> LLLL <b>S</b> T <b>AS</b> E <b>EL</b> SE <b>II</b> KL <b>P</b> R <b>LE</b> TT <b>S</b> YELGNEFVFM
GmERF#261	120	6.14e <sup>-22</sup>	RMMKASGNDE <b>K</b> SG <b>IA</b> S <b>DD</b> GD <b>SG</b> CD <b>DF</b> W <b>GE</b> IE <b>L</b> PE <b>LM</b> D <b>GE</b> CCWGPCAGAS
Consenso			[SP] [SV] TxSSDDLxS [AE] S [DE] [ED] [LF] [SGW] [EG] [IE] [IV] E LP [EN] [LI] E [TD] [ST] [FS]

Tabela S35. Motivo III-6

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#006	048	4.85e <sup>-12</sup>	EVILASARPK <b>K</b> R <b>A</b> G <b>R</b> R <b>V</b> F <b>K</b> E <b>T</b> RHPVYRGVRR
GmERF#061	006	2.03e <sup>-12</sup>	SARPK <b>K</b> R <b>A</b> G <b>R</b> R <b>I</b> F <b>K</b> E <b>T</b> RHPVYRGVRR
GmERF#119	050	1.98e <sup>-14</sup>	EVRLAVRHPK <b>K</b> R <b>A</b> G <b>R</b> K <b>K</b> F <b>R</b> E <b>T</b> RHPVYRGVRR
GmERF#142	022	1.98e <sup>-14</sup>	DYFLAASNPK <b>K</b> R <b>A</b> G <b>R</b> K <b>K</b> F <b>R</b> E <b>T</b> RHPVYRGVRR
GmERF#153	038	4.42e <sup>-10</sup>	NSNPLSDSPN <b>Q</b> P <b>S</b> R <b>K</b> K <b>F</b> R <b>E</b> T RHPVYRGVRQ
GmERF#156	049	1.11e <sup>-12</sup>	PSLSQLVSHK <b>R</b> K <b>G</b> R <b>K</b> K <b>F</b> K <b>E</b> T RHPVYRGVRQ
GmERF#157	001	1.89e <sup>-13</sup>	<b>R</b> K <b>A</b> G <b>R</b> K <b>K</b> F <b>R</b> E <b>T</b> RHPVYRGVRQ
GmERF#173	002	1.98e <sup>-13</sup>	<b>K</b> <b>R</b> R <b>A</b> G <b>R</b> K <b>K</b> F <b>H</b> E <b>T</b> RHPVYKGVRO
GmERF#184	013	1.89e <sup>-13</sup>	KKVVVLIPHK <b>R</b> K <b>A</b> G <b>R</b> K <b>K</b> F <b>R</b> E <b>T</b> RHPVYRGVRQ
GmERF#223	050	1.98e <sup>-14</sup>	EVRLAVRHPK <b>K</b> R <b>A</b> G <b>R</b> K <b>K</b> F <b>R</b> E <b>T</b> RHPVYRGVRR
GmERF#229	012	2.03e <sup>-12</sup>	VNTLASAHPK <b>K</b> R <b>A</b> G <b>R</b> R <b>I</b> F <b>K</b> E <b>T</b> RHPVYRGVRR
GmERF#266	039	1.98e <sup>-14</sup>	DFFLAASNPK <b>K</b> R <b>A</b> G <b>R</b> K <b>K</b> F <b>R</b> E <b>T</b> RHPVYRGVRR
Consenso			[KR] [RK] AGR [KR] KF [RK] ET

Tabela S36. Motivo III-7

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#003	180	2.09e <sup>-16</sup>	LPPLGTSFDS PDRSSEFVFLDMDGWPYSHP WATALMMVGT
GmERF#050	150	2.25e <sup>-24</sup>	PSLESTDDGS VDLKKEFVFDVSLDAWYQPP FGFDTQDTG
GmERF#078	156	1.05e <sup>-22</sup>	PSLESTDDGS VVLEKEFVFDVSLDAWYQPP FGFDTQDTG
GmERF#088	152	1.86e <sup>-20</sup>	VELPNIEDSS VDSSTPEFVLVDVDSWVFPPM GMEFYWFRDR
GmERF#107	197	3.62e <sup>-21</sup>	PNIEDSSFDS VDSRPEFVLVDVDSWVFPPM GMGSEGIEIC
GmERF#185	160	7.87e <sup>-21</sup>	LPNIEESFDS VESRSEFALIDSVDVSWVYPPP MTVGSICDMM
GmERF#194	190	6.05e <sup>-25</sup>	IKLPRLETTS YELGNEFVFMDSQDAWYQPP IPWLHTTYDG
GmERF#202	163	5.07e <sup>-21</sup>	LPNIEESFES VESRSEFTLIDSVDVSWVYPPP MAVGSIGDMM
GmERF#235	155	2.79e <sup>-24</sup>	IELPRLETTS YELGNEFVFMDSQDTWYQPP IPWLHTTYDG
<b>Consenso</b>			[VY] [DE] [LS] [RG] [SKNP]EFV[FL] [VIM]D[SV] [VLQ] D[ASG]W[MV] [YF] [PQ] P[PM]

Tabela S37. Motivo IV-1

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#024	028	8.58e <sup>-44</sup>	RKSRSRGDGS KSVAE <sup>T</sup> IAK <sup>W</sup> KEY <sup>N</sup> EHLYSGKDD <sup>S</sup> RTTRKAPAKGSKK <sup>G</sup> CMK GKGGPQNSQC
GmERF#087	015	5.04e <sup>-46</sup>	VKNHNKGDGS KSLAD <sup>T</sup> LAK <sup>W</sup> KEY <sup>N</sup> AWLESNN <sup>E</sup> A <sup>E</sup> K <sup>P</sup> ARKVPAKGSKK <sup>G</sup> CMK GKGGPENSRC
GmERF#154	015	3.81e <sup>-47</sup>	VKNHNKGDGS KSLAD <sup>T</sup> LAK <sup>W</sup> KEY <sup>N</sup> AWLESNN <sup>E</sup> A <sup>E</sup> K <sup>P</sup> VRKVPAGS <sup>K</sup> KK <sup>G</sup> CMK GKGGPENLRC
GmERF#158	013	1.32e <sup>-46</sup>	AKTHNKGDGS KSLAK <sup>I</sup> LAK <sup>W</sup> KEY <sup>N</sup> AQIDSSSDADK <sup>P</sup> VRKVPAGS <sup>K</sup> KK <sup>G</sup> CMK GKGGPENSRC
GmERF#183	013	2.59e <sup>-46</sup>	AKAHNKGDGS KSLAK <sup>I</sup> LAK <sup>W</sup> KEY <sup>N</sup> AQIDSSSDADK <sup>P</sup> IRKVPAGS <sup>K</sup> KK <sup>G</sup> CMK GKGGPENSRC
GmERF#189	028	1.01e <sup>-41</sup>	RKSRSRGYGT GSVAE <sup>T</sup> IAK <sup>W</sup> KEY <sup>N</sup> EHLYSGKDD <sup>S</sup> RTTRKAPAKGSKK <sup>G</sup> CMK GKGGPQNSQC
<b>Consenso</b>			KS [LV]A [DEK] [TI] [LI]AKWKEYN [AE] [HQW] [LI] [DEY] S [GNS] [KNS] [DE] [AD] [DES] [KR] [PT] [TV]RK [VA] PAKGSKK <sup>G</sup> CMK

Tabela S38. Motivo IV-2

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#022	121	6.84e <sup>-24</sup>	ISMFP <sup>S</sup> RGGL LNLNAQ <sup>P</sup> SVH <sup>V</sup> IHQ <sup>L</sup> QELK <sup>Q</sup> NGVVDLENFD
GmERF#025	141	2.26e <sup>-19</sup>	PFNSAPAT <sup>T</sup> N H <sup>N</sup> INTV <sup>Q</sup> VH <sup>H</sup> PH <sup>H</sup> YR <sup>C</sup> PE <sup>V</sup> VV QVPHSESDLG
GmERF#036	121	4.72e <sup>-26</sup>	FISMFP <sup>S</sup> CGL LNVNAQ <sup>P</sup> SVH <sup>L</sup> IHQ <sup>R</sup> LQELK <sup>R</sup> NSVVSQSPFN
GmERF#131	095	7.57e <sup>-26</sup>	ISMFP <sup>S</sup> SGGL LNLNAQ <sup>P</sup> SVH <sup>V</sup> IHQ <sup>R</sup> LQELK <sup>Q</sup> NGVVATSONQ
GmERF#190	150	6.32e <sup>-19</sup>	SPPVLC <sup>P</sup> NDN H <sup>N</sup> NNIV <sup>Q</sup> VH <sup>H</sup> PH <sup>H</sup> YR <sup>C</sup> PE <sup>L</sup> LV QVQVAHTLST
GmERF#253	119	4.72e <sup>-26</sup>	FISMFP <sup>S</sup> CGL LNVNAQ <sup>P</sup> SVH <sup>L</sup> IHQ <sup>R</sup> LQELK <sup>R</sup> NSVVSQSSSS
<b>Consenso</b>			[LH]N [LV]NA [QV] [PQ] [SV] [VH]H [LPV] [IH]HQ <sup>R</sup> [LC] [QP]E [LE]K [Q <sup>R</sup> V]

Tabela S39. Motivo IV-3

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#022	162	2.37e <sup>-54</sup>	IQNHA <sup>E</sup> ILPP LQEGNS <sup>R</sup> VSSQ <sup>K</sup> TIVGDFE <sup>E</sup> K <sup>P</sup> QID <sup>L</sup> LEFLQ <sup>Q</sup> MGILKE <sup>E</sup> REEE <sup>K</sup> TN <sup>S</sup> SGS <sup>S</sup> TTV <sup>S</sup> E <sup>A</sup> ASRDESSEQA
GmERF#036	170	3.37e <sup>-61</sup>	TQTVGSK <sup>K</sup> DG ENPPK <sup>D</sup> QML <sup>S</sup> EEV <sup>L</sup> GDLQ <sup>E</sup> K <sup>P</sup> QID <sup>L</sup> HEFLQ <sup>Q</sup> MGILKE <sup>E</sup> RES <sup>E</sup> RTD <sup>S</sup> SGS <sup>S</sup> STM <sup>P</sup> E <sup>A</sup> LLRDDNDHHL
GmERF#131	159	4.73e <sup>-54</sup>	QNHA <sup>E</sup> TLVPP LQEGNS <sup>Q</sup> VSSQ <sup>K</sup> TIIGDFE <sup>E</sup> K <sup>P</sup> QID <sup>L</sup> LEFLQ <sup>Q</sup> MGILKE <sup>E</sup> REEE <sup>K</sup> TN <sup>S</sup> SGS <sup>S</sup> TTV <sup>S</sup> E <sup>A</sup> VSRDDQSEQQ
GmERF#197	147	7.49e <sup>-63</sup>	NVDSKNH <sup>G</sup> ED ENPPK <sup>D</sup> VQ <sup>T</sup> SLEE <sup>V</sup> LGD <sup>L</sup> Q <sup>E</sup> K <sup>P</sup> QID <sup>L</sup> HEFLQ <sup>Q</sup> MGILKE <sup>E</sup> RQ <sup>S</sup> KRTD <sup>S</sup> SGS <sup>S</sup> TV <sup>C</sup> E <sup>A</sup> V LTDDCDNLGV
GmERF#253	170	4.32e <sup>-64</sup>	NVDSKNH <sup>G</sup> ED ENPPK <sup>D</sup> VQ <sup>T</sup> SS <sup>E</sup> EV <sup>L</sup> GDLQ <sup>E</sup> K <sup>P</sup> QID <sup>L</sup> HEFLQ <sup>Q</sup> MGILKE <sup>E</sup> RQ <sup>S</sup> ERTD <sup>S</sup> SGS <sup>S</sup> TV <sup>R</sup> E <sup>A</sup> V LTDDCDHLGV
<b>Consenso</b>			[EL] [NQ] [PE] [PG] [KN] [DS] [VDQR] [QV] [STM] [SL] [QSL] [EK] [ET] [VI] [LIV]G [DG] [LDF] [QEF] EK <sup>P</sup> QID <sup>L</sup> [HL]EFLQ <sup>Q</sup> MGILKE <sup>E</sup> R [EQ] [SE] [EK] [RK] T [DN] SSGS [TS] [TV] [VCMR] [ESP] [EA] [AV]

Tabela S40. Motivo IV-4

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#022	256	7.35e <sup>-25</sup>	HEHDHG <sup>V</sup> ADS FYRASE <sup>E</sup> IQ <sup>F</sup> EAY <sup>D</sup> IN <sup>E</sup> DLT <sup>F</sup> S <sup>P</sup> SIW <sup>N</sup> Y
GmERF#036	257	2.33e <sup>-36</sup>	EALIDMH <sup>G</sup> IA GIQ <sup>E</sup> SEV <sup>T</sup> Q <sup>L</sup> EAY <sup>D</sup> PN <sup>D</sup> DLT <sup>F</sup> S <sup>T</sup> SIW <sup>N</sup> F
GmERF#197	232	5.10e <sup>-36</sup>	EALIE <sup>M</sup> DGLA GIQ <sup>E</sup> SEV <sup>T</sup> Q <sup>L</sup> EAY <sup>E</sup> PN <sup>D</sup> DLT <sup>F</sup> S <sup>T</sup> SIW <sup>N</sup> F
GmERF#253	255	5.10e <sup>-36</sup>	EALIE <sup>M</sup> HGLA GIQ <sup>E</sup> SEV <sup>T</sup> Q <sup>L</sup> EAY <sup>E</sup> PN <sup>D</sup> DLT <sup>F</sup> S <sup>T</sup> SIW <sup>N</sup> F
<b>Consenso</b>			[GF] [IY] [QR] [EA]SE [VE] [TI]Q [LF]EAY [DE] [PI]N [DE]DLT <sup>F</sup> S [TP]SIW <sup>N</sup> [FY]

Tabela S41. Motivo IV-5

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#022	099	4.31e <sup>-22</sup>	IFTNSTAT <sup>K</sup> S HKY <sup>F</sup> KW <sup>F</sup> PSK <sup>N</sup> FISM <sup>F</sup> PS <sup>R</sup> GG LLNLNAQ <sup>P</sup> SV
GmERF#036	100	5.37e <sup>-28</sup>	PRSTST <sup>T</sup> TITS SGK <sup>F</sup> KW <sup>F</sup> PSK <sup>N</sup> FISM <sup>F</sup> PS <sup>C</sup> GL LNVNAQ <sup>P</sup> SVH
GmERF#197	087	5.37e <sup>-28</sup>	LQPMST <sup>T</sup> STIM SGK <sup>F</sup> KW <sup>F</sup> PSK <sup>N</sup> FISM <sup>F</sup> PS <sup>C</sup> GL LNRLQ <sup>E</sup> LKRN
GmERF#253	098	5.37e <sup>-28</sup>	LQPRST <sup>T</sup> STIT SGK <sup>F</sup> KW <sup>F</sup> PSK <sup>N</sup> FISM <sup>F</sup> PS <sup>C</sup> GL LNVNAQ <sup>P</sup> SVH
<b>Consenso</b>			[SH] [GK] [KY]FKW <sup>F</sup> PSK <sup>N</sup> FISM <sup>F</sup> PS [CR]G [LG]

Tabela S42. Motivo IV-6

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#045	161	7.35e <sup>-40</sup>	NEVEFR <sup>P</sup> SWA TMNEGL <sup>P</sup> V <sup>F</sup> DD <sup>S</sup> IWAE <sup>A</sup> AAMSL <sup>D</sup> FP <sup>H</sup> IAA <sup>E</sup> TAI <sup>Y</sup> ASGNNLADVG
GmERF#073	196	2.46e <sup>-39</sup>	NDGDFR <sup>P</sup> SWG TMNEGL <sup>P</sup> V <sup>F</sup> DD <sup>S</sup> IWAE <sup>A</sup> AAMSL <sup>D</sup> FP <sup>Q</sup> IAA <sup>E</sup> TAI <sup>Y</sup> SSGNNLADVG
GmERF#191	221	4.49e <sup>-36</sup>	DNMFRS <sup>F</sup> WGP TMDD <sup>T</sup> MP <sup>V</sup> FD <sup>Q</sup> SIW <sup>T</sup> E <sup>A</sup> AAMSL <sup>D</sup> L <sup>P</sup> IIAD <sup>N</sup> GFY <sup>G</sup> RGVGGNKG <sup>G</sup>
GmERF#240	185	5.03e <sup>-32</sup>	DDMFRP <sup>F</sup> WGP TMDD <sup>T</sup> MP <sup>V</sup> S <sup>D</sup> ESIW <sup>T</sup> E <sup>A</sup> AAMSL <sup>D</sup> Y <sup>P</sup> IVGGGG <sup>F</sup> G NNRKFF <sup>F</sup> G <sup>E</sup> A
<b>Consenso</b>			TM [DN] [DE] [GT] [LM]PV [FS]D [DEQ]SIW [AT]EAAMSLD [FLY]P [IHQ] [IV] [AG] [ADG] [EGN] [GT] [AFG] [IFY] [GY]

Tabela S43. Motivo IV-7

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#087	159	7.43e <sup>-40</sup>	PAASQSGHAL MILESSECMILPNNSSGGDAEEDDDMEDLSLSLSVSRVKE EGEDESGTSS
GmERF#154	159	5.17e <sup>-39</sup>	PSAKESGSAL VILERSECMMLWNNSSGGDAEEDDGMEDLSLSLSVSRKEEGE DESGTSSSYL
GmERF#158	146	2.11e <sup>-40</sup>	SPVANHCSS MAVSANESMISPSNSGVGAEDDVEDMEPIISLSLTVKENGES GESGISSSPP
GmERF#183	146	1.32e <sup>-44</sup>	SPSANHCSS MAVSANESMISPSNSGVDAEEDVDMEPIISLSLSVSRKEENGES GESGISSSPP
<b>Consenso</b>			[MV] [AI] [LV] [ES] [ARS] [NS] E [CS] M [IM] [LS] [PW] [NS] NSG [GV] [DG] A [AE] [ED] D [DV] [DG] ME [DP] [IL] SLSL [ST] VK [HR] [EV] [NEK] [GH] E

Tabela S44. Motivo IV-8

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#045	008	4.97e <sup>-26</sup>	MLKSGIG IERKQLKKPAQASSRKGCMR GKGGPENASC
GmERF#073	008	7.98e <sup>-27</sup>	MLKSGMG IERKQVKKPAQASSRKGCMR GKGGPENASC
GmERF#191	003	4.83e <sup>-27</sup>	MG IQERKKVKKPAQASSRKGCMR GKGGPENAKC
GmERF#240	003	4.83e <sup>-27</sup>	MG IQERKKVKKPAQASSRKGCMR GKGGPENATC
<b>Consenso</b>			I [EQ] ERK [KQ] [VL] KKPAQASSRKGCMR



Tabela S45. Motivo V-1

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#048	165	8.16e <sup>-19</sup>	VENAQQFKPL EEDHIEQMIPELLHY GSIELCSVIP
GmERF#049	158	1.10e <sup>-16</sup>	VEVEVEEEEE EDRVALQMIPELLNW N
GmERF#053	192	1.50e <sup>-15</sup>	QELLRTEMDE EERIALQMIPELLSR NCPTSPSFNN
GmERF#076	166	8.16e <sup>-19</sup>	VENAQQFKPL EEDHIEQMIPELLHY GSIELCSVIP
GmERF#077	170	1.10e <sup>-16</sup>	GATENNEVEE EDRVALQMIPELLNW N
GmERF#089	169	8.16e <sup>-19</sup>	VSNQQFQPVV EDDHIEQMIPELLDY GSIELCPVSS
GmERF#091	174	6.54e <sup>-16</sup>	KPEESTGLDE EQKMALQMIPELLNR N
GmERF#108	185	6.54e <sup>-16</sup>	PEEESTGLDE EQKMALQMIPELLNR N
GmERF#110	170	8.16e <sup>-19</sup>	VSHQQFQPVV EDDHIEQMIPELLDY GSIELCSVGL
GmERF#174	198	2.15e <sup>-16</sup>	QEVVRGEIDE EERIALQMIPELLND RNCPSPSFDN
GmERF#186	154	7.55e <sup>-17</sup>	AQQQQFQPVV QDDHIEQMIPELLDY GSMELSFVDS
GmERF#193	171	2.31e <sup>-19</sup>	ENVPLQFKPL EEDHIEQMIPELLDY GSIELCSVIP
GmERF#200	169	7.55e <sup>-17</sup>	AQQQQFQPVV QDDHIEQMIPELLDY GSMEISFVHS
GmERF#201	161	3.78e <sup>-16</sup>	KVESKNGLDE EQKIALQMIPELLNR N
GmERF#226	200	2.15e <sup>-16</sup>	QEVVRGEIDE EERIALQMIPELLND RNCPSPSFNN
GmERF#236	177	5.18e <sup>-16</sup>	VGHGDYGYDE EDRIAMQMIPELLNW NHPCGSSN
GmERF#237	120	2.31e <sup>-19</sup>	ENVPLQFKPL EEDHIEQMIPELLDY GSIELCSVIS
Consenso			E [DE] [DR] [HI] [AI] [EL] QMI [EQ] ELL [ND] [YR]

Tabela S46. Motivo V-2

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#019	248	1.21e <sup>-18</sup>	VSESGSHQNN PSWGFEPSPSQLPDFERIKVERQISASLY AINGVQEYMD
GmERF#048	067	4.82e <sup>-28</sup>	TRARTNFPYN PNASQSSSSKLLSATLTAKLHRCYMASLQ MTRPSSQAEF
GmERF#076	067	4.82e <sup>-28</sup>	TRARTNFPYN PNASQSSSSKLLSATLTAKLHRCYMASLQ MTRPSSQPET
GmERF#089	067	8.45e <sup>-29</sup>	PKARTNFPYN PNEPHSSSSKLLSATLTAKLHRCYMASLS LQMAKQKPPQ
GmERF#110	067	8.45e <sup>-29</sup>	SKARTNFPYN PNEPHSSSSKLLSATLTAKLHRCYMASLS LQMAKQKPPQ
GmERF#162	230	2.70e <sup>-23</sup>	EMEEEPVEPE NENYGDPEPSMLRAVMKRMKYERKFSASLY AFNGIPECLK
GmERF#186	067	2.14e <sup>-27</sup>	PKARTNFPYN PNEPLSSSSKLLSATLTAKLHRCYMASLS LQMSKQKQSQ
GmERF#193	069	2.46e <sup>-23</sup>	ARTNFPNPN VPHSSSSSSKLLSATLTAKLHRCYMASLQ ITRQALECES
GmERF#200	083	1.77e <sup>-28</sup>	PKARTNFPYN PNEPQSSSSKLLSATLTAKLHRCYMASLS LQMSKQKQSQ
GmERF#216	215	2.47e <sup>-19</sup>	AQILDIPKID DGLLTNAADLLESEFERMKVERQISASLY AINGVQEYME
GmERF#221	253	3.13e <sup>-21</sup>	FIGSESGSQN NPSWGFEPSPSQLPDFERIKVERQISASLY AINGVQEYMD
GmERF#247	183	1.05e <sup>-25</sup>	IAEDIEEQLE PENYGDPEPSMLREVMKRMKYERKFSASLY AFNGIPECLK
GmERF#249	190	5.69e <sup>-23</sup>	EMEEEPVEPE NENYDDEPSMLRAVMKRMKYERKFSASLY AFNGIPECLK
GmERF#252	195	8.02e <sup>-20</sup>	MAQIFDIPKI DDGLTNAADLLESEFERMKVERQISASLY AINGVQEYME
GmERF#263	200	5.93e <sup>-19</sup>	AEEMVGSIEE EEKYRVDDSELGRESFRMKYERKFSACLY TLIGVSECLS
Consenso			[PN] [NE] [EN] [PYS] [GQ] [SD] [SE] [SP] S [KM] L [LR] S [AEV] [TFM] [LEK] [RT] [AM] K [LVY] [EH] [RK] [CKQ] [FHIY] [SM] AS L [YSQ]

Tabela S47. Motivo V-3

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#019	091	1.88e <sup>-21</sup>	GSNTRTNFIT RVSLDSPLASRIQNLVNNRKGTKTKQ EQYVVETCDT
GmERF#067	066	2.89e <sup>-17</sup>	TRTNFSTQGC G VATNSPISLKLRLNLLDRKAMSNQSQ SADTSASPNV
GmERF#104	070	9.03e <sup>-17</sup>	TRTNFSTHKY GIATNSPISLKLRLNLLDRKAMSNQSQ IQIQRQSSST
GmERF#162	078	4.28e <sup>-25</sup>	NTRTNFWPCS QSSSTPALSSKITNLLLQRLKERNNN NTCSSSSSSS
GmERF#216	083	6.03e <sup>-23</sup>	GSNTRTNFIT HVSLDSPLASRIQNLVNNRKGTKTKQ DTNANANNAP
GmERF#221	093	6.66e <sup>-21</sup>	GSNTRTNFIT RVSLDSPLASRIQNLVNNRKGTKTKH QEQDVETCHT
GmERF#247	083	6.77e <sup>-23</sup>	NTRTNFWPSS QSSSTPALSSKITNLLLQRLKARNNN NNNNPCSLSS
GmERF#249	078	4.28e <sup>-25</sup>	NTRTNFWPCS QSSSTPALSSKITNLLLQRLKERNNN NNNNACSSSS
GmERF#250	078	4.28e <sup>-25</sup>	NMRTNFWPCS QSSSTPALSSKITNLLLQRLKERNNN NNNNACSSSS
GmERF#252	084	7.60e <sup>-23</sup>	GSNTRTNFIT HVSLDSPLASRIQNLVNNRKGTKTKQ DANANANNAP
GmERF#263	076	9.73e <sup>-16</sup>	TRTNFSSQS SPNSVPALPLKIAKLLLLRLKARNMA SSCVICYSSS
Consenso			Q [VS] S [TL] [DS] [SP] [PA] L [SA] [SL] [KR] I [RT] NLL [LN] [QN] R [LK] [KG] E [RK] NN [NQ]

Tabela S48. Motivo V-4

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#049	081	2.78e <sup>-65</sup>	ADHDNNNTFL <b>SEKALSELLSTKLRKYCKDEAPSLTCLRLDADNSHIGVWQKGAGPHSGSNWVMRVELGKK</b> QIITEVEVEV
GmERF#053	090	8.09e <sup>-62</sup>	DHETSSSSTT <b>TSKLELEILHAKLRKCGKLFPSMTCLRLDPENSHIGVWQKRAGQRSDSNWMMVPPPLGK</b> KGGANEGGTG
GmERF#077	082	1.11e <sup>-63</sup>	DDNNNTDTFL <b>SEKALSELLSTKLRKYCKDEAPSLTCLRLDADNSHIGVWQKGAGPHSGSNWVMRLELGGK</b> QITEEGSEST
GmERF#091	080	9.29e <sup>-69</sup>	GNHSSSTSSS <b>STTTLSAVLSAKLRKCKSPSPSLTCLRLDTENSHIGVWQKRAGPRSDSNWIMMVELEKK</b> NNKGPSSESE
GmERF#108	081	1.57e <sup>-66</sup>	GNHSSNTSSS <b>STTTLSAVLSAKLRKCKSPSPSLTCLRLDTENSHIGVWQKRAGPRSDSNWIMMVELENK</b> NNDNEGDHHQ
GmERF#174	084	2.29e <sup>-67</sup>	STNSEDTPST <b>TSKDLEEILHAKLRKCKVSPSPMTCLRLDTENSHIGVWQKRAGRRSDSNWMTVQLGKK</b> TVNNNETHQA
GmERF#201	076	4.44e <sup>-56</sup>	DNQKGNHISS <b>SSPTFSSALNAKLRKCKSLSPSLTCLRLDTENSHIGVWQKRAGPRSESNWIMTVELERS</b> NVNVVSDSQV
GmERF#226	084	7.43e <sup>-68</sup>	STTSEDTPST <b>TSKDLEEILHAKLRKCGKVFSPSPMTCLRLDPENSHIGVWQKRAGRRSDSNWMTVQLGKK</b> SVNVNNDTH
GmERF#236	084	1.50e <sup>-60</sup>	TTTPCVDVFS <b>SENVLSEHLSRKLKRCCKDESPSLTCLRLDADNSHIGVWQKGAGPHSDSNWIVRVEIGKK</b> LHAEGESSV
<b>Consenso</b>			[ST] [SPT] [KT] [TAD] L [SE] [EA] [ILV] L [SH] [AT] KL [RK] K [CY] [CG] K [DSV] P [SA] PS [LM] TCLRLD [TAP] [ED] NSHIGVWQK [RG] AG [PR] [RH] S [DG] SNW [VI] M [MRT] V [EQ] L [GE] KK

Tabela S49. Motivo V-5

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#019	277	3.42e <sup>-18</sup>	VERQISASLY <b>AINGVQEYMDTLQDS</b> NNNNGAIWDF
GmERF#162	259	6.34e <sup>-15</sup>	YERKFSASLY <b>AFNGIPECLKGLADQ</b> LTSLQMACSK
GmERF#216	244	6.97e <sup>-19</sup>	VERQISASLY <b>AINGVQEYMETVQDS</b> NDTLWDLPL
GmERF#221	282	1.47e <sup>-15</sup>	VERQISASLY <b>AINGVQEYMDTVTVQ</b> DSNNNVAIWD
GmERF#247	212	1.38e <sup>-14</sup>	YERKFSASLY <b>AFNGIPECLKLELES</b> GNTNGRGNEE
GmERF#249	219	1.38e <sup>-14</sup>	YERKFSASLY <b>AFNGIPECLKLELES</b> GNMKGRGLAD
GmERF#252	224	6.97e <sup>-19</sup>	VERQISASLY <b>AINGVQEYMETVQDS</b> NETLWDLPL
<b>Consenso</b>			A [IF] NG [VI] [QP] E [YC] [ML] [KDE] [TL] [VKL] [QL] [DE] [SQ]

Tabela S50. Motivo V-6

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#048	123	1.69e <sup>-41</sup>	VITSSASNFS <b>MKSNETETLLLPQKRQEEQEAEGNWWVFKVKVVENAQQFKP</b> LEEDHIEQMI
GmERF#076	124	9.14e <sup>-42</sup>	VITSSASNFS <b>MKSNETETMLLPQKRQEEQEGEGNWWVFKVKVVENAQQFKP</b> LEEDHIEQMI
GmERF#089	125	2.20e <sup>-43</sup>	SNPFASANA <b>AGSSADTSFRWPDNRHEEFRWLEGNWVGVEGQVEVSNQQFQ</b> PVLEDDHIEQ
GmERF#110	126	1.58e <sup>-44</sup>	SNPFANFANVI <b>AGSSADTGFWRWPDNRHEEFRWLEGNWVGVEGQVEVSHQQFQ</b> PVLEDDHIEQ
<b>Consenso</b>			[AM] [GK] S [NS] [AE] [DT] [ET] [TGS] [FLM] [LR] [LW] P [DQ] [KN] R [HQ] EE [EF] [QR] [EW] [LAG] EGNWV [FG] [KV] [EK] [GV] [KQ] VE [NV] [AS] [QHN] Q [FQ] [FK] [PQ]

Tabela S51. Motivo VI-1

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#008	021	2.11e <sup>-26</sup>	HSKLLIPSEE <b>DPPKHPYRLV</b> RITVTDTEATDSSSDEEQ HTSTRHRHRK
GmERF#023	028	2.11e <sup>-26</sup>	SSHFDNDPPC <b>KTKTKTQR</b> LLRIIVTDHDATDSDSDEEE EQQQQTTRVK
GmERF#028	013	7.18e <sup>-25</sup>	YQRKTIITKL <b>VKSSKFLPKV</b> IRISVTDQDATDSSSDEEG EHARVKKVIN
GmERF#042	019	1.03e <sup>-20</sup>	KVVTYKHVKP <b>GDILNAAFKV</b> VVISVMDPYATDDEDEDEG LVIHPRMKKF
GmERF#058	029	3.03e <sup>-24</sup>	LISPKKSSQT <b>NSSTQTE</b> RRIVRISYTDPDATDSSSDEEG FPLVRKRMKR
GmERF#072	025	4.56e <sup>-26</sup>	TKLFMEEDTL <b>LKKKRHYPKL</b> IRIRVTDADATDSSSDDEE PSVSTRRRRV
GmERF#081	027	1.96e <sup>-24</sup>	KKLISPKKSS <b>QTNSTTE</b> RRIVRISYTDPDATDSSSDEEG FPFVRQRMKR
GmERF#102	025	1.75e <sup>-27</sup>	TKLFMEEDTL <b>LKKKRHYPKL</b> IRIRVTDADATDSSSDDEE PSMSSTRRRV
GmERF#147	021	7.56e <sup>-29</sup>	HTKLLIPPEE <b>DPPKHPYRLV</b> RITVTDTEATDSSSDEEE EQTYHCNSST
GmERF#152	041	1.62e <sup>-18</sup>	FDGDPCKTA <b>GKTKPPHTK</b> LLRVILTDDHDATDSDSSGDD EDPKNTKKHK
GmERF#163	028	7.52e <sup>-22</sup>	ETVTKKHKMS <b>KNNTSLSEK</b> VISILMTDQYATDSSSDEER VTSRRRIKRY
GmERF#196	025	2.73e <sup>-22</sup>	QTMVKKHKMS <b>KDNFCSFK</b> VISISLIDPYATDSSSDEEN VMRRRVKRY
GmERF#199	028	2.69e <sup>-23</sup>	SSHFDNDPPC <b>KTKTKTQR</b> LLRIIITDHDATDSDSDEEE QQQQQQKTRR
GmERF#260	020	1.71e <sup>-20</sup>	VVVTYKHVKP <b>GILNAAAFK</b> VVISVMDPYATDDEDEDEG LVNHPRMKKF
GmERF#275	023	2.40e <sup>-28</sup>	KLLIPPGPEE <b>DPPKHPYPKL</b> LRIITVTDTEATDSSSDEEQ TNFSTRHRHR
<b>Consenso</b>			[KDG] [KPT] [KNP] [KT] [HK] [THP] Y [PR] [KR] [LVI] [VIL] RI [SIT] VTD [PHT] [DY] ATDS [SD] SD [ED] [ED] [EG]

Tabela S52. Motivo VI-2

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#008	054	9.65e <sup>-14</sup>	SSDEEQHTST <b>RHRHRKFV</b> NEISIE SSCASENDGA
GmERF#023	062	1.17e <sup>-11</sup>	SDEEEQQQQ <b>TTRVKREITQ</b> INMQ LPLSYNSSFP
GmERF#028	043	8.21e <sup>-13</sup>	TDSSSDEEGE <b>HARVKKVINE</b> IKIV CVNNKKNHKN
GmERF#042	051	1.21e <sup>-14</sup>	DEDEDEGLVI <b>HPRMKKFV</b> NEIRIV EKSSTNCESA
GmERF#058	062	1.80e <sup>-16</sup>	SSDEEGFPLV <b>RKRMRKYVN</b> QIEIE TAAEKVVRKR
GmERF#072	060	3.81e <sup>-15</sup>	DDDEPSVSST <b>RRRVKVFV</b> NEITIQ GGGGGGGCVN
GmERF#081	060	2.09e <sup>-16</sup>	SSDEEGFPFV <b>RQRMKRYVN</b> QIEIE TAAAEKVVRK
GmERF#102	060	9.72e <sup>-15</sup>	DDDEPSMSST <b>RRRVKNFV</b> NEITIQ GGGGGGGDVN
GmERF#147	060	9.65e <sup>-14</sup>	EQTYHCNSST <b>RHRHRKFV</b> NEISIE SCSSSENDGVV
GmERF#152	078	1.10e <sup>-10</sup>	DDEDPKNTKK <b>HKKVKRQIT</b> HTITIN LPFFSKSPTP
GmERF#163	060	7.18e <sup>-14</sup>	SSSDEERVTS <b>RRRIKRYVN</b> RIELQ PAIKSVATRK
GmERF#196	057	2.07e <sup>-14</sup>	SSSDEENVMR <b>RRRVKRYVN</b> RIELQ PAFKPVTRK
GmERF#199	064	6.17e <sup>-11</sup>	DEEQQQQQK <b>TRRVNREITQ</b> INMQ LPLSHNSSFS
GmERF#260	052	1.21e <sup>-14</sup>	EEDEDEGLVN <b>HPRMKKFV</b> NEIRIV EKSSTNSES
GmERF#275	057	3.50e <sup>-12</sup>	SDEEQTNFST <b>RHRHRKLINE</b> ISIA RASDTRRGQK
<b>Consenso</b>			[RH] [RH] R [VMH] [KR] [RK] [FY] [VI] [NT] [EQ] I [EST] I [QEV]

Tabela S53. Motivo VI-3

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#008	193	1.10e <sup>-17</sup>	YASGEETNDN <b>NNNLFSP</b> SVLHSCSLSEE AEVKAESPFP
GmERF#023	209	9.00e <sup>-14</sup>	NLSSDGGASY <b>SDLVASP</b> SVLAYECDSTP FDGFRYLDVE
GmERF#028	178	4.90e <sup>-13</sup>	LECDSVNGT <b>SSACASP</b> SVLVFQPWIEP PLLEETFKEE
GmERF#042	188	1.14e <sup>-14</sup>	AVSGENDSSE <b>ESHLSSP</b> SVLGLQSVELM LDEVSETDLP
GmERF#058	192	1.71e <sup>-17</sup>	SGYDSGDDHC <b>QNLSSP</b> SVLHFRSNESES DQKSEQVLR
GmERF#072	184	6.95e <sup>-16</sup>	TGYCSGEESR <b>NNDLRS</b> PSTVLGCRSVYEE AESVTANATG
GmERF#081	205	4.02e <sup>-17</sup>	SGYDSSEDHC <b>HNLS</b> PSTSVLQFRSNESES ENSESQQKAE
GmERF#102	185	8.83e <sup>-16</sup>	GYCSGEESRN <b>NDDLRS</b> PSTVLGCRSVSEE AVTANDVFGG
GmERF#147	181	4.62e <sup>-17</sup>	PPSGEETHCN <b>SKNIF</b> PSTSVLHCCSLSEE AESVTAKDDD
GmERF#152	207	1.56e <sup>-12</sup>	DTPPPETVFS <b>GEG</b> LSSPSTSVLTYCDGDST PFDGFRYGEV
GmERF#163	216	2.05e <sup>-13</sup>	SCFYDSEDEF <b>CLNLS</b> PSTSVLRFNESAEL EKQYEPFGR
GmERF#196	200	4.20e <sup>-14</sup>	SGCYDSGDEC <b>CLNLS</b> PSTSVLQFSAEPYK PDEPFDPMPW
GmERF#199	208	9.00e <sup>-14</sup>	NLSSDGGASY <b>SDLVASP</b> SVLAYECDSTP FDGFRYLDVD
GmERF#260	187	1.04e <sup>-12</sup>	VSGENDSSEG <b>SS</b> FSPSTSVLKLQVPEIM LEEFSETDLL
GmERF#275	148	1.87e <sup>-13</sup>	ITPLLNEDTH <b>NKFL</b> PSTSVLHVSEKVK VRGCSVPVDS
<b>Consenso</b>			[SN] [DS] NL [SAF] SPTSVL [HG] [FY] [CQR] S [VN] SE [EP]

Tabela S54. Motivo VI-4

Nome Genérico	Início	Valor $p$	Sítio
GmERF#008	004	2.55e <sup>-12</sup>	MTA PPSKYTHHSKLLIPS EEDPPKHPYP
GmERF#023	003	1.80e <sup>-11</sup>	MS V#SKLKHHPMNLATP SSHFDNDPPC
GmERF#058	008	4.50e <sup>-14</sup>	MEDSILC K#TV#RSVTKKLLISP KKSSQTNST
GmERF#072	007	3.52e <sup>-12</sup>	MASSSI KNT#H#LNRTKLFMEE DTLK#K#RHY
GmERF#081	008	2.60e <sup>-15</sup>	MEDSILC K#TV#H#TVTKKLLISP KKSSQTNSTT
GmERF#102	007	3.37e <sup>-14</sup>	MASSSI K#T#H#LNRTKLFMKE ETLK#K#K#H#Y
GmERF#147	004	1.07e <sup>-15</sup>	MTA PPTKYTHHTKLLIPP EEDPPKHHYP
GmERF#163	013	2.58e <sup>-13</sup>	PQNRSNSILC KYTE#H#TVTKK#H#MKS KNNTSLSPKV
GmERF#196	010	3.36e <sup>-13</sup>	MEQQNSILC KYTE#H#Q#TV#M#K#H#MKS KDNTFCSPKV
GmERF#199	003	1.67e <sup>-11</sup>	MS V#AKLKHHPMNLASP SSHFEDDPPC
GmERF#275	004	1.07e <sup>-15</sup>	MAA PPTKYTHHTKLLIPP GPEEDPPKHP
<b>Consenso</b>			[KP] [HP] TK [HY] T [HT] [HV] TK [LK] L [IM] [KPS] [PS]

Tabela S55. Motivo VI-5

Nome Genérico	Início	Valor $p$	Sítio
GmERF#008	261	8.70e <sup>-12</sup>	FLTACEDFGF N#NRD#FFQ#D#D DLF
GmERF#054	225	4.12e <sup>-15</sup>	ITACPSPSSL CQVDDFFQDIL LASDPVTTTN
GmERF#058	305	4.12e <sup>-15</sup>	DKACPPPSTL CQVDDFFQDIL LASDPLVLL
GmERF#072	313	2.24e <sup>-12</sup>	ENLDLGFTSW H#RECD#N#F#Q#D#I#G DLFVWDPLVA
GmERF#081	284	4.12e <sup>-15</sup>	PEACPSPSSL CQVDDFFQDIL FASDPL
GmERF#102	271	2.24e <sup>-12</sup>	EDLDLGFTSW H#RECD#N#F#Q#D#I#G GDLFVWDPLV
GmERF#147	268	1.66e <sup>-11</sup>	EDFGFKSWHT D#R#N#R#D#F#F#Q#D#I#D DLFVSDPLLA
GmERF#163	316	1.87e <sup>-15</sup>	EKFLPSATL CQVDDYFQDIL LGSDPLVAL
GmERF#196	255	6.32e <sup>-12</sup>	LFDETT#P#F#S#I S#V#D#D#Y#F#Q#D#I#L LGSDPLV#L#F#C
<b>Consenso</b>			[CH] [QR] [VEN] [DCR] D [FNY] FQDI [LDG]

Tabela S56. Motivo VI-6

Nome Genérico	Início	Valor $p$	Sítio
GmERF#008	174	1.32e <sup>-13</sup>	T#P#P#A#T#C#H#N#T#D P#P#E#V#A#N#G#Y#G#Y#A#S#G#E#E T#N#D#N#N#N#L#F#S
GmERF#054	122	1.43e <sup>-12</sup>	P#T#E#V#C#E#M#K#V#V V#T#E#E#A#F#N#S#S#Y#D#S#S#E#D H#C#H#K#L#S#S#P#T
GmERF#058	175	1.17e <sup>-14</sup>	E#S#P#S#Q#A#T#T#V#A V#T#E#E#A#S#G#S#G#Y#D#S#G#D#D H#C#Q#H#N#L#S#S#P#T
GmERF#072	167	1.94e <sup>-15</sup>	G#A#D#A#L#T#N#F#V#T P#P#R#E#N#R#K#T#G#Y#C#S#G#E#E S#R#N#D#L#R#S#P#T
GmERF#081	188	2.66e <sup>-15</sup>	P#A#E#V#C#E#M#K#V#V V#T#E#E#A#S#G#S#G#Y#D#S#S#E#D H#C#H#N#L#S#S#P#T
GmERF#102	167	5.57e <sup>-16</sup>	G#A#D#A#L#T#N#F#I#T P#P#P#E#N#R#K#T#G#Y#C#S#G#E#E S#R#N#D#D#L#R#S#P#T
<b>Consenso</b>			[PV] [PT] [EP] E [AN] [RS] [GK] [ST] GY [DC] S [GS] E [DE]



Tabela S57. Motivo VII-1

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#015	001	5.13e <sup>-25</sup>	MCGGAIIDFI PRRGRRRLTA SELWPNSFAK
GmERF#044	001	3.26e <sup>-20</sup>	MCGGAIISDFI PAAAIAGSRR LTADYLWPD
GmERF#092	001	2.80e <sup>-21</sup>	MCGGAIISDFI PAGFAGGAQR VTADILWPNL
GmERF#114	001	1.14e <sup>-23</sup>	MCGGAIISDFI GVKRGRNLAA QELWSELDPF
GmERF#125	001	5.13e <sup>-25</sup>	MCGGAIIDFI PRRGRRRLTA SELWPNSFGK
GmERF#210	001	1.14e <sup>-23</sup>	MCGGAIISDFI GVKRGRNLAA QELWSELDPF
GmERF#214	001	2.26e <sup>-21</sup>	MCGGAIISDFI PAGFASGARR VTADILWPSL
GmERF#258	001	1.14e <sup>-23</sup>	MCGGAIIDFI PRQRERTLTA SELWPKRSEP
<b>Consenso</b>			MCGGAI [SA]DFI [PG] [ARV] [GKR] [RGP] [GA]R [GNR] [LA] [TAR] [AR]

Tabela S58. Motivo VII-2

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#015	219	4.42e <sup>-15</sup>	SASGSEGAYS TTEFMGSVQNGNGYLGTVM EKKEKETTEVI
GmERF#044	321	2.59e <sup>-24</sup>	TSLESLLSGD TTQGGNLMNLWCFDDIPSM AGGVF
GmERF#092	369	4.62e <sup>-19</sup>	ALASLLSEDA SQDAAGNEMNLSFDDLP SM AGVF
GmERF#114	212	7.17e <sup>-16</sup>	LENELPVSNN ICAEWDNMDLLWMLDDVVVP NRHLIY
GmERF#125	221	5.93e <sup>-15</sup>	LASGSEGAYS TTEFMGSVQNGNGYLGTVM EKNEKETTEVI
GmERF#210	198	3.84e <sup>-16</sup>	QLERFLGLDN MGAEWDNMDLLWMLDDVVVP NRHLIY
GmERF#214	359	1.82e <sup>-21</sup>	TLASLLGGDA THDAAGNEMNLSFDDLP SM AGVF
GmERF#262	235	2.65e <sup>-23</sup>	TSLESLLSGD TTQGGNLMNLWCFDDIPSM SSGVF
<b>Consenso</b>			T [TG] [ADEQ] [ADEE] [AGMW] [GD] [NS] [LMP] [MDQ] [ND] [LG] [WN] [SGM] [FLY] [DL] [DG] [VIL] [PTV] [SV] [MP]

Tabela S59. Motivo VII-3

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#015	245	1.49e <sup>-41</sup>	GTVMEKKEKE TEVIEAEEKKNVLELSEELMAYENYMKFYQIPYYDGGST TNNVQESLVG
GmERF#044	270	1.55e <sup>-34</sup>	FVQANQQQP NSQNVVSI EDDSAKTLSEELVDIESELKFFQMPYLEGSWG DTSLESLLSG
GmERF#092	316	5.13e <sup>-34</sup>	GAADQNQKKN NLLDMASVQDDSAKTLSEELADIE SQLKFFETPSFLDEAW ADAALASLLS
GmERF#125	247	1.63e <sup>-42</sup>	VTAMEKNEKE TEVIEAEEKKNVLELSEELMAYENYMKFYQIPYYDGGST APPNNVQESL
GmERF#214	306	4.14e <sup>-34</sup>	GAADQVQKTN NSQDVVAAQDDSAKTLSEELADIE SQLKFFETPSFLDEAW ADATLASLLG
GmERF#258	145	9.44e <sup>-30</sup>	SSNADENCGS HVAYSVTEANNKVEKLEELMAYENLMRFYQIPENVGDL WAFPDTSPL
<b>Consenso</b>			[NT] [ES] [QV] [DI] [EV] [AV] SE [EQ] [DK] [DN] [KS] [AV] K [TE] LSEEL [MA] [AD] [IY] E [NS] [QY] [LM] KF [FY] [QE] [IT] P [YS] [FY] [DL] [GD] [EQ] [AS] [TW]

Tabela S60. Motivo VII-4

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#015	050	2.05e <sup>-16</sup>	SHIATQQPST LKRSQPPKVSQVENKPVKRQ RKNLYRGIRQ
GmERF#044	075	5.37e <sup>-17</sup>	KPFafaASSR LSTAAKSVAFQGRAEISANRK RKNQYRGIRQ
GmERF#092	094	8.64e <sup>-21</sup>	NKAASKPLSR GATTVKSVESKQAEKCAKRK RKNQYRGIRQ
GmERF#125	050	4.47e <sup>-17</sup>	SHIADQQRST LKMSPPPKASEQVENKPVKRQ RKNLYRGIRQ
GmERF#214	090	1.62e <sup>-18</sup>	RSSNKAASKP LSRGSATVESKQAEKCAKRK RKNQYRGIRQ
GmERF#262	017	6.94e <sup>-19</sup>	RFLHFSNYIK GSTTAKSVAFQVRAEKFANRK RKNQYRGIRQ
<b>Consenso</b>			[LG] [SK] [TR] [ST] A [KP] [SP] [VK] [AE] [SF] [EKQ] [GQ] [QRV] [AE] [EN] K [CP] [AV] [KN] R [KQ]

Tabela S61. Motivo VII-5

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#015	034	6.01e <sup>-13</sup>	LWPNSFAKDD DFDDYSHIATQQ PSTLKRSP
GmERF#044	255	5.05e <sup>-18</sup>	EISSMLSAPL DCESEFVQANQQ QPNSQNVVSI
GmERF#092	298	3.52e <sup>-11</sup>	ISSMLSAPL DGESEFVQGAADQ NQKKNLLDM
GmERF#125	034	7.89e <sup>-13</sup>	LWPNSFGKDG DFDDYSHIADQQ RSTLKMSPPP
GmERF#262	196	5.05e <sup>-18</sup>	EISSMLSAPL DCESEFVQANQQ PNSQNV MAYL
<b>Consenso</b>			D [CFG] [ED] [SFL] [DHQ] [FY] [VS] [QH] [ING] A [NADT] [QD] Q

Tabela S62. Motivo VII-6

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#044	183	1.56e <sup>-61</sup>	ENLKTVPNL <b>G</b> H <b>K</b> F <b>S</b> A <b>G</b> N <b>N</b> <b>F</b> M <b>D</b> L <b>V</b> E <b>Q</b> K <b>P</b> L <b>V</b> S <b>Q</b> <b>Y</b> A <b>N</b> M <b>A</b> S <b>F</b> P <b>G</b> S <b>G</b> N <b>L</b> R <b>S</b> L <b>P</b> S <b>S</b> D <b>D</b> A <b>T</b> L <b>Y</b> F <b>S</b> S <b>D</b> Q <b>S</b> N <b>S</b> F <b>D</b> <b>Y</b> A <b>P</b> E <b>I</b> S <b>S</b> M <b>L</b> S <b>A</b>
GmERF#092	214	7.51e <sup>-62</sup>	K <b>P</b> K <b>I</b> N <b>Q</b> M <b>F</b> N <b>F</b> <b>G</b> D <b>N</b> L <b>E</b> G <b>Y</b> S <b>F</b> I <b>D</b> Q <b>V</b> E <b>Q</b> K <b>P</b> L <b>V</b> N <b>Q</b> <b>Y</b> V <b>N</b> R <b>A</b> F <b>F</b> A <b>G</b> N <b>G</b> V <b>Q</b> V <b>S</b> P <b>V</b> T <b>P</b> S <b>A</b> D <b>V</b> T <b>A</b> Y <b>F</b> S <b>S</b> E <b>H</b> S <b>S</b> S <b>S</b> F <b>D</b> <b>Y</b> S <b>D</b> L <b>G</b> W <b>G</b> E <b>Q</b> V <b>P</b>
GmERF#214	211	1.46e <sup>-68</sup>	K <b>P</b> K <b>M</b> N <b>Q</b> M <b>F</b> N <b>F</b> <b>G</b> H <b>N</b> L <b>E</b> G <b>Y</b> S <b>F</b> I <b>D</b> Q <b>V</b> E <b>Q</b> K <b>P</b> L <b>V</b> N <b>Q</b> <b>Y</b> V <b>N</b> R <b>A</b> F <b>F</b> P <b>G</b> N <b>G</b> V <b>Q</b> V <b>S</b> P <b>V</b> T <b>P</b> S <b>A</b> D <b>V</b> T <b>A</b> Y <b>F</b> S <b>S</b> E <b>H</b> S <b>S</b> N <b>S</b> F <b>D</b> <b>Y</b> S <b>D</b> L <b>G</b> W <b>G</b> E <b>Q</b> V <b>P</b>
GmERF#262	118	3.60e <sup>-62</sup>	R <b>S</b> K <b>V</b> N <b>P</b> Q <b>E</b> N <b>L</b> <b>S</b> H <b>K</b> F <b>G</b> A <b>C</b> N <b>N</b> <b>F</b> M <b>D</b> L <b>V</b> E <b>Q</b> K <b>P</b> L <b>V</b> N <b>Q</b> <b>Y</b> A <b>N</b> M <b>A</b> S <b>F</b> P <b>G</b> S <b>G</b> N <b>L</b> T <b>S</b> L <b>P</b> S <b>S</b> D <b>D</b> V <b>T</b> L <b>Y</b> F <b>S</b> S <b>D</b> Q <b>S</b> N <b>S</b> F <b>G</b> W S <b>E</b> Q <b>G</b> P <b>K</b> T <b>P</b> E <b>I</b>
<b>Consenso</b>			[GS] [HD] [KN] [FL] [EGS] [AG] [GY] [NY] [NS] [HP] [IM] D [LQ] VEQKPLV [NS] QY [AV] N [MPR] A [PS] F [PA] G [NS] G [NV] [GQ] [LV] [SRT] [PS] [LV] [PT] [PS] S [AD] D [VA] T [AL] YFSS [DE] [HQ] [GS] S [NS] SF [DG] [YW]

Tabela S63. Motivo VII-7

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#044	243	2.66e <sup>-12</sup>	S <b>D</b> Q <b>G</b> S <b>N</b> S <b>F</b> D <b>Y</b> <b>A</b> P <b>E</b> I <b>S</b> S <b>M</b> L <b>S</b> A <b>P</b> L <b>D</b> C <b>E</b> S <b>H</b> F <b>V</b> Q
GmERF#092	285	4.05e <sup>-13</sup>	D <b>L</b> G <b>W</b> G <b>E</b> Q <b>V</b> P <b>K</b> <b>T</b> P <b>E</b> I <b>S</b> S <b>M</b> L <b>S</b> A <b>A</b> P <b>L</b> D <b>G</b> E <b>S</b> Q <b>F</b> V
GmERF#214	282	1.44e <sup>-11</sup>	D <b>L</b> G <b>W</b> G <b>E</b> Q <b>V</b> P <b>K</b> <b>T</b> P <b>E</b> I <b>S</b> S <b>L</b> L <b>S</b> A <b>A</b> P <b>L</b> E <b>G</b> A <b>A</b> D <b>Q</b> V
GmERF#262	184	4.05e <sup>-13</sup>	S <b>F</b> G <b>W</b> S <b>E</b> Q <b>G</b> P <b>K</b> <b>T</b> P <b>E</b> I <b>S</b> S <b>M</b> L <b>S</b> A <b>P</b> L <b>D</b> C <b>E</b> S <b>H</b> F <b>V</b> Q
<b>Consenso</b>			[TA] PEISS [ML] LSA

Tabela S64. Motivo VIII-1 (Motivo EAR)

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#018	193	2.37e <sup>-11</sup>	LFDDGETELP <b>ELPPDITSSMGYEMG</b> NGFGANATGS
GmERF#026	188	4.57e <sup>-13</sup>	DCVLISSFRQ <b>ELFFDLNLPSSDADD</b> DNDDDLPATA
GmERF#027	196	8.82e <sup>-14</sup>	DNIVSSSFRP <b>ELFFDLNALPFDDAA</b> ADDDLRRTAL
GmERF#083	169	2.00e <sup>-11</sup>	VDYERVPHRK <b>VLDDLNVPPPEVA</b>
GmERF#113	183	7.78e <sup>-14</sup>	DNVASSFPRE <b>ELFFDLNALPLDDAD</b> VATDDLFCV
GmERF#160	112	1.56e <sup>-10</sup>	TLDSSSPSP <b>PELDTLTLPLSVAV</b> TVFPVARPVL
GmERF#179	208	3.21e <sup>-12</sup>	EGEVKNGSR <b>IFDLNHPPEIA</b>
GmERF#180	072	6.32e <sup>-13</sup>	KAKTNFPPAP <b>ELCLDLNVSSDHRW</b> PHVPPRSVL
GmERF#181	114	2.89e <sup>-10</sup>	STLDSSPPP <b>PELDTLTLPLSSAV</b> FPVARPVLFF
GmERF#187	194	3.27e <sup>-13</sup>	DNIVSSSFRP <b>ELFFDLNALPFDDAA</b> AVVDDLRC
GmERF#188	183	8.82e <sup>-14</sup>	DCVLTSSFRQ <b>ELFFDLNLPDDAA</b> DDDDDDDVPA
GmERF#204	075	6.32e <sup>-13</sup>	KAKTNFPPAP <b>ELCLDLNVSSDHRW</b> PHVPPRSVL
GmERF#205	205	2.65e <sup>-12</sup>	EGVVKNGSR <b>MFDLNHPPEIA</b>
GmERF#217	151	6.09e <sup>-10</sup>	NNKDSTIELP <b>ELPPDITSSVGYEGF</b> YNNDGGYWE
GmERF#220	193	3.34e <sup>-11</sup>	LFDDGETQLP <b>ELPPDITSSMGYEMS</b> NGLYGNDVGF
GmERF#242	185	1.27e <sup>-14</sup>	DNVASSFRE <b>ELFFDLNVLPLDDAD</b> VATDDLFCV
<b>Consenso</b>			PLP[LF]DL[NT][LV][PL]P[LP]D[ED]AA

Tabela S65. Motivo VIII-2

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#026	161	1.21e <sup>-22</sup>	PIIPLNDD <b>DC#SDCSSSVVDDGQDCVL</b> ISSFRQPLPF
GmERF#027	169	9.73e <sup>-25</sup>	YPRTPPLVPE <b>DC#SDCSSSVVDDGDDNIV</b> SSSFRPPLPF
GmERF#083	148	4.97e <sup>-12</sup>	EVRASFDPV <b>AGFSDSSSVVDYERVPHRK</b> VLDDLNVPP
GmERF#085	176	3.76e <sup>-14</sup>	QSFPLRFEPG <b>FVQSDSSSVVDCQPKREL</b> NLDLNLAPPN
GmERF#090	100	4.36e <sup>-13</sup>	KPKTARSNAR <b>KCSSVEQLFSGVQLRKDDGI</b> GENGNVE
GmERF#109	110	4.36e <sup>-13</sup>	KPKTARNAR <b>KCSSVEQLFSGVQLRKDDGI</b>
GmERF#112	183	1.29e <sup>-12</sup>	HFHNNNNKSS <b>SCYVNSSSSAAANNCLQGCC</b> FTNPNLGMEV
GmERF#113	155	2.12e <sup>-20</sup>	YPRTPVPIPE <b>DCRSDCSSSVVDDGEGDNV</b> ASSFPREPLP
GmERF#179	181	7.89e <sup>-15</sup>	HHPVAASEFH <b>ATTSDDSSSVIDLN#NEGEV</b> KGNGSRIFDL
GmERF#187	167	5.80e <sup>-23</sup>	YPRTPPLVPE <b>DC#SDCSSSVVDDSDNIV</b> SSSFRPPLPF
GmERF#188	157	9.05e <sup>-21</sup>	PIIPLDDDD <b>DC#SDCSSSVVDDQDCVLT</b> SSFRQPLPFD
GmERF#205	178	6.61e <sup>-15</sup>	PHPVAASEYH <b>ATTSDDSSSVIDLN#NEGVV</b> KGNGSRMFDL
GmERF#242	157	7.83e <sup>-21</sup>	YPRTPVPIPE <b>DCRSDCSSSVVDDGEGDNV</b> ASSFRREPLP
GmERF#243	140	8.14e <sup>-13</sup>	HNNNNNSKSS <b>SCYVNSSSSSSATNCLQGCC</b> FTNANFGMEV
<b>Consenso</b>			[DA]CHSD[CS]DSSS[SV]V[VD]D[DN][GR]DD[DG]xV

Tabela S66. Motivo VIII-3

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#026	115	6.00e <sup>-22</sup>	LSTKFEVVQ <b>VNRPTSSGMSSTVESFSGPR</b> VPSSSSRAKP
GmERF#027	117	1.05e <sup>-21</sup>	HHNNNNLNN <b>PQRPTSSGMSSTVESFSGPR</b> PPTTTTTTTI
GmERF#083	086	2.58e <sup>-13</sup>	TKAKTNFPTH <b>AAAARSPQSSTLDSSSPPP</b> LDLTAAPLT
GmERF#085	100	6.90e <sup>-16</sup>	NVNVNKNKSV <b>NVKSHPQSSTVESATPER</b> EATRRSTSAA
GmERF#113	111	7.17e <sup>-22</sup>	FANTGEDFHD <b>HRRPTSSGMSSTVESFSGPR</b> AAVPATAPVA
GmERF#160	091	4.64e <sup>-15</sup>	NFPPTSELIL <b>NNNIRSPQSSTLDSSSPPP</b> PPPPLDLTLT
GmERF#179	087	5.34e <sup>-17</sup>	AKTNFPLE <b>NVKNSSPQSSTVESSSRDR</b> DVAADSSPLD
GmERF#181	094	1.48e <sup>-14</sup>	TPSELILNN <b>NINARSPQSTLDSSSPPP</b> PPPLDLTLTP
GmERF#187	119	1.05e <sup>-21</sup>	SNNNDNLNN <b>PQRPTSSGMSSTVESFSGPR</b> PPTTTTTTTT
GmERF#188	111	9.82e <sup>-21</sup>	GLSTKFEPAQ <b>VNRPTSSGMSSTVESFSGPR</b> VPPSSSSSRK
GmERF#205	084	1.75e <sup>-16</sup>	AKTNFPLE <b>NFKNASPQSSTVESSSRDR</b> DVAADSSPLD
GmERF#242	113	7.17e <sup>-22</sup>	GAGAGEDFQD <b>HRRPTSSGMSSTVESFSGPR</b> AAVPAKAPVI
<b>Consenso</b>			NN[RK]P[TR]S[PS][GS][MQ]SST[VL][ED]S[FS]S [GP]PR

Tabela S67. Motivo VIII-4

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#001	249	6.28e <sup>-21</sup>	KMNTHC GTNF <b>SEKGGDDGSGGFFSRESSDSGLLEEIVNK</b> FLPKTKPNKC
GmERF#016	223	1.08e <sup>-22</sup>	KMNTFPTCGT <b>NFAEKGGDDGSGGFFSRESSDSGLLEEIVNK</b> FLPRTKPSKC
GmERF#101	169	1.84e <sup>-30</sup>	PSPLNNDMCV <b>ETTYGSAQDDNFFFSNSGYLECIVPD</b> DCFRPASSSS
GmERF#112	212	4.71e <sup>-25</sup>	CCFTNP NLGM <b>EVAEAIDEESDFFPIESSDSGLLEEIVHK</b> FLPKSKPKKE
GmERF#160	153	3.28e <sup>-19</sup>	RADAMIAVSR <b>REMGCFERPAADFRRNDSDSYNNRRVLLD</b> LDLNVPPLPE
GmERF#175	161	1.50e <sup>-31</sup>	ISSHIENPSD <b>MSAYGSPQDDNFFFSNSGYLECIVQD</b> NCFRKSNASA
GmERF#181	153	6.00e <sup>-20</sup>	RADTMVAVSR <b>RETCGFERPAADFRRGDSDSYNNRRILLD</b> LDLNVPPLPE
GmERF#209	160	3.19e <sup>-28</sup>	ISSHIENPSG <b>MSSSGSPQDDNFFFSNSGYLECIVQD</b> NCSRKSNASA
GmERF#224	177	1.84e <sup>-30</sup>	PSPLNNDMCV <b>ETTYGSAQDDNFFFSNSGYLECIVPD</b> NCFRPASSSS
GmERF#243	169	3.24e <sup>-24</sup>	CCFTNANFGM <b>EVAEAIDEESEFFPRESSDSGLLEEIVH</b>
<b>Consenso</b>			[EMR] [ESTV] [AT] [EYC] [GA] [SFI] [DAEP] [QDER] [DEGP] [DSA] [NAG] [FD] F [FPRS] [RS] [ESN] [DS] S [DN] S [GD] [YL] [LN] [ER] [CER] I [VL] [HLNPQ] [DK]

Tabela S68. Motivo VIII-5

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#001	120	8.24e <sup>-19</sup>	TSPQPSSATT <b>EFLFNFNNSNNFHKHSVFNH</b> HHGNRHITGS
GmERF#016	121	3.01e <sup>-19</sup>	TSPQPSSATT <b>EFLFNFNNNNNFHKHSLFNH</b> HPPRNPSSLN
GmERF#027	096	5.06e <sup>-17</sup>	PFCYPHPTTD <b>PFYTGFDQHHHHNNNNLNN</b> PQRPTSSGMS
GmERF#112	160	2.19e <sup>-18</sup>	NNNNVSSSLN <b>MVLEFRDFLNSSSHHF#NNNNK</b> SSSCYVNSSS
GmERF#175	124	5.37e <sup>-15</sup>	FHNALTPMNT <b>QPLLPESHNTHTQPTNQSGL</b> SHHLGAISSH
GmERF#187	096	2.32e <sup>-14</sup>	PFCYQHPTTD <b>PFYAGFDHQNNNNNDNL</b> NNPQRPTSSG
GmERF#209	124	2.15e <sup>-14</sup>	FHNALTPMNA <b>QPLLPESYNTHNPQPINQSGL</b> SHLGNISSHI
GmERF#217	123	5.59e <sup>-16</sup>	ASVAASQFSQ <b>QQQQPDNNNNNIINNSNNK</b> DSTIELPPLP
GmERF#243	115	5.68e <sup>-18</sup>	NPSSSSSSSLN <b>MLLEFRDFLNSSSHHF#NNNNN</b> SKSSSCYVNS
<b>Consenso</b>			[QEMP] [FHP] [LF] [FLY] [PR] [DGNP] [FS] [HNL] [ND] [SNT] [HNS] [NS] [HF] [HQ] [NFKP] [HN] [NS] [NQ] [NFS] [NG] [LHKN]

Tabela S69. Motivo VIII-6

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#026	205	1.66e <sup>-13</sup>	LPPSDADDDN <b>DDDLFATALCL</b>
GmERF#027	212	6.29e <sup>-13</sup>	NALPFDDAAA <b>DDDLRRATALCL</b>
GmERF#113	200	5.96e <sup>-14</sup>	ALPLDDADVA <b>TDDLFC TVLCL</b>
GmERF#187	212	1.66e <sup>-14</sup>	LPFDDAAVV <b>DDDLRATALCL</b>
GmERF#188	202	1.33e <sup>-12</sup>	PDAAAYDDDD <b>DDDVFATALCL</b>
GmERF#242	202	5.96e <sup>-14</sup>	VLPLDDADVA <b>TDDLFC TVLCL</b>
<b>Consenso</b>			[DT] DDL [FPR] [CA] T [AV] LCL

Tabela S70. Motivo VIII-7

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#101	019	1.26e <sup>-37</sup>	DTPFKGYDPS <b>QTQMCLSLLRNTSPCGERRGRKQAEFG</b> RFLGVRRRPW
GmERF#175	018	1.43e <sup>-36</sup>	DASLKGYESN <b>QTQMNLSSLQRNMSPSGERRGRKQAEFG</b> RFLGVRRRPW
GmERF#209	018	1.43e <sup>-36</sup>	DASLKGYESN <b>QTQMNLSSLQRNMSPSGERRGRKQAEFG</b> RFLGVRRRPW
GmERF#224	019	1.26e <sup>-37</sup>	DTAFKGYDPS <b>QTQMCLSLLRNTSPCGERRGRKQAEFG</b> RFLGVRRRPW
<b>Consenso</b>			QTQM [CN] LSLLRN [MT] SP [CS] GERRGR [GR] KQAEFG

Tabela S71. Motivo VIII-8

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#001	370	4.30e <sup>-21</sup>	SVQAMALGNE <b>QIMNFAENCVVDFVQYHEL</b> LNAFAIRMQN
GmERF#016	295	7.49e <sup>-24</sup>	GFNSMAMEND <b>QNIINNAENCVVDFVQYQEL</b> LNAFAIRMQN
GmERF#112	118	9.35e <sup>-25</sup>	AADPHHHHLF <b>QPYNINPKHCHVTRFVNQNL</b> NADFSANSHN
GmERF#243	072	2.41e <sup>-26</sup>	DAADPHHHHLF <b>QPYNINPKHCHVTRWI PNQNL</b> SADFSSTNGHN
<b>Consenso</b>			Q [PIN] [YIM] [NIM] [IN] [NH] [AP] [EK] [HN] C [HV] V [ET] [DR] [VFW] [FIV] [PQ] [NY] [QH] [EN] L



Tabela S72. Motivo IX-1

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#013	172	1.64e <sup>-17</sup>	KKVEETHGDV <b>R</b> LENVVVLED <b>L</b> GA <b>E</b> YLE <b>H</b> LLM MSSHDTTSCS
GmERF#014	128	6.31e <sup>-17</sup>	SSASALASPS <b>S</b> T <b>N</b> Q <b>V</b> IE <b>F</b> E <b>F</b> Y <b>F</b> DD <b>K</b> LL <b>E</b> ELL <b>G</b> YEDKNKPKD
GmERF#038	203	6.35e <sup>-16</sup>	ACGSNGKWD <b>T</b> <b>S</b> T <b>Q</b> N <b>L</b> L <b>V</b> LED <b>L</b> G <b>P</b> E <b>Y</b> L <b>D</b> Q <b>L</b> L <b>S</b> FTCPGSWC
GmERF#039	125	6.31e <sup>-17</sup>	GEGSSQQQ <b>Q</b> G <b>N</b> DD <b>T</b> F <b>E</b> L <b>E</b> CL <b>D</b> N <b>K</b> LL <b>E</b> ELL <b>Q</b> QAPDRSGN
GmERF#040	219	1.78e <sup>-19</sup>	DQREQLDFSR <b>T</b> S <b>Q</b> N <b>V</b> LV <b>F</b> ED <b>L</b> GA <b>E</b> YLE <b>Q</b> LL <b>T</b> LTS
GmERF#115	103	1.21e <sup>-18</sup>	STANGSS <b>F</b> AP <b>T</b> H <b>K</b> E <b>V</b> IE <b>F</b> E <b>F</b> CL <b>D</b> N <b>K</b> VLE <b>D</b> LL <b>E</b> SEVKRRNED
GmERF#116	092	7.04e <sup>-18</sup>	NSSTKGN <b>G</b> SS <b>F</b> D <b>K</b> E <b>V</b> IE <b>L</b> E <b>Y</b> L <b>D</b> D <b>K</b> VLE <b>L</b> LL <b>E</b> LEEKRRNED
GmERF#117	108	5.52e <sup>-19</sup>	TGVVGS <b>S</b> SA <b>Q</b> Q <b>G</b> E <b>V</b> IE <b>F</b> E <b>F</b> Y <b>L</b> DD <b>K</b> VLE <b>D</b> LL <b>E</b> LEEKRGKKE
GmERF#123	137	6.92e <sup>-17</sup>	KRRHSLAA <b>K</b> R <b>K</b> K <b>E</b> D <b>V</b> LV <b>F</b> H <b>D</b> L <b>G</b> A <b>D</b> Y <b>L</b> E <b>H</b> LLM CSDLPSTTIV
GmERF#124	095	2.26e <sup>-20</sup>	SSSITHAG <b>G</b> G <b>Q</b> Q <b>R</b> Q <b>V</b> F <b>E</b> F <b>E</b> Y <b>L</b> DD <b>K</b> VLE <b>L</b> LL <b>E</b> FEENKVQKG
GmERF#126	104	5.19e <sup>-20</sup>	SSSNSML <b>K</b> ND <b>H</b> G <b>K</b> Q <b>V</b> IE <b>F</b> E <b>F</b> CL <b>D</b> D <b>K</b> LL <b>E</b> LL <b>D</b> CDDYAYEKDL
GmERF#127	163	2.53e <sup>-18</sup>	SKGK <b>K</b> CK <b>G</b> KE <b>T</b> S <b>E</b> S <b>V</b> VV <b>L</b> ED <b>L</b> G <b>V</b> D <b>Y</b> L <b>E</b> Q <b>L</b> L <b>S</b> LSDQS
GmERF#128	131	4.59e <sup>-16</sup>	FSGEIMQ <b>G</b> D <b>H</b> DD <b>T</b> F <b>E</b> L <b>E</b> Y <b>F</b> D <b>N</b> KLL <b>E</b> ELL <b>Q</b> MQDNRFH
GmERF#129	238	1.06e <sup>-18</sup>	HGMESQ <b>N</b> VNI <b>I</b> S <b>Q</b> N <b>V</b> LV <b>L</b> ED <b>L</b> G <b>S</b> E <b>Y</b> L <b>E</b> Q <b>L</b> L <b>S</b> LTSSSDDDYF
GmERF#134	189	2.24e <sup>-18</sup>	NNN <b>S</b> KL <b>Q</b> SS <b>T</b> V <b>D</b> N <b>A</b> V <b>V</b> F <b>E</b> D <b>L</b> G <b>P</b> D <b>Y</b> L <b>E</b> Q <b>L</b> L <b>M</b> SSDHHIPTTF
GmERF#135	109	4.78e <sup>-17</sup>	HASEST <b>S</b> AA <b>Q</b> Q <b>R</b> V <b>F</b> E <b>I</b> P <b>C</b> L <b>D</b> D <b>K</b> VLE <b>D</b> LL <b>G</b> LSEEHKKKK
GmERF#136	118	3.18e <sup>-19</sup>	ADVEARG <b>R</b> SE <b>Q</b> G <b>R</b> E <b>V</b> F <b>E</b> F <b>E</b> Y <b>L</b> DD <b>K</b> LL <b>E</b> LL <b>D</b> CENKKKRG
GmERF#165	105	5.19e <sup>-20</sup>	SSSNSML <b>K</b> SD <b>H</b> G <b>K</b> Q <b>V</b> IE <b>F</b> E <b>F</b> CL <b>D</b> D <b>K</b> LL <b>E</b> LL <b>D</b> CDDYAYEKDL
GmERF#166	185	4.03e <sup>-18</sup>	CRENKQ <b>E</b> SEA <b>P</b> T <b>P</b> S <b>V</b> VV <b>L</b> ED <b>L</b> G <b>V</b> E <b>Y</b> L <b>E</b> Q <b>L</b> L <b>S</b> ISDQSVSL
GmERF#167	188	1.21e <sup>-18</sup>	KAKET <b>S</b> SED <b>P</b> T <b>P</b> S <b>V</b> LV <b>F</b> ED <b>L</b> G <b>A</b> D <b>Y</b> L <b>E</b> Q <b>L</b> L <b>S</b> ISDQS
GmERF#168	140	1.32e <sup>-19</sup>	KCKG <b>K</b> GT <b>S</b> ED <b>P</b> T <b>Q</b> S <b>V</b> VV <b>L</b> ED <b>L</b> G <b>V</b> D <b>Y</b> L <b>E</b> Q <b>L</b> L <b>S</b>
GmERF#169	138	1.54e <sup>-19</sup>	RGKET <b>S</b> SED <b>S</b> T <b>Q</b> S <b>V</b> VV <b>F</b> ED <b>L</b> G <b>V</b> D <b>Y</b> L <b>E</b> H <b>L</b> L <b>S</b> VSDQSESTSP
GmERF#170	124	5.40e <sup>-16</sup>	VRPGEIMQ <b>G</b> G <b>D</b> L <b>D</b> D <b>T</b> F <b>E</b> L <b>E</b> Y <b>L</b> D <b>N</b> KLL <b>E</b> ELL <b>Q</b> MQDNRFYK
GmERF#171	229	6.31e <sup>-18</sup>	GHGMELE <b>N</b> IV <b>L</b> S <b>Q</b> N <b>V</b> LV <b>L</b> ED <b>L</b> G <b>S</b> E <b>Y</b> L <b>E</b> Q <b>L</b> L <b>S</b> LTSSDDDFLC
GmERF#177	233	1.10e <sup>-13</sup>	ATKKLE <b>C</b> MVE <b>N</b> E <b>L</b> G <b>A</b> F <b>V</b> F <b>Q</b> D <b>L</b> G <b>S</b> D <b>Y</b> L <b>D</b> S <b>L</b> L <b>S</b> SF
GmERF#207	231	6.55e <sup>-14</sup>	ATKKLE <b>C</b> MVE <b>N</b> E <b>L</b> G <b>A</b> F <b>V</b> F <b>H</b> D <b>L</b> G <b>S</b> D <b>Y</b> L <b>D</b> S <b>L</b> L <b>S</b> SF
GmERF#211	103	4.41e <sup>-20</sup>	ISSNG <b>S</b> YAA <b>Q</b> Q <b>K</b> E <b>V</b> IE <b>F</b> E <b>F</b> CL <b>D</b> N <b>K</b> VLE <b>D</b> LL <b>E</b> SEVKRRNED
GmERF#212	092	2.69e <sup>-17</sup>	NSSTK <b>G</b> NESS <b>F</b> D <b>T</b> E <b>V</b> IE <b>L</b> E <b>Y</b> L <b>D</b> D <b>K</b> VLE <b>L</b> LL <b>E</b> LEEKRRNED
GmERF#254	092	1.32e <sup>-19</sup>	SKK <b>S</b> SS <b>N</b> TE <b>I</b> G <b>K</b> Q <b>V</b> VV <b>F</b> E <b>F</b> CL <b>D</b> D <b>K</b> LL <b>E</b> LL <b>D</b>
GmERF#255	180	1.29e <sup>-15</sup>	AKTSKR <b>K</b> QDL <b>E</b> SS <b>V</b> VV <b>L</b> ED <b>L</b> G <b>A</b> D <b>Y</b> L <b>E</b> ALL <b>S</b> TSADQSYKIA
GmERF#256	151	2.07e <sup>-19</sup>	DQREQLDF <b>S</b> S <b>S</b> S <b>Q</b> N <b>V</b> LV <b>F</b> ED <b>L</b> GA <b>E</b> YLE <b>Q</b> L <b>L</b> S LTS
GmERF#259	132	2.00e <sup>-17</sup>	MKRK <b>S</b> ET <b>K</b> KH <b>Q</b> L <b>H</b> D <b>V</b> M <b>V</b> LED <b>L</b> GA <b>E</b> YLE <b>H</b> LL <b>T</b> LSQDMDMDAC
GmERF#272	118	1.20e <sup>-17</sup>	ANVEARG <b>R</b> SE <b>Q</b> G <b>R</b> E <b>V</b> F <b>E</b> F <b>E</b> Y <b>L</b> DD <b>R</b> LL <b>E</b> LL <b>D</b> CEDKKKRG
GmERF#273	102	1.18e <sup>-16</sup>	LATP <b>S</b> SS <b>S</b> GA <b>Q</b> Q <b>R</b> V <b>F</b> E <b>I</b> P <b>C</b> L <b>D</b> D <b>K</b> VLE <b>D</b> LL <b>G</b> QSEEHKKKNK
GmERF#274	184	2.24e <sup>-18</sup>	KNN <b>S</b> KL <b>Q</b> SS <b>T</b> V <b>D</b> N <b>A</b> V <b>V</b> F <b>E</b> D <b>L</b> G <b>P</b> D <b>Y</b> L <b>E</b> Q <b>L</b> L <b>M</b> SSDHHIPTTF
Consenso			QxQ [NE] V [FVIL] [VE] [FL] E [DYC] L [GD] [DA] [KDE] [YLV] LE [QED] LLS

Tabela S73. Motivo IX-2

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#013	134	5.06e <sup>-21</sup>	VKESLR <b>D</b> MNC <b>T</b> L <b>S</b> Q <b>L</b> D <b>G</b> S <b>S</b> E <b>V</b> V <b>A</b> L <b>K</b> R <b>R</b> H <b>S</b> L <b>G</b> R <b>K</b> M <b>A</b> A <b>K</b> K <b>K</b> KVEETHGDVR
GmERF#030	109	1.01e <sup>-14</sup>	PPER <b>P</b> CF <b>L</b> PS <b>P</b> P <b>P</b> V <b>T</b> E <b>D</b> S <b>S</b> Q <b>P</b> Q <b>L</b> K <b>R</b> R <b>N</b> L <b>A</b> D <b>L</b> L <b>N</b> R <b>L</b> A <b>K</b> NRSQM <b>Q</b> K <b>H</b> M <b>T</b>
GmERF#038	157	3.28e <sup>-15</sup>	VVMESL <b>Q</b> DM <b>G</b> <b>F</b> K <b>A</b> L <b>K</b> N <b>D</b> L <b>S</b> E <b>V</b> L <b>A</b> L <b>K</b> R <b>M</b> H <b>V</b> T <b>R</b> T <b>K</b> S <b>R</b> A <b>S</b> R <b>S</b> GNKKV <b>K</b> R <b>A</b> C <b>G</b>
GmERF#040	177	2.36e <sup>-26</sup>	VVRESL <b>K</b> DMP <b>S</b> K <b>P</b> W <b>E</b> E <b>G</b> S <b>S</b> E <b>V</b> L <b>A</b> L <b>K</b> R <b>K</b> H <b>T</b> M <b>R</b> T <b>K</b> S <b>N</b> S <b>K</b> K <b>K</b> TKVDQ <b>R</b> E <b>Q</b> L <b>D</b>
GmERF#122	166	6.39e <sup>-15</sup>	LAEN <b>S</b> PP <b>S</b> SP <b>S</b> W <b>S</b> C <b>S</b> D <b>D</b> S <b>S</b> E <b>S</b> Q <b>A</b> K <b>R</b> R <b>K</b> S <b>L</b> A <b>D</b> L <b>L</b> N <b>K</b> L <b>A</b> K NRS <b>H</b> V <b>K</b> VER
GmERF#127	128	1.99e <sup>-22</sup>	VKESL <b>Q</b> E <b>I</b> Q <b>Y</b> <b>S</b> C <b>F</b> F <b>N</b> G <b>S</b> S <b>S</b> E <b>P</b> A <b>L</b> A <b>L</b> K <b>E</b> R <b>H</b> C <b>I</b> Q <b>R</b> K <b>L</b> S <b>S</b> K <b>G</b> K KCKG <b>K</b> E <b>T</b> S <b>E</b> S
GmERF#129	194	2.81e <sup>-24</sup>	SLK <b>D</b> M <b>Q</b> S <b>N</b> N <b>K</b> <b>P</b> I <b>S</b> L <b>E</b> D <b>G</b> S <b>S</b> E <b>V</b> L <b>A</b> L <b>K</b> R <b>K</b> H <b>T</b> M <b>R</b> R <b>K</b> S <b>K</b> A <b>T</b> N <b>K</b> KNK <b>R</b> D <b>H</b> G <b>M</b> E <b>S</b>
GmERF#134	148	1.93e <sup>-23</sup>	SLK <b>E</b> M <b>N</b> Y <b>A</b> H <b>D</b> <b>D</b> S <b>N</b> N <b>E</b> E <b>G</b> C <b>S</b> E <b>V</b> V <b>A</b> L <b>K</b> R <b>K</b> H <b>S</b> L <b>R</b> R <b>K</b> I <b>N</b> V <b>R</b> K <b>K</b> KNNNN <b>S</b> KL <b>Q</b>
GmERF#166	145	2.83e <sup>-25</sup>	KRV <b>K</b> ES <b>L</b> LE <b>I</b> <b>Q</b> Y <b>G</b> C <b>S</b> K <b>G</b> C <b>S</b> E <b>P</b> A <b>L</b> A <b>I</b> K <b>E</b> R <b>H</b> Y <b>I</b> R <b>R</b> K <b>L</b> S <b>S</b> K <b>A</b> R KCRE <b>N</b> K <b>Q</b> E <b>S</b> E
GmERF#167	147	2.44e <sup>-27</sup>	K <b>T</b> V <b>K</b> ES <b>L</b> Q <b>E</b> I <b>Q</b> Y <b>S</b> C <b>S</b> N <b>G</b> S <b>S</b> E <b>P</b> A <b>L</b> A <b>L</b> K <b>E</b> R <b>H</b> C <b>I</b> Q <b>R</b> K <b>L</b> S <b>L</b> K <b>G</b> R KCK <b>A</b> K <b>E</b> T <b>S</b> S
GmERF#168	101	7.82e <sup>-31</sup>	KRV <b>K</b> DSL <b>Q</b> E <b>I</b> <b>Q</b> Y <b>S</b> C <b>C</b> N <b>G</b> S <b>S</b> E <b>P</b> A <b>L</b> A <b>L</b> K <b>E</b> R <b>H</b> C <b>I</b> Q <b>R</b> K <b>F</b> S <b>S</b> K <b>A</b> K KCKG <b>K</b> GT <b>S</b> ED
GmERF#169	097	1.50e <sup>-32</sup>	K <b>T</b> V <b>K</b> SL <b>Q</b> E <b>I</b> <b>Q</b> Y <b>S</b> C <b>C</b> N <b>G</b> S <b>S</b> E <b>P</b> A <b>L</b> A <b>L</b> K <b>E</b> R <b>H</b> C <b>I</b> Q <b>R</b> K <b>L</b> S <b>S</b> K <b>A</b> K MCR <b>G</b> K <b>E</b> T <b>S</b> S
GmERF#171	186	7.06e <sup>-24</sup>	ESL <b>K</b> D <b>M</b> Q <b>N</b> N <b>N</b> <b>K</b> P <b>S</b> L <b>E</b> D <b>G</b> S <b>S</b> E <b>V</b> L <b>A</b> L <b>K</b> R <b>K</b> H <b>T</b> M <b>R</b> R <b>K</b> S <b>K</b> A <b>T</b> N <b>K</b> K <b>T</b> K <b>R</b> G <b>H</b> G <b>M</b> E <b>L</b>
GmERF#177	183	1.93e <sup>-14</sup>	ER <b>G</b> AN <b>A</b> L <b>A</b> C <b>D</b> <b>H</b> K <b>G</b> T <b>N</b> <b>L</b> A <b>C</b> T <b>C</b> I <b>G</b> N <b>K</b> E <b>K</b> L <b>I</b> N <b>S</b> R <b>K</b> R <b>A</b> S <b>M</b> E <b>R</b> EAF <b>D</b> D <b>M</b> V <b>N</b> E <b>Q</b>
GmERF#207	181	2.66e <sup>-14</sup>	EM <b>V</b> AN <b>A</b> L <b>A</b> C <b>D</b> <b>H</b> K <b>A</b> V <b>T</b> H <b>L</b> A <b>C</b> T <b>C</b> I <b>G</b> N <b>K</b> E <b>K</b> L <b>I</b> N <b>S</b> R <b>K</b> R <b>A</b> S <b>R</b> E <b>R</b> EAF <b>D</b> D <b>I</b> M <b>N</b> E <b>Q</b>
GmERF#245	165	1.03e <sup>-15</sup>	K <b>R</b> S <b>L</b> L <b>E</b> I <b>S</b> S <b>P</b> <b>S</b> S <b>S</b> C <b>S</b> D <b>D</b> S <b>S</b> E <b>S</b> Q <b>G</b> T <b>K</b> R <b>R</b> K <b>S</b> L <b>A</b> E <b>L</b> L <b>N</b> K <b>L</b> A <b>K</b> NRS <b>Q</b> V <b>K</b> VE <b>C</b>
GmERF#255	141	1.80e <sup>-22</sup>	R <b>V</b> K <b>E</b> S <b>L</b> Q <b>E</b> M <b>N</b> <b>Y</b> S <b>G</b> C <b>S</b> R <b>G</b> C <b>S</b> E <b>A</b> L <b>E</b> L <b>K</b> E <b>R</b> H <b>N</b> I <b>R</b> R <b>K</b> L <b>S</b> S <b>S</b> S <b>K</b> AK <b>T</b> S <b>K</b> R <b>K</b> Q <b>D</b> L
GmERF#256	109	7.31e <sup>-28</sup>	V <b>V</b> SE <b>S</b> L <b>K</b> D <b>M</b> P <b>S</b> K <b>P</b> W <b>E</b> E <b>G</b> S <b>S</b> E <b>V</b> L <b>A</b> L <b>K</b> R <b>K</b> H <b>T</b> M <b>R</b> R <b>K</b> S <b>N</b> S <b>K</b> K <b>K</b> SK <b>I</b> D <b>Q</b> R <b>E</b> Q <b>L</b> D
GmERF#259	103	1.84e <sup>-20</sup>	V <b>R</b> Q <b>S</b> L <b>R</b> N <b>M</b> E <b>Q</b> <b>Y</b> G <b>S</b> F <b>Q</b> E <b>G</b> S <b>S</b> E <b>V</b> L <b>E</b> L <b>K</b> R <b>H</b> S <b>M</b> K <b>R</b> K <b>S</b> E <b>T</b> K <b>K</b> H <b>Q</b> L <b>H</b> D <b>V</b> M <b>V</b> LED
GmERF#274	147	1.80e <sup>-22</sup>	ESL <b>K</b> EM <b>N</b> Y <b>A</b> N <b>D</b> S <b>N</b> N <b>E</b> E <b>G</b> C <b>S</b> E <b>V</b> V <b>A</b> L <b>K</b> R <b>K</b> H <b>S</b> L <b>R</b> R <b>K</b> I <b>S</b> V <b>K</b> K <b>N</b> N <b>N</b> S <b>K</b> L <b>Q</b> S <b>S</b> T <b>V</b>
Consenso			[SQ] [KSY] SC [ES] [EDN] [GD] [SC] SP [VA] L [AG] LK [RE] [RK] H [SCT] [ILM] [RQ] RK [LS] [SN] [SA] K [AK] [KR]

Tabela S74. Motivo IX-3

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#005	005	1.37e <sup>-17</sup>	MYGQ <b>SSYESDLALLDSIRRHLLGDS</b> EEHKFGAPNV
GmERF#032	003	6.64e <sup>-16</sup>	MA <b>ATTSCDFSSLEPLQHYLLKHG</b> SKHLVSNNGVE
GmERF#033	003	2.93e <sup>-16</sup>	MA <b>TTISCDFSSVESIQHYLLELE</b> HDFNTLMNAS
GmERF#034	007	5.74e <sup>-17</sup>	MNNFNT <b>TLESDDFSLLESIQRYLLDDN</b> EDFNALTEVL
GmERF#063	001	1.13e <sup>-15</sup>	<b>MANAAEVSAKRIKRLHLLGEL</b> SPLATPQNKF
GmERF#064	010	4.10e <sup>-18</sup>	MYGQSSSS <b>SSYESDLALLDSIRRHLLGES</b> ESIFGAPNFG
GmERF#095	003	5.74e <sup>-17</sup>	MA <b>TTISCDFSSMESIQHYLLEPE</b> HDSNALMNIS
GmERF#096	007	9.37e <sup>-17</sup>	MNNFNT <b>TLESDDFALLESIQRYLLDDN</b> EDFNSLTEVL
GmERF#122	004	7.71e <sup>-17</sup>	MAS <b>ATVSDDFEFLESVQQYLLGHD</b> SVNLMPEQQ
GmERF#133	001	3.47e <sup>-17</sup>	<b>MTTAEKTSNLDLIRQLHLLGEN</b> IISDSSSFVS
GmERF#138	001	5.19e <sup>-17</sup>	<b>MATAEEISTLKLRLQYLLGDI</b> SDPFLTNLIP
GmERF#149	001	2.68e <sup>-15</sup>	<b>MANALEVSALNRKRLHLLGEL</b> SPLTQFQEPN
GmERF#150	005	3.27e <sup>-18</sup>	MYGQ <b>SCYESDLALLDSIRRHLLGDS</b> EEHRFGAPNV
GmERF#231	001	1.35e <sup>-15</sup>	<b>MANAAEVSAALNRKRLHLLGEL</b> SPLATPLNYF
GmERF#232	006	3.47e <sup>-17</sup>	MYGRS <b>DSYESDLALLDSIRRHLLGES</b> ELIFGAPNFG
GmERF#244	010	2.63e <sup>-14</sup>	MDNYDATT <b>ATTESDMTSLDLIQHYMHDE</b> ANALNNNVEG
GmERF#245	004	7.98e <sup>-18</sup>	MVS <b>ATVSDDFAFLESVQQYLLGHD</b> SINLMSETHQ
GmERF#265	001	8.53e <sup>-20</sup>	<b>MTTAEETSNDLIRQLHLLGEN</b> IISDSSFFSN
GmERF#269	001	5.19e <sup>-17</sup>	<b>MATAEEISTLKLRLQYLLGDI</b> SDPFLTNLIP
<b>Consenso</b>			[MAT] [TA] [TY] [AES] [SE] [DE] [FL] [SA] LL [DE] [SL] I [QR] [QHR] [YH] LLG [DE] [NS]

Tabela S75. Motivo IX-4

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#005	060	6.55e <sup>-30</sup>	SRLYPCLSND <b>WGELPLKEDDSEDMLLYGVLDAVNVGWV</b> PSLDASSPES
GmERF#013	012	2.38e <sup>-21</sup>	DCSSVLHSSN <b>SSFFPFNENDPEEMLLYGMITSYQEKVII</b> KEEEVNSEEN
GmERF#029	052	3.65e <sup>-26</sup>	SFRNIFLTEN <b>WAEPLKENDTNDMVIYGALRDAATGWF</b> PANGNEVSNV
GmERF#064	058	7.92e <sup>-34</sup>	SSLDSCLSDN <b>WGELPFKEDDTEDMVLYGVLDAVNVGWV</b> PSLDAGSPES
GmERF#094	021	2.07e <sup>-26</sup>	SFRNIFLAKN <b>WAEPLKEDNAEDMVIYGVLREAAATMGWL</b> PANGNGVSKV
GmERF#123	006	1.87e <sup>-24</sup>	MDSSS <b>HQLLPFNENDPEDMLLYGIITSCQEKKVT</b> INQEQVNKKR
GmERF#134	028	1.38e <sup>-24</sup>	ESSSPEAFSW <b>EGYLPFNENDPEEMLLYGMIAATTEES</b> GERASSEESA
GmERF#150	060	4.81e <sup>-31</sup>	GRLYPCLSND <b>WGELPLMEDDSEDMLLYGVLDAVNVGWV</b> PSLDASSPES
GmERF#178	012	2.07e <sup>-24</sup>	GSKKTYEEEE <b>NLLLPFNENDPQDMVIYQVLEANALTNT</b> FLPQPTRNIT
GmERF#206	023	3.41e <sup>-24</sup>	TVLFSSHQTO <b>TELLPFNENDPQDMVIYQVLEANALTNT</b> FLPQRLHQPO
GmERF#232	054	2.38e <sup>-33</sup>	SSLDSCLSDD <b>WGELPFKEDDSEDMVLYGVLDAVNVGWV</b> PSLDAGSPES
GmERF#270	010	4.05e <sup>-18</sup>	MNGVGVAEQ <b>WSDVPPFRMDDADDMIIFNSLCDAKTTPRI</b> RDQSRNKLLR
GmERF#274	028	1.37e <sup>-25</sup>	ESSSPEAFSW <b>EGYLPFNENDPEEMLLYGMIAATTVDBS</b> FERTSSEESE
<b>Consenso</b>			WG [EL] LP [FL] [NK] E [ND] D [PS] E [DE] M [LV] [LI] YG [VM] [LI] R [DE] AV [NTA] VGW [VT]

Tabela S76. Motivo IX-5

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#005	108	1.20e <sup>-25</sup>	SFSSGFIPAV <b>TVKSEPDLEFAPEPVYIPPAQAEPVAVV</b> PAKGKHYRGV
GmERF#063	065	4.93e <sup>-22</sup>	FTDLLEPEIE <b>FPLFEFDSKPOVIDLETPKTLISAEKKPQ</b> FNRKPSLLIA
GmERF#064	103	1.01e <sup>-29</sup>	SPESVSSGFP <b>AVKLEPDLMPALISPCPPPPAAAAEKKVV</b> PAKGKHYRGV
GmERF#150	108	6.49e <sup>-24</sup>	SFSSAFMPPV <b>TVKSETDLFEFAPEPICNPPVQGFAPAVV</b> PAKGKHYRGV
GmERF#231	056	8.79e <sup>-23</sup>	SSSVSLNHVF <b>TDLFEFDSKPOIIDLQTPKTLTSAQKKPQ</b> LNRKPSLLIA
GmERF#232	099	2.45e <sup>-29</sup>	SPESVSSGFP <b>AVKLEPDMVMPALINPCPPPPAPAVEEKKVV</b> PPKGKHYRGV
<b>Consenso</b>			[TA] V [KL] [FLS] E [PF] D [LS] [FKM] P [AQ] [LP] [IE] [DP] [LP] C [PT] P [PK] [AT] [AL] [AQ] S [AE] E [KP] [KA] [VP] [VQ]

Tabela S77. Motivo IX-6

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#063	241	3.06e <sup>-25</sup>	EHDVSCFRGM PLTPSMWTGFWDSDVKDIFN IPPLSPLSPF
GmERF#132	248	1.43e <sup>-25</sup>	SNKHVKKEEC PLTPSCWKGFWDTDVMTIF SVPLSPLSP
GmERF#149	236	4.94e <sup>-20</sup>	ESDVSLIKNN PLTPSSWLGFWGSDANGIFT VPPLSPLSPH
GmERF#231	238	3.06e <sup>-25</sup>	EQDVSCFREM PLTPSMWTGFWDSDVKDIFN VPPLSPLSPF
GmERF#264	234	1.43e <sup>-25</sup>	KEECSPKAVC PLTPSCWKGFWDTDVMTIF SVPLSPLSP
<b>Consenso</b>			PLTPS [CMS]W [KTL] GFW [DG] [ST]D [VA] [KMN] [GD] [IT] [FI] [FNT]

Tabela S78. Motivo IX-7

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#127	010	5.21e <sup>-30</sup>	VLFDMSST NAVLDPIHDSLSFDMVDFSTAPTTEGNH NDEAKQLVVK
GmERF#166	031	2.92e <sup>-31</sup>	SEILFDSFSS NEVRDTHDLSFDMVDFSTPTEGNHN KTRQLVVKSE
GmERF#167	032	4.22e <sup>-33</sup>	EILFDSMSSN NAVLDPIQDLSFDMVDFSTPTEGNHN EAKKLIVKSE
GmERF#255	030	2.05e <sup>-25</sup>	DILFESFSSK GVDHDPFHDCLSFDMIDNSRDPQESHQ QVIEEAMKTK
<b>Consenso</b>			[NG] [APV] [VD] [LHR]D [PT] [IFV] [HQ]D [SCL]LSFDM [VI]D [FN] S [TR] [ADPT] P [TQ] [ET] [EG] [NGS] [HN] [HN]

Tabela S79. Motivo X-1

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#020	076	8.29e <sup>-18</sup>	QGGTLVSGYN <b>RGD</b> FDL <b>RE</b> MVSAL <b>TH</b> VV <b>VS</b> SGSGQRSTEW
GmERF#021	014	3.44e <sup>-16</sup>	SSSSFFPAHI <b>TTD</b> Q <b>EF</b> SLIVAAL <b>TN</b> VV <b>SG</b> STSSSSVPEF
GmERF#056	025	1.86e <sup>-17</sup>	EYGLFPIYSE <b>RS</b> Q <b>Q</b> DMSAMV <b>SALT</b> Q <b>VI</b> GG SNSDSLHMHE
GmERF#057	016	1.90e <sup>-20</sup>	ALLSRLFPGI <b>NRE</b> REMSAMV <b>SALT</b> H <b>VV</b> SG EVPTGDSSSV
GmERF#069	033	7.69e <sup>-23</sup>	LSMPMMFSGL <b>NRE</b> KEMSAMISAL <b>TH</b> V <b>V</b> CG EDEHHGADYS
GmERF#082	025	1.86e <sup>-17</sup>	EYDLFPIYSE <b>RS</b> Q <b>Q</b> DMSAMV <b>SALT</b> Q <b>VI</b> GG SNSDSLQQHE
GmERF#099	012	5.96e <sup>-17</sup>	SSSSFSPATL <b>TTE</b> Q <b>EL</b> SVIVAAL <b>TN</b> V <b>V</b> AG STSASSSFSL
GmERF#106	033	1.36e <sup>-19</sup>	LSMPMMFSGF <b>NRE</b> KEMSAMIT <b>AL</b> TR <b>VV</b> CG DYSLVHHNID
GmERF#218	183	2.35e <sup>-12</sup>	LMQYAQLLWS <b>RDD</b> DL <b>Q</b> RIAS <b>GL</b> H <b>Q</b> HHH NEAFYDSSSS
GmERF#222	075	1.02e <sup>-17</sup>	QGGSLVSGYN <b>RGD</b> PE <b>L</b> REIV <b>SALT</b> H <b>VV</b> SS GSGQRSTEW
GmERF#251	194	4.65e <sup>-12</sup>	LMQYAQLLWS <b>RDD</b> DL <b>Q</b> RVAS <b>GL</b> HH <b>Q</b> HH HHHHPNEAF
<b>Consenso</b>			[RN]R[DE]Q[ED][LM]SA[MI]VSALT[HQ]VVSG

Tabela S80. Motivo X-2

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#017	145	2.79e <sup>-15</sup>	SLHNNSSLISA <b>EG</b> F <b>PN</b> LE <b>EY</b> AR <b>LL</b> SC SDDDDFQ RVA
GmERF#020	292	7.30e <sup>-14</sup>	IQQPQIQGSS <b>DL</b> IR <b>DY</b> LQ <b>YS</b> QL <b>LQ</b> S DFQQQQIQQQ
GmERF#111	119	1.08e <sup>-15</sup>	APPSLSTHDH <b>QV</b> F <b>PN</b> LY <b>Q</b> YA <b>Q</b> ILSS GDEEFPYYS
GmERF#140	139	6.89e <sup>-12</sup>	MQGTRFQDSS <b>DF</b> LR <b>DY</b> WG <b>YS</b> QL <b>LR</b> S TGEFHGLDHW
GmERF#218	168	1.63e <sup>-18</sup>	QPLLSSSAQE <b>EG</b> F <b>PN</b> LMQYA <b>QL</b> LWS RDDDDLQRIA
GmERF#219	156	2.55e <sup>-15</sup>	SSNNSSLSS <b>EG</b> F <b>PN</b> LE <b>EY</b> AR <b>LL</b> NC SDDDDFQ RVA
GmERF#222	277	7.30e <sup>-14</sup>	PENPQIQGSS <b>DL</b> IR <b>DY</b> LQ <b>YS</b> QL <b>LQ</b> S DFQQQQIQQQ
GmERF#248	171	1.08e <sup>-15</sup>	SLSTHDQGH <b>QV</b> F <b>PN</b> LY <b>Q</b> YA <b>Q</b> ILSS DDAQLPYFYS
GmERF#251	179	1.63e <sup>-18</sup>	QPLLSSSSQE <b>EG</b> F <b>PN</b> LMQYA <b>QL</b> LWS RDDDDLQ RVA
<b>Consenso</b>			[EDQ][GLV][FI][PR][ND][LY][ELMY][QE]Y [AS][QR][LI]L[SQW][SC]

Tabela S81. Motivo X-3

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#017	177	7.03e <sup>-20</sup>	RVALGLYQHH <b>NN</b> ED <b>FI</b> C <b>SS</b> Q <b>PP</b> P <b>V</b> PF <b>V</b> SSSSSSSAMT
GmERF#020	166	2.11e <sup>-16</sup>	QQSAPRLFR <b>NIG</b> DF <b>MC</b> N <b>LC</b> V <b>LS</b> Y <b>H</b> F <b>Y</b> F NQRQKKPPP
GmERF#021	064	3.64e <sup>-22</sup>	ASLMPPPLPS <b>I</b> TC <b>REC</b> NI <b>AG</b> CL <b>GC</b> N <b>FF</b> P EEQRKKYRGV
GmERF#099	066	8.93e <sup>-22</sup>	CLMPPPPPPS <b>MD</b> TC <b>REC</b> NI <b>AG</b> CL <b>GC</b> N <b>FF</b> P EEKKKQKRAK
GmERF#219	188	9.32e <sup>-17</sup>	RVALGLYQHH <b>NN</b> ED <b>YI</b> G <b>SS</b> Q <b>PP</b> P <b>V</b> PF <b>V</b> SSSSSSSAM
<b>Consenso</b>			[NIM][NDHI][ETG][DC][FRY][EIM][CY][NG][ISH] [ASL][GQC][CPV][LP][GPS][CVY][NPH]F[FY] [PVF]

Tabela S82. Motivo X-4

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#017	011	2.45e <sup>-18</sup>	MSNNKNNTFF <b>PT</b> PH <b>Q</b> I <b>ST</b> SE <b>FV</b> NS <b>Q</b> NP PQATNQGIDD
GmERF#218	016	3.91e <sup>-22</sup>	TGDDSTLSSP <b>IP</b> Q <b>Y</b> G <b>Q</b> I <b>S</b> NS <b>CL</b> I <b>ET</b> Q <b>D</b> P SLQPLNQGNN
GmERF#219	011	4.88e <sup>-17</sup>	MSTNQNTTFF <b>PL</b> PH <b>Q</b> I <b>TT</b> SE <b>FV</b> NS <b>Q</b> NN PSQATNQGID
GmERF#251	028	3.91e <sup>-22</sup>	TSDDSTPSTP <b>IP</b> Q <b>Y</b> G <b>Q</b> I <b>S</b> NS <b>CL</b> I <b>ET</b> Q <b>D</b> P SQQPLNQGNN
<b>Consenso</b>			[IP][PLT][PQ][HY]GQI[ST][NT]S[CE][FL][IV] [EN][ST]Q[DN][PN]

Tabela S83. Motivo X-5

Nome Genérico	Início	Valor <i>p</i>	Sítio
GmERF#020	130	8.00e <sup>-20</sup>	CLSSFSSPSS <b>SG</b> AS <b>W</b> V <b>G</b> H <b>K</b> R <b>G</b> RE <b>E</b> E <b>E</b> SS <b>S</b> SH NMIQQQQQSA
GmERF#056	001	6.17e <sup>-22</sup>	<b>ME</b> GR <b>S</b> I <b>S</b> H <b>S</b> S <b>E</b> RE <b>E</b> E <b>Y</b> GL <b>F</b> PI YSERSQQDMS
GmERF#082	001	2.60e <sup>-22</sup>	<b>ME</b> GR <b>S</b> I <b>S</b> H <b>S</b> S <b>E</b> RE <b>E</b> E <b>Y</b> DL <b>F</b> PI YSERSQQDMS
GmERF#222	131	3.66e <sup>-21</sup>	AFSSSSPSPS <b>SG</b> AS <b>W</b> V <b>G</b> H <b>K</b> R <b>G</b> RE <b>E</b> E <b>E</b> NS <b>T</b> SH NLMQQQQQSA
<b>Consenso</b>			[MS][EG][AG][RS][SW][IV][GS]H[KS][RS][EG]REE E[EY][DGNS][LS][FST][PS][HI]



Tabela S84. Sumário dos Motivos Identificados em Membros da Superfamília de Fatores de Transcrição AP2/ERF

Motivos	Membros da Superfamília AP2/ERF com Motivos Similares em Diferentes Plantas <sup>1</sup>		Possível Função dos Motivos	
	Número de Acesso	Principais Correspondências	Número do Bloco	Melhor Homologia do Motivo - Banco de Dados LAMA
I-1	-	Sem Correspondência Identificada	-	Sem Correspondência Identificada
I-2	-	Sem Correspondência Identificada	IPB001427B	Ribonuclease Pancreática
I-3	-	Sem Correspondência Identificada	IPB006926B	VPS16, N-terminal
I-4	-	Sem Correspondência Identificada	IPB007582B	Região WD40 Associada com a Subunidade TFIID
I-5	-	Sem Correspondência Identificada	IPB003102A	Coativador CBP, pKID
I-6	-	Sem Correspondência Identificada	IPB001471A	Proteína com Elemento de Interação Responsivo à Etileno, Assinatura
I-7	I3SH08	Proteína Não Caracterizada ( <i>Lotus japonicus</i> )	-	Sem Correspondência Identificada
	XP_003591645	Fator de Transcrição Responsivo à Etileno RAP2-3 ( <i>Medicago truncatula</i> )		
II-1	XP_007144182	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )	IPB004518C	Proteína da Família MAZ G
	-	Sem Correspondência Identificada		
II-2	W6D834	Proteína do Grupo DREB ( <i>Ammopiptanthus mongolicus</i> )	IPB007001F	Proteína Bacteriana Shufflon, N-terminal Motivo EAR <sup>2</sup>
	XP_004499041	Fator de Transcrição Responsivo à Etileno Semelhante à RAP2-1 ( <i>Cicer arietinum</i> )		
	XP_004499041	Proteína com Elemento de Interação Responsivo à Desidratação ( <i>Halimodendron halodendron</i> )		
	XP_003589156	Fator de Transcrição Responsivo à Etileno ( <i>Medicago truncatula</i> )		
	XP_003603078	Fator de Transcrição Responsivo à Etileno RAP2-1 ( <i>Medicago truncatula</i> )		
	V5NAL1	Fator de Transcrição Responsivo à Etileno RAP2 ( <i>Phaseolus vulgaris</i> )		
III-1	XP_007140279	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )	IPB004926A	Proteína Abundante no Final da Embriogênese 3
	XP_004508937	Fator de Transcrição Responsivo à Etileno Semelhante à ERF034 ( <i>Cicer arietinum</i> )		
	XP_006421614	Proteína Não Caracterizada ( <i>Citrus clementina</i> )		
	XP_006490227	Fator de Transcrição Responsivo à Etileno Semelhante à ERF035 ( <i>Citrus sinensis</i> )		
	XP_006477148	Fator de Transcrição Responsivo à Etileno Semelhante à ERF038 ( <i>Citrus sinensis</i> )		
	V5RF18	Fator de Transcrição Responsivo à Etileno 4 ( <i>Coffea arabica</i> )		
	XP_004301177	Fator de Transcrição Responsivo à Etileno Semelhante à ERF034 ( <i>Fragaria vesca</i> )		
	W6FIZ0	Fator com Elemento de Interação Responsivo à Desidratação ( <i>Morus notabilis</i> )		
	XP_008795185	Proteína com Elemento de Interação Responsivo à Desidratação 3 ( <i>Phoenix dactylifera</i> )		
	XP_008238439	Fator de Transcrição Responsivo à Etileno ERF039 ( <i>Prunus mume</i> )		
	XP_007209502	Proteína Não Caracterizada ( <i>Prunus persica</i> )		
	XP_006359467	Proteína com Elemento de Interação Responsivo à Desidratação 3 ( <i>Solanum tuberosum</i> )		
	G3K514	Proteína com Elemento de Interação Responsivo à Desidratação DREB1 ( <i>Solanum tuberosum</i> )		
	XP_006347933	Fator de Transcrição Responsivo à Etileno ERF034 ( <i>Solanum tuberosum</i> )		
	M1KCM7	Fator Responsivo à Etileno 13 ( <i>Tamarix hispida</i> )		
	XP_002271812	Fator de Transcrição Responsivo à Etileno ERF034 ( <i>Vitis vinifera</i> )		
	XP_003635442	Fator de Transcrição Responsivo à Etileno ERF034 ( <i>Vitis vinifera</i> )		
A5AEB6	Proteína Não Caracterizada ( <i>Vitis vinifera</i> )			
III-2	-	Sem Correspondência Identificada	-	Motivo LWSY <sup>2</sup>
III-3	-	Sem Correspondência Identificada	IPB007441B	Proteína de Utilização de Etanolamina, EUT H
III-4	-	Sem Correspondência Identificada	IPB002653B	Dedo de Zinco, Semelhante à A20
III-5	-	Sem Correspondência Identificada	IPB007576D	CITED
III-6	XP_002865886	Fator de Interação - Repetição C 4 ( <i>Arabidopsis lyrata</i> )	IPB006174A	Metiltransferase 16S rRNA m5C967
	A3FMQ2	Fator de Interação - Repetição C ( <i>Ageratina adenofores</i> )		
	Q9FJ93	Proteína com Elemento de Interação Responsivo à Desidratação 1D ( <i>Arabidopsis thaliana</i> )		
	A5HIK0	CBF/DREB1 Fator de Transcrição 1 ( <i>Betula pendula</i> )		
	A5HIK2	CBF/DREB1 Fator de Transcrição 3 ( <i>Betula pendula</i> )		
	A5IZK1	DREB1-1 ( <i>Brassica juncea</i> )		
	A5IZK2	DREB1-2 ( <i>Brassica juncea</i> )		
	XP_009106963	Proteína com Elemento de Interação Responsivo à Desidratação 1D ( <i>Brassica rapa</i> )		
	XP_004496614	Proteína com Elemento de Interação Responsivo à Desidratação 1A ( <i>Cicer arietinum</i> )		
	XP_006421846	Proteína com Elemento de Interação Responsivo à Desidratação 1D ( <i>Citrus sinensis</i> )		
	XP_008439162	Proteína com Elemento de Interação Responsivo à Desidratação 1A ( <i>Cucumis melo</i> )		
	XP_008467214	Proteína com Elemento de Interação Responsivo à Desidratação 1D ( <i>Cucumis melo</i> )		
	XP_004140794	Proteína com Elemento de Interação Responsivo à Desidratação 1C ( <i>Cucumis sativus</i> )		
	XP_004143704	Proteína com Elemento de Interação Responsivo à Desidratação 1D ( <i>Cucumis sativus</i> )		
	XP_004308117	Proteína com Elemento de Interação Responsivo à Desidratação 1D ( <i>Fragaria vesca</i> )		
	XP_008376515	Proteína com Elemento de Interação Responsivo à Desidratação 1D ( <i>Malus domestica</i> )		
	A2WZ14	Proteína com Elemento de Interação Responsivo à Desidratação 1F ( <i>Oryza sativa</i> )		
	XP_008807522	Proteína com Elemento de Interação Responsivo à Desidratação 1F ( <i>Phoenix dactylifera</i> )		
	XP_008234474	Proteína com Elemento de Interação Responsivo à Desidratação 1B ( <i>Prunus mume</i> )		
	XP_002510847	Proteína com Elemento de Interação Responsivo à Desidratação 1A, Putativa ( <i>Ricinus communis</i> )		
	XP_004971403	Proteína com Elemento de Interação Responsivo à Desidratação 1F ( <i>Setaria italica</i> )		
	XP_004228864	Proteína com Elemento de Interação Responsivo à Desidratação 1C ( <i>Solanum lycopersicum</i> )		
	XP_004244599	Proteína com Elemento de Interação Responsivo à Desidratação 1F ( <i>Solanum lycopersicum</i> )		
	XP_006358387	Proteína com Elemento de Interação Responsivo à Desidratação 1D ( <i>Solanum tuberosum</i> )		
	XP_006358389	Proteína com Elemento de Interação Responsivo à Desidratação 1F ( <i>Solanum tuberosum</i> )		
	XP_007015927	Repetição C/Fator de Interação DRE 1, Putativa ( <i>Theobroma cacao</i> )		
	XP_002267961	Proteína com Elemento de Interação Responsivo à Desidratação 1D ( <i>Vitis vinifera</i> )		
	XP_002276153	Proteína com Elemento de Interação Responsivo à Desidratação 1E ( <i>Vitis vinifera</i> )		
	XP_002277273	Proteína com Elemento de Interação Responsivo à Desidratação 1F ( <i>Vitis vinifera</i> )		
XP_008674480	Proteína com Elemento de Interação Responsivo à Desidratação 1F ( <i>Zea mays</i> )			
III-7	-	Sem Correspondência Identificada	IPB002120C	Receptor de Liberação do Hormônio Tirotrófina, Assinatura
IV-1	-	Sem Correspondência Identificada	IPB007151E	Proteína MPP10
IV-2	XP_007162244	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )	-	Sem Correspondência Identificada
IV-3	-	Sem Correspondência Identificada	IPB003076B	Receptor Alfa Ativado por Proliferação de Peroxissomo, Assinatura
IV-4	-	Sem Correspondência Identificada	IPB007782Q	Carboxilase Gama Dependente de Vitamina K
IV-5	-	Sem Correspondência Identificada	IPB006952A	Fosfodiesterase cGMP Retinal, Subunidade Gama
IV-6	-	Sem Correspondência Identificada	IPB013674L	RNA Polimerase de Luteovirus P1-P2/Replicase
IV-7	-	Sem Correspondência Identificada	IPB000565F	Subunidade B da DNA Girase, Assinatura
IV-8	XP_004499266	Proteína com Elemento de Interação Responsivo à Desidratação 2D ( <i>Cicer arietinum</i> )	IPB008109D	Purinoreceptor P2Y13 Purinoreceptor, Assinatura
V-1	XP_004515544	Fator de Transcrição Responsivo à Etileno ERF003 ( <i>Cicer arietinum</i> )	IPB013681P	Fator de Transcrição de Mielina 1
	XP_006465255	Fator de Transcrição Responsivo à Etileno ERF003 ( <i>Citrus sinensis</i> )		
	XP_006475843	Fator de Transcrição Responsivo à Etileno WIN1 ( <i>Citrus sinensis</i> )		
	XP_008375424	Fator de Transcrição Responsivo à Etileno ERF003 ( <i>Malus domestica</i> )		
	XP_008792041	Fator de Transcrição Responsivo à Etileno WIN1 ( <i>Phoenix dactylifera</i> )		
	NP_001234848	Fator Responsivo à Etileno 1 ( <i>Solanum lycopersicum</i> )		
	XP_006353664	Fator de Transcrição Responsivo à Etileno ERF003 ( <i>Solanum tuberosum</i> )		
	XP_002268413	Fator de Transcrição Responsivo à Etileno WIN1 ( <i>Vitis vinifera</i> )		
XP_004305484	Fator de Transcrição Responsivo à Etileno ERF003 ( <i>Fragaria vesca</i> )			

**Tabela S84.** Sumário dos Motivos Identificados em Membros da Superfamília de Fatores de Transcrição AP2/ERF (cont.)

Motivos	Membros da Superfamília AP2/ERF com Motivos Similares em Diferentes Plantas <sup>1</sup>		Possível Função dos Motivos	
	Número de Acesso	Principais Correspondências	Número do Bloco	Melhor Homologia do Motivo - Banco de Dados LAMA
V-2	-	Sem Correspondência Identificada	-	Sem Correspondência Identificada
V-3	XP_007151265	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )	IPB010490C	Complexo Oligomérico Conservado COG6
V-4	XP_007135744	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )	IPB002693H	Fosfoproteína P <i>Paramyxovirus</i>
V-5	XP_004489494	Semelhante ao Fator de Transcrição Responsivo à Etileno RAP2-12 ( <i>Cicer arietinum</i> )	IPB008538C	Proteína Sem Função Conhecida DUF820
	XP_003618556	Fator de Transcrição Responsivo à Etileno ( <i>Medicago truncatula</i> )		
	XP_007151265	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )		
	XP_007161956	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )		
V-6	-	Sem Correspondência Identificada	IPB002813G	Proteína de Biossíntese de Arginina ARG J
VI-1	-	Sem Correspondência Identificada	IPB007991C	Fator de Iniciação Transcricional Específico da RNA Polimerase I RRN3
VI-2	XP_008437369	Semelhante ao Fator de Transcrição Responsivo à Etileno CRF2 ( <i>Cucumis melo</i> )	IPB008905B	Fator Eucariótico 3 de Iniciação da Tradução, Subunidade 8, N-terminal
	XP_004143957	Semelhante ao Fator de Transcrição Responsivo à Etileno CRF2 ( <i>Cucumis sativus</i> )		
	XP_004308016	Semelhante ao Fator de Transcrição Responsivo à Etileno CRF4 ( <i>Fragaria vesca</i> )		
	XP_008392454	Semelhante ao Fator de Transcrição Responsivo à Etileno CRF2 ( <i>Malus domestica</i> )		
	XP_007208855	Proteína Não Caracterizada ( <i>Prunus persica</i> )		
	XP_008239077	Semelhante ao Fator de Transcrição Responsivo à Etileno CRF2 ( <i>Prunus mume</i> )		
	XP_002509918	Interação com DNA, Putativa ( <i>Ricinus communis</i> )		
	XP_004234184	Semelhante ao Fator de Transcrição Responsivo à Etileno CRF4 ( <i>Solanum lycopersicum</i> )		
	XP_006350598	Semelhante ao Fator de Transcrição Responsivo à Etileno CRF4 ( <i>Solanum tuberosum</i> )		
VI-3	-	Sem Correspondência Identificada	IPB001038B	Glicoproteína C, Assinatura
VI-4	-	Sem Correspondência Identificada	-	Sem Correspondência Identificada
VI-5	XP_006422111	Fator de Transcrição Responsivo à Etileno 1A ( <i>Medicago truncatula</i> )	-	Sem Correspondência Identificada
	XP_007137563	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )		
	XP_007141007	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )		
VI-6	-	Sem Correspondência Identificada	IPB004972I	Núcleo da Proteína P4B de <i>Poxvirus</i>
VII-1	XP_004495934	Semelhante ao Fator de Transcrição Responsivo à Etileno RAP2-12 ( <i>Cicer arietinum</i> )	IPB004679A	Proteína Carreadora de Citrato
	B3IX38	Fator de Transcrição AP2-EREBP ( <i>Lotus japonicus</i> )		
VII-2	XP_007139068	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )	-	Sem Correspondência Identificada
VII-3	-	Sem Correspondência Identificada	IPB010528B	TOL A
VII-4	-	Sem Correspondência Identificada	-	Sem Correspondência Identificada
VII-5	-	Sem Correspondência Identificada	-	Sem Correspondência Identificada
VII-6	-	Sem Correspondência Identificada	IPB005007F	Proteína L3/FP4 de <i>Poxvirus</i>
VII-7	K0FAV2	Fator com Elemento de Interação Responsivo à Etileno 1 ( <i>Arachis hypogaea</i> )	-	Sem Correspondência Identificada
	XP_004510558	Semelhante ao Fator de Transcrição Responsivo à Etileno RAP2-12 Isoforma X1 ( <i>Cicer arietinum</i> )		
	Q8GTE5	Proteína Semelhante ao Fator de Transcrição EREBP ( <i>Cicer arietinum</i> )		
	I0B4A0	Fator Responsivo à Etileno 3 ( <i>Diospyros kaki</i> )		
	XP_007135380	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )		
	XP_007139068	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )		
VIII-1	XP_002533237	Fator de Transcrição Responsivo à Etileno, Putativo ( <i>Ricinus communis</i> )	IPB010941G	Poli Beta Hidroxibutirato Polimerase, N-terminal Motivo EAR <sup>2</sup>
	-	Sem Correspondência Identificada		
VIII-2	-	Sem Correspondência Identificada	IPB006986D	NAB1, C-terminal
VIII-3	-	Sem Correspondência Identificada	IPB000684E	RNA Polimerase II Eucariótica, Repetição de Heptapeptídeo
VIII-4	-	Sem Correspondência Identificada	-	Sem Correspondência Identificada
VIII-5	-	Sem Correspondência Identificada	IPB011591A	Neurotoxina Botulínica
VIII-6	XP_004304075	Semelhante ao Fator de Transcrição Responsivo à Etileno 3 ( <i>Fragaria vesca</i> )	IPB002723D	Proteína Sem Função Conhecida DUF43
	XP_007142329	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )		
VIII-7	XP_007154243	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )	IPB001471A	Proteína com Elemento de Interação Responsivo à Etileno, Assinatura
VIII-8	-	Sem Correspondência Identificada	IPB004935B	Proteína de 45/70 kDa de <i>Tymovirus</i>
IX-1	-	Sem Correspondência Identificada	-	Sem Correspondência Identificada
IX-2	-	Sem Correspondência Identificada	-	Sem Correspondência Identificada
IX-3	-	Sem Correspondência Identificada	IPB009539M	Estrabismo
IX-4	G8EVZ8	Fator Responsivo à Etileno 11 ( <i>Medicago sativa</i> )	IPB001404A	Proteína de Choque Térmico HSP90
	XP_003608658	Fator de Transcrição Responsivo à Etileno 1A ( <i>Medicago truncatula</i> )		
IX-5	-	Sem Correspondência Identificada	IPB006569B	Proteína de Regulação de Pré-mRNA Nuclear
IX-6	7DM20	Fator Responsivo à Etileno 1 ( <i>Cicer arietinum</i> )	IPB003327A	Zíper de Leucina, MYC
	XP_004513203	Semelhante ao Fator de Transcrição Responsivo à Etileno ERF104 ( <i>Cicer arietinum</i> )		
IX-7	-	Sem Correspondência Identificada	IPB002693A	Fosfoproteína P de <i>Paramyxovirus</i>
X-1	-	Sem Correspondência Identificada	IPB005640B	Lectina Hepática, N-terminal
X-2	XP_007139323	Proteína Não Caracterizada ( <i>Phaseolus vulgaris</i> )	IPB003888J	Domínio Rico em FY, N-terminal
X-3	-	Sem Correspondência Identificada	IPB013881D	Fator de Splicing de Pré-mRNA 3
X-4	-	Sem Correspondência Identificada	IPB003433P	Proteína de Capsídeo VP4
X-5	-	Sem Correspondência Identificada	-	Sem Correspondência Identificada

<sup>1</sup> Todas as plantas com sequências depositadas no banco de dados, excluindo *Glycine max.*

<sup>2</sup> Classificação de acordo com Nakano et al. (2006)

**Tabela S85.** Ciclo Limiar (Ct) dos Genes Alvo e Endógenos em Folha de Soja sob Déficit Hídrico - PCR em Tempo Real<sup>1</sup>

Símbolo	Nome Genérico	D.c. <sup>2</sup>	Tempo de Déficit Hídrico (minuto)													
			0		25		50		75		100		125		150	
			BR16 <sup>3</sup>	E48 <sup>4</sup>	BR16 <sup>3</sup>	E48 <sup>4</sup>	BR16 <sup>3</sup>	E48 <sup>4</sup>	BR16 <sup>3</sup>	E48 <sup>4</sup>	BR16 <sup>3</sup>	E48 <sup>4</sup>	BR16 <sup>3</sup>	E48 <sup>4</sup>	BR16 <sup>3</sup>	E48 <sup>4</sup>
ACO	GmACO#012	1:80	22,1±0,2	23,4±1,1	20,6±0,4	22,4±0,4	21,6±0,4	22,5±0,5	22,3±0,4	22,5±0,2	22,1±0,1	20,7±0,1	20,6±0,4	19,9±0,1	18,9±0,1	20,6±0,1
ACS	GmACS#006	1:40	30,9±0,2	35,0±0,8	29,8±0,2	30,2±0,2	29,5±0,2	31,4±0,1	31,1±0,1	31,3±0,2	30,4±0,2	30,7±0,6	31,4±0,3	32,9±1,2	32,4±0,4	33,2±0,4
CTR	GmCTR#003	1:40	30,2±0,1	30,6±0,1	29,9±0,1	30,7±0,1	30,9±0,2	30,5±0,1	31,1±0,1	30,4±0,1	31,2±0,2	29,8±0,2	30,0±0,2	31,9±0,1	31,1±0,1	31,8±0,1
ERF069	GmERF#069	1:80	30,7±0,2	31,0±0,1	29,5±0,3	30,2±0,4	28,5±1,6	28,3±0,5	26,6±0,7	24,1±0,1	25,8±0,8	26,7±0,2	24,4±0,8	23,8±0,2	21,0±0,1	25,3±0,5
ERF223	GmERF#223	1:80	34,9±0,2	27,5±0,1	29,9±0,1	30,9±0,2	24,8±0,2	26,6±0,1	26,4±0,1	29,3±0,1	28,1±0,1	26,7±0,1	27,2±0,1	27,7±0,1	27,3±0,2	25,7±0,2
ETR	GmETR#007	1:40	21,9±0,3	22,9±0,2	22,0±0,1	22,3±0,1	22,7±0,4	22,5±0,1	22,7±0,3	22,1±0,3	23,3±0,1	22,2±0,2	22,3±0,2	22,3±0,1	20,6±0,1	22,6±0,7
MAT	GmMAT#008	1:80	19,2±0,2	19,3±0,4	17,9±0,3	19,0±0,2	18,7±0,1	18,8±0,1	18,1±0,3	18,5±0,2	19,6±0,1	17,2±0,3	17,7±0,3	19,4±0,5	17,8±0,4	18,0±0,2
RAV	GmRAV#001	1:80	27,2±1,5	26,5±1,0	25,4±0,5	26,6±0,8	24,9±0,9	25,7±0,5	24,5±1,2	25,9±0,4	23,8±0,1	22,7±0,1	22,5±0,2	24,4±0,1	19,9±0,1	22,5±0,1
-	ACT11	1:40	18,4±0,5	18,9±0,8	17,7±0,6	18,3±0,1	18,5±0,4	18,5±0,1	19,1±0,5	18,4±0,2	18,7±1,2	18,1±0,6	18,8±0,3	19,7±0,5	19,6±0,1	20,3±0,4
-	UBC2	1:40	21,2±0,1	22,1±0,5	20,7±1,0	21,5±0,3	21,4±0,5	21,6±0,2	21,8±0,7	21,5±0,1	20,8±1,2	19,8±1,3	21,0±0,5	22,1±0,9	21,2±0,1	22,3±0,3
-	ACT11	1:80	20,6±0,4	21,1±0,9	19,3±0,7	20,3±0,1	20,7±0,6	20,6±0,3	21,0±0,6	20,7±0,4	21,9±0,3	20,1±0,1	20,4±0,7	21,7±0,9	20,2±0,6	21,3±0,5
-	UBC2	1:80	23,2±0,2	24,2±0,6	22,2±1,1	23,4±0,3	23,3±0,7	23,7±0,4	23,4±0,7	23,7±0,3	24,0±0,1	22,2±0,1	22,5±0,9	24,0±1,3	21,7±0,5	23,1±0,7

<sup>1</sup> O ciclo limiar (Ct) apresentado na tabela foi obtido pela média dos valores de todas as amostras técnicas e biológicas ± o valor do desvio padrão.

<sup>2</sup> Diluição do cDNA.

<sup>3</sup> Cultivar de soja BR16.

<sup>4</sup> Cultivar de soja EMBRAPA48.

**Tabela S86.** Ciclo Limiar (Ct) dos Genes Alvo e Endógenos em Raiz de Soja sob Déficit Hídrico - PCR em Tempo Real<sup>1</sup>

Símbolo	Nome Genérico	D.c. <sup>2</sup>	Tempo de Déficit Hídrico (minuto)													
			0		25		50		75		100		125		150	
			BR16 <sup>3</sup>	E48 <sup>4</sup>	BR16 <sup>3</sup>	E48 <sup>4</sup>	BR16 <sup>3</sup>	E48 <sup>4</sup>	BR16 <sup>3</sup>	E48 <sup>4</sup>	BR16 <sup>3</sup>	E48 <sup>4</sup>	BR16 <sup>3</sup>	E48 <sup>4</sup>	BR16 <sup>3</sup>	E48 <sup>4</sup>
ACO	GmACO#012	1:80	23,3±0,1	24,2±0,7	23,0±0,2	22,2±1,0	21,0±0,4	21,2±0,3	22,4±0,3	22,9±0,5	21,3±0,1	22,0±0,3	21,3±0,4	20,5±0,3	19,8±0,3	20,6±0,2
ACS	GmACS#006	1:40	30,5±0,3	30,3±0,3	27,8±0,2	26,4±0,1	27,2±0,1	27,6±0,1	28,9±0,2	30,9±0,3	32,9±0,1	30,6±0,3	29,4±0,2	30,0±0,2	29,6±0,6	31,5±0,3
CTR	GmCTR#003	1:40	29,1±0,2	28,1±0,1	30,0±0,3	29,5±0,1	30,9±0,2	30,7±0,3	31,8±0,1	30,9±0,4	33,3±0,2	31,4±0,4	31,7±0,3	31,9±0,4	32,1±0,4	32,0±0,1
ERF069	GmERF#069	1:80	32,4±0,2	29,9±0,1	29,2±0,1	27,0±0,6	23,4±0,1	23,8±0,7	23,1±0,2	25,1±0,4	23,9±0,2	23,4±0,4	23,3±0,4	24,6±0,2	23,8±0,6	24,5±0,2
ERF223	GmERF#223	1:80	34,3±0,1	28,9±0,2	28,0±0,6	27,5±0,5	28,0±1,1	27,7±0,9	30,2±1,3	31,3±0,4	34,2±0,1	34,2±0,6	31,9±0,2	32,1±0,7	30,4±1,0	30,5±1,1
ETR	GmETR#007	1:40	22,4±0,1	21,5±0,1	23,6±0,1	25,5±0,1	25,3±0,1	23,1±0,1	23,7±0,1	23,6±0,1	28,7±0,1	24,9±0,1	23,7±0,1	23,4±0,1	24,0±0,1	25,7±0,1
MAT	GmMAT#008	1:80	20,0±0,1	19,4±0,5	19,6±0,5	19,3±0,5	19,3±0,2	19,3±0,3	20,0±0,4	19,9±0,1	20,1±0,4	19,7±0,3	19,9±0,5	20,2±0,4	19,9±0,2	19,7±0,2
RAV	GmRAV#001	1:80	26,4±0,1	24,8±0,2	26,2±0,1	25,8±0,1	24,4±0,1	23,8±0,1	25,0±0,1	24,8±0,2	24,8±0,1	25,5±0,1	23,5±0,2	25,7±0,1	24,7±0,1	24,6±0,1
-	CYP2	1:40	19,1±0,1	19,0±0,6	19,6±0,4	18,9±0,3	20,1±0,1	19,3±0,5	20,5±0,7	19,6±0,3	24,2±1,4	20,0±0,8	20,2±0,3	20,3±0,4	20,8±0,1	21,4±0,5
-	ELF1A	1:40	19,2±0,2	19,1±0,7	19,7±0,3	19,1±0,4	19,9±0,1	19,9±0,6	21,1±0,7	20,3±0,2	23,9±0,9	21,1±0,6	21,3±0,2	21,4±0,4	22,3±0,2	22,5±0,4
-	CYP2	1:80	21,6±0,1	20,4±0,1	21,7±0,3	21,3±0,3	22,1±0,3	21,3±0,5	21,8±0,2	22,3±0,5	23,0±0,5	21,7±0,2	21,4±0,8	22,2±0,4	22,0±0,4	21,5±1,1
-	ELF1A	1:80	22,5±0,1	20,9±0,1	22,4±0,3	22,0±0,8	22,5±0,3	22,2±0,4	23,0±0,1	23,0±0,3	24,2±0,2	23,2±0,3	22,8±0,7	23,4±0,5	23,4±0,1	23,1±0,6

<sup>1</sup> O ciclo limiar (Ct) apresentado na tabela foi obtido pela média dos valores de todas as amostras técnicas e biológicas ± o valor do desvio padrão.

<sup>2</sup> Diluição do cDNA.

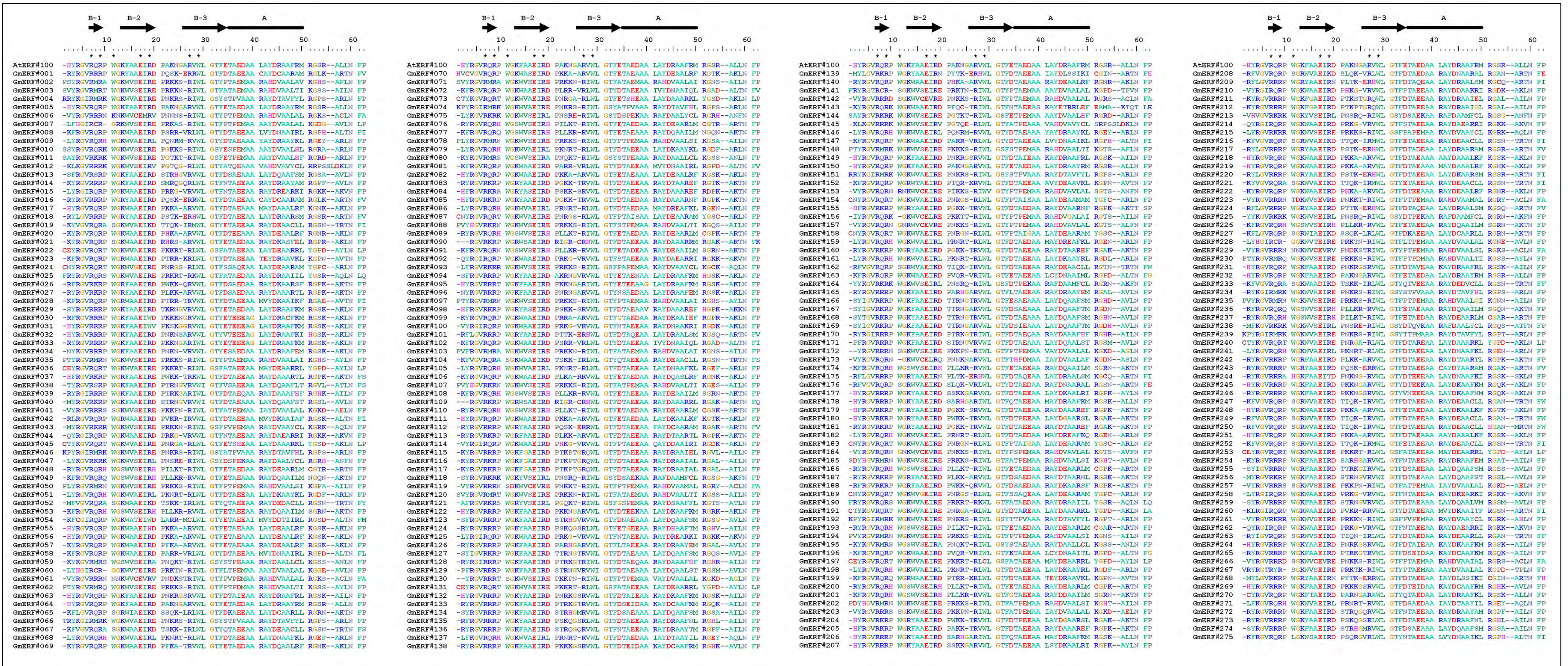
<sup>3</sup> Cultivar de soja BR16.

<sup>4</sup> Cultivar de soja EMBRAPA48.



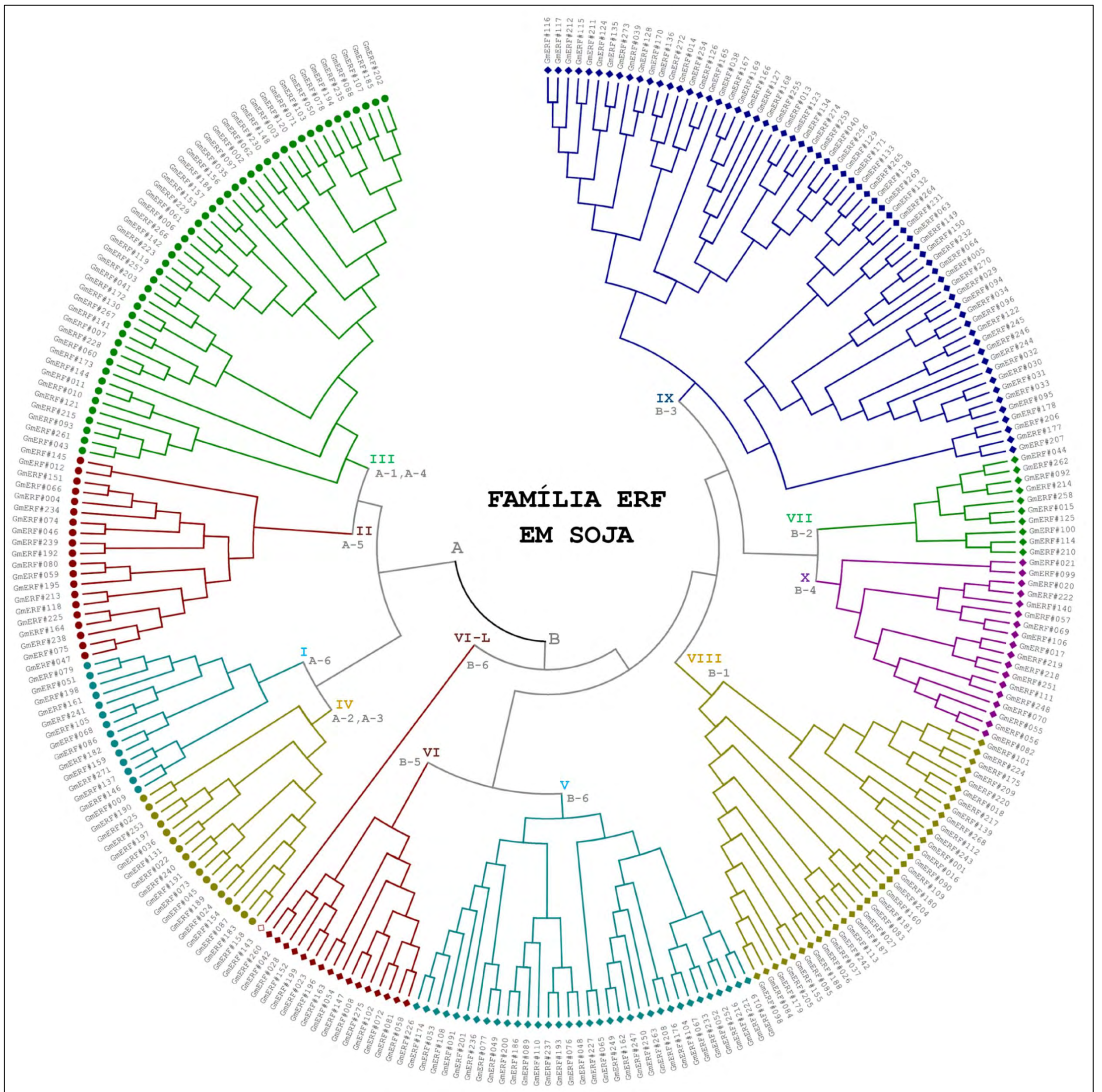






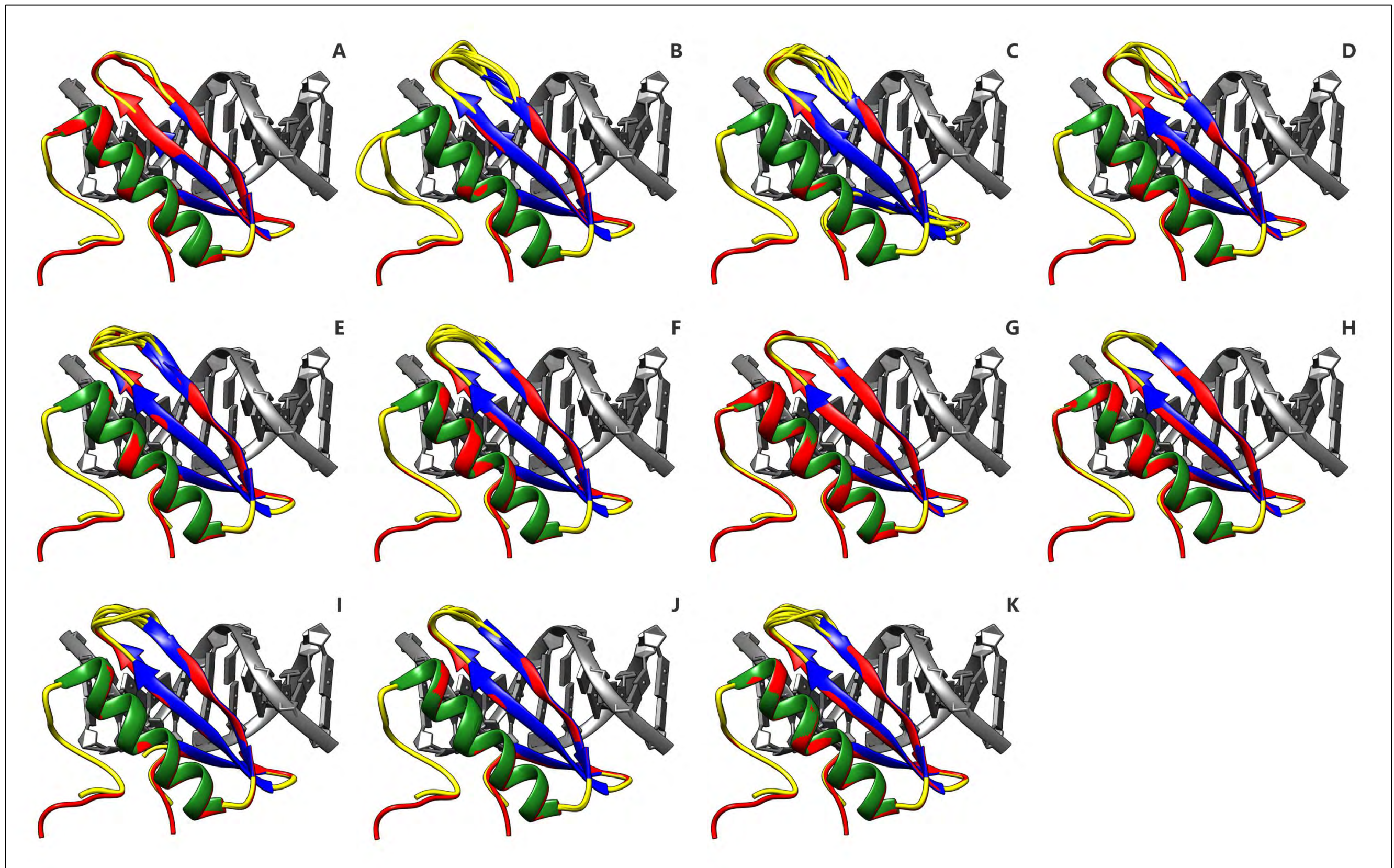
**Figura S2. Alinhamento da Sequência de Aminoácidos do Domínio AP2 da Família ERF de Fatores de Transcrição de Soja.** O alinhamento foi obtido comparando-se a sequência de aminoácidos do domínio AP2 do fator de transcrição AtERF#100 (At4g17500 - AtERF1A) de *Arabidopsis thaliana*, identificado por Allen et al. (1998), com seus homólogos em soja. Foram selecionadas apenas sequências de soja com domínio AP2 único e completo. Cada fator de transcrição é representado por um nome genérico (ver Tabela S14). Na figura, A representa região do domínio que apresenta alta conservação com sequências consenso de  $\alpha$ -hélices e B com de folhas  $\beta$ . Asteriscos (\*) indicam resíduos de aminoácidos conservados e com possível importância estrutural no domínio.





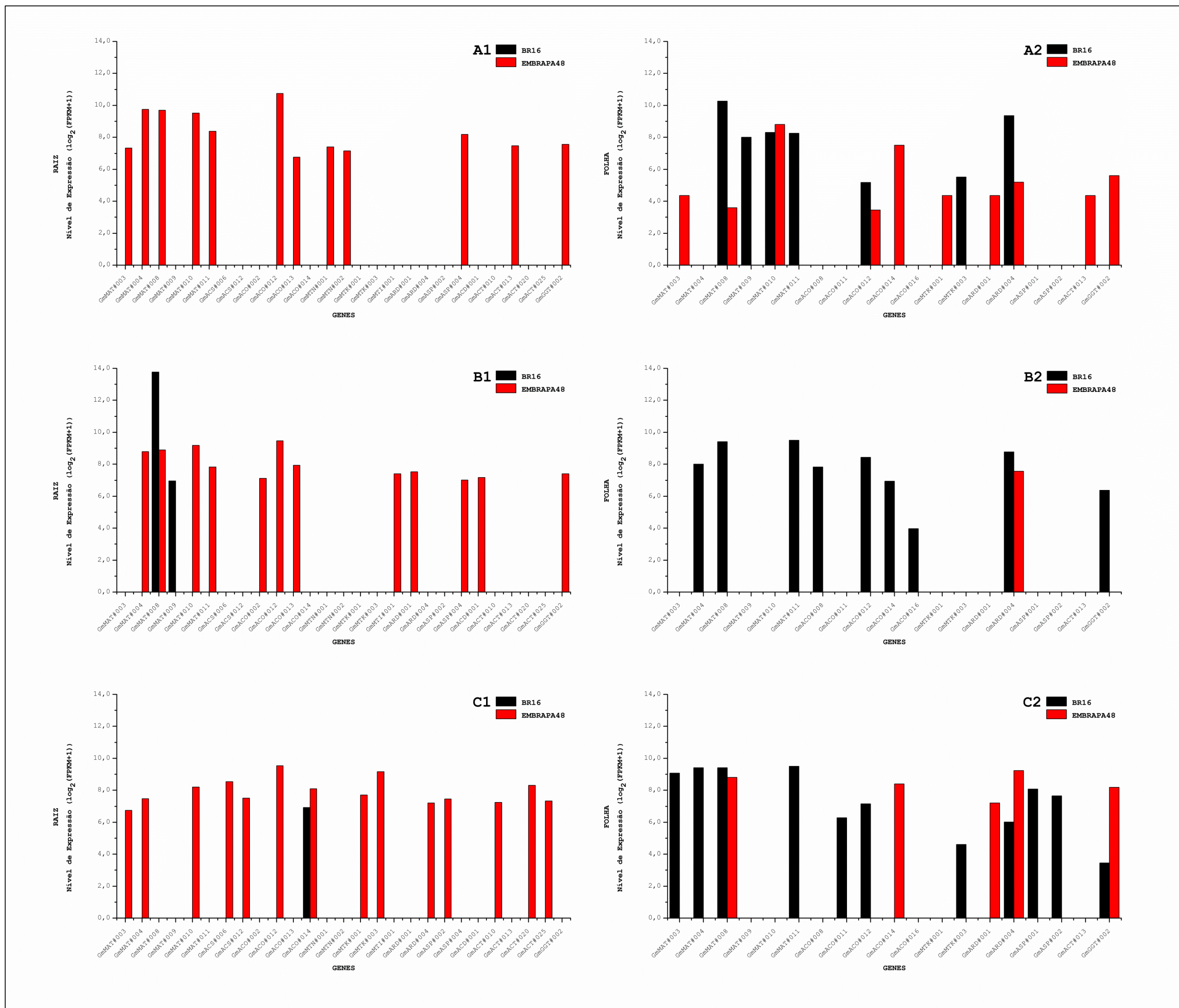
**Figura S3. Classificação dos Fatores de Transcrição ERF de Soja.** O dendograma representa os 275 membros da família ERF obtido a partir do alinhamento dos domínios AP2 de soja (ver Figura S2). As duas classificações apresentadas foram originadas pela comparação prévia dos genes ERF de soja com seus respectivos homólogos em *Arabidopsis thaliana*. A primeira classificação, baseada no trabalho de Sakuma et al. (2002), divide os ERFs de soja em dois grupos principais: **grupo A** (subfamília CBF/DREB; A-1 até A-6; **representados por círculos**) e grupo B (subfamília ERF; B-1 até B-6; **representados por losangos**). Nakano e colaboradores (2006) propuseram outra classificação para ERFs, e com base nela os 275 ERFs de soja foram subdivididos em onze grupos: grupos I-X e VI-L (destacados com cores diferentes). Cada fator de transcrição é representado por um nome genérico (ver Tabela S14).





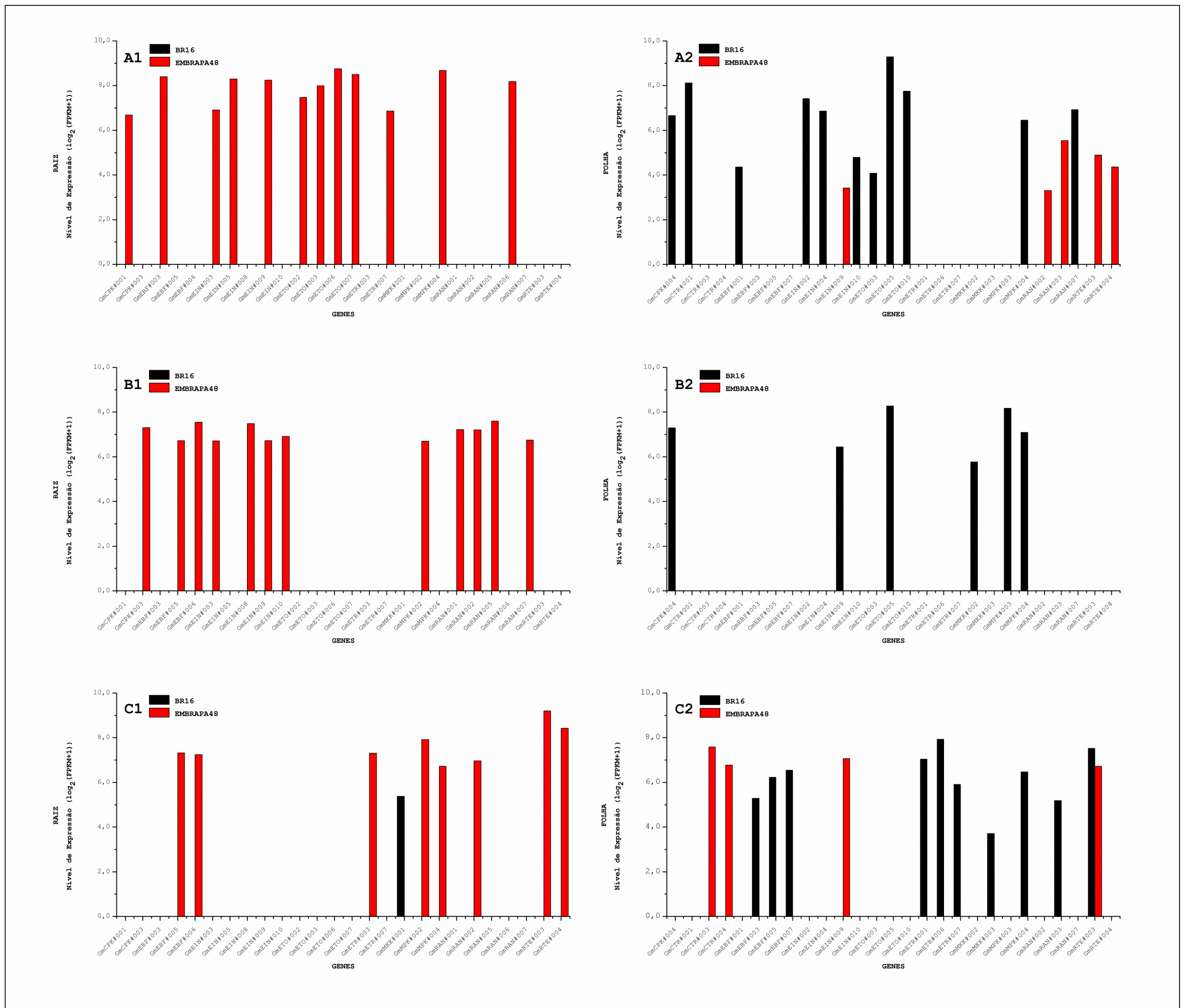
**Figura S4. Estrutura Terciária dos Domínios AP2 de Soja.** A figura ilustra as estruturas terciárias dos domínios AP2 dos 275 fatores de transcrição responsivos à etileno (ERF) de soja, obtidas por modelagem molecular. Os domínios em questão são apresentados de acordo com a classificação proposta por Nakano e colaboradores (2006), descrita na Figura S3. Cada membro de seu referido grupo apresenta sua estrutura sobreposta à estrutura cristalográfica deste domínio proposta para *Arabidopsis thaliana*, obtida em banco de dados de proteínas (PDB). Nos modelos, as estruturas secundárias em **verde**, **azul** e **amarelo** correspondem respectivamente à  $\alpha$ -hélices, folhas  $\beta$  e alças dos domínios AP2 de soja. Já as estruturas secundárias em **vermelho** correspondem ao modelo de *Arabidopsis thaliana*. A estrutura em **cinza** representa a sequência de DNA reconhecida pelo domínio. As Figuras S4-A à S4-F representam os membros dos grupos I ao VI; S4-G, o membro do grupo VI-L; e S4-H à S4-K os membros dos grupos VII ao X.





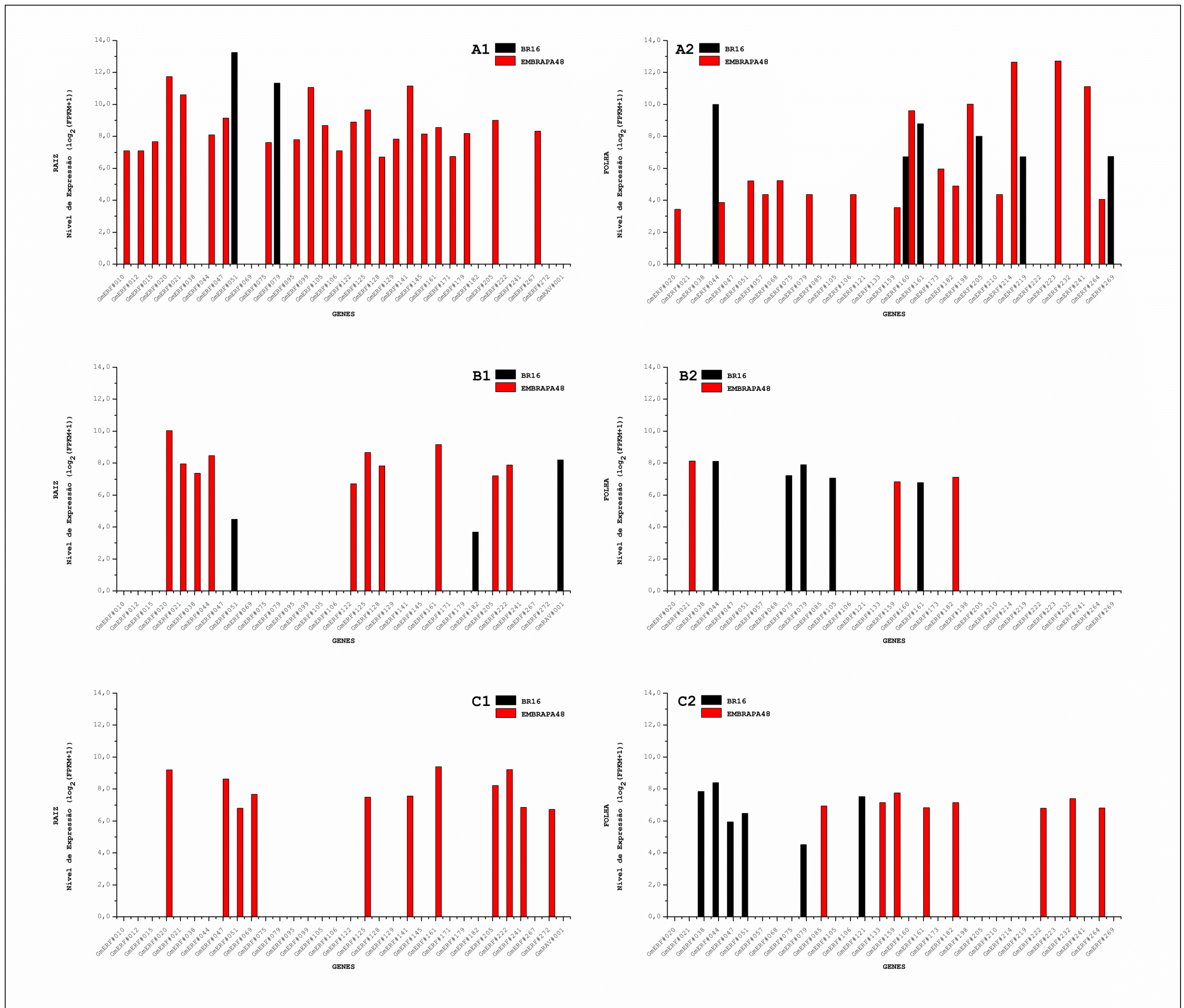
**Figura S5. Expressão Diferencial de Genes Relacionados à Biossíntese de Etileno em Transcriptomas de Soja Submetida à Déficit Hídrico.** Os gráficos representam os níveis de expressão dos genes relacionados à biossíntese do etileno em transcriptomas de raiz e folha de duas cultivares de soja: BR16 e EMBRAPA48, sensível e tolerante à seca, respectivamente. Cada gene é representado por um nome genérico (ver Tabela S11). As siglas correspondem a: **A1** - raiz/25 à 50 minutos sob condições de seca; **A2** - folha/25 à 50 minutos sob condições de seca; **B1** - raiz/75 à 100 minutos sob condições de seca; **B2** - folha/75 à 100 minutos sob condições de seca; **C1** - raiz/125 à 150 minutos sob condições de seca; **C2** - folha/125 à 150 minutos sob condições de seca; **FPKM** - fragmentos por quilobase de transcrito por milhão de fragmentos mapeados.





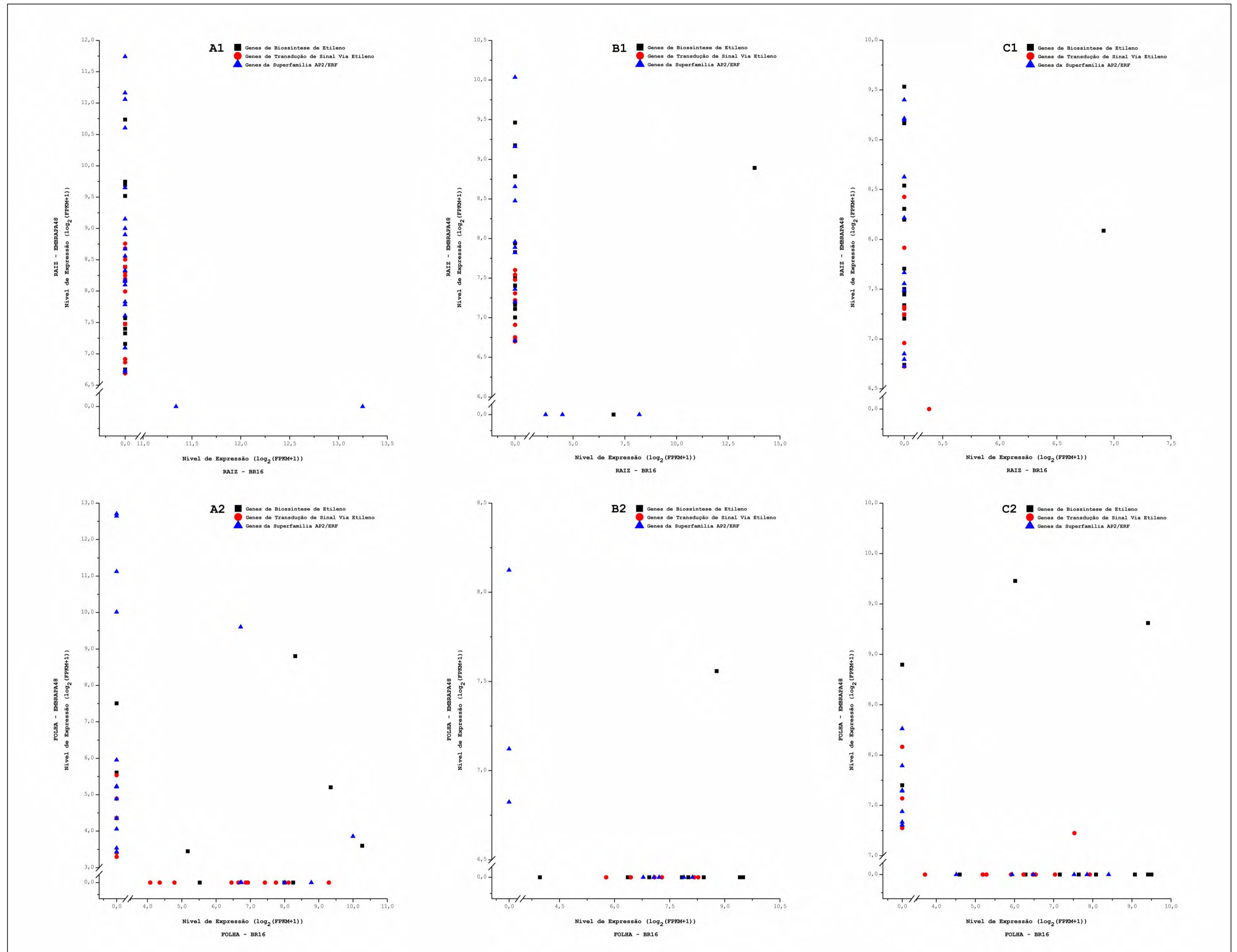
**Figura S6. Expressão Diferencial dos Genes Relacionados à Transdução de Sinal Mediada por Etileno em Transcriptomas de Soja Submetida à Déficit Hídrico.** Os gráficos representam os níveis de expressão dos genes relacionados à cascata de transdução de sinal mediada por etileno em transcriptomas de raiz e folha de duas cultivares de soja: BR16 e EMBRAPA48, sensível e tolerante à seca, respectivamente. Cada gene é representado por um nome genérico (ver Tabela S12). As siglas correspondem a: **A1** - raiz/25 à 50 minutos sob condições de seca; **A2** - folha/25 à 50 minutos sob condições de seca; **B1** - raiz/75 à 100 minutos sob condições de seca; **B2** - folha/75 à 100 minutos sob condições de seca; **C1** - raiz/125 à 150 minutos sob condições de seca; **C2** - folha/125 à 150 minutos sob condições de seca; FPKM - fragmentos por quilobase de transcrito por milhão de fragmentos mapeados.





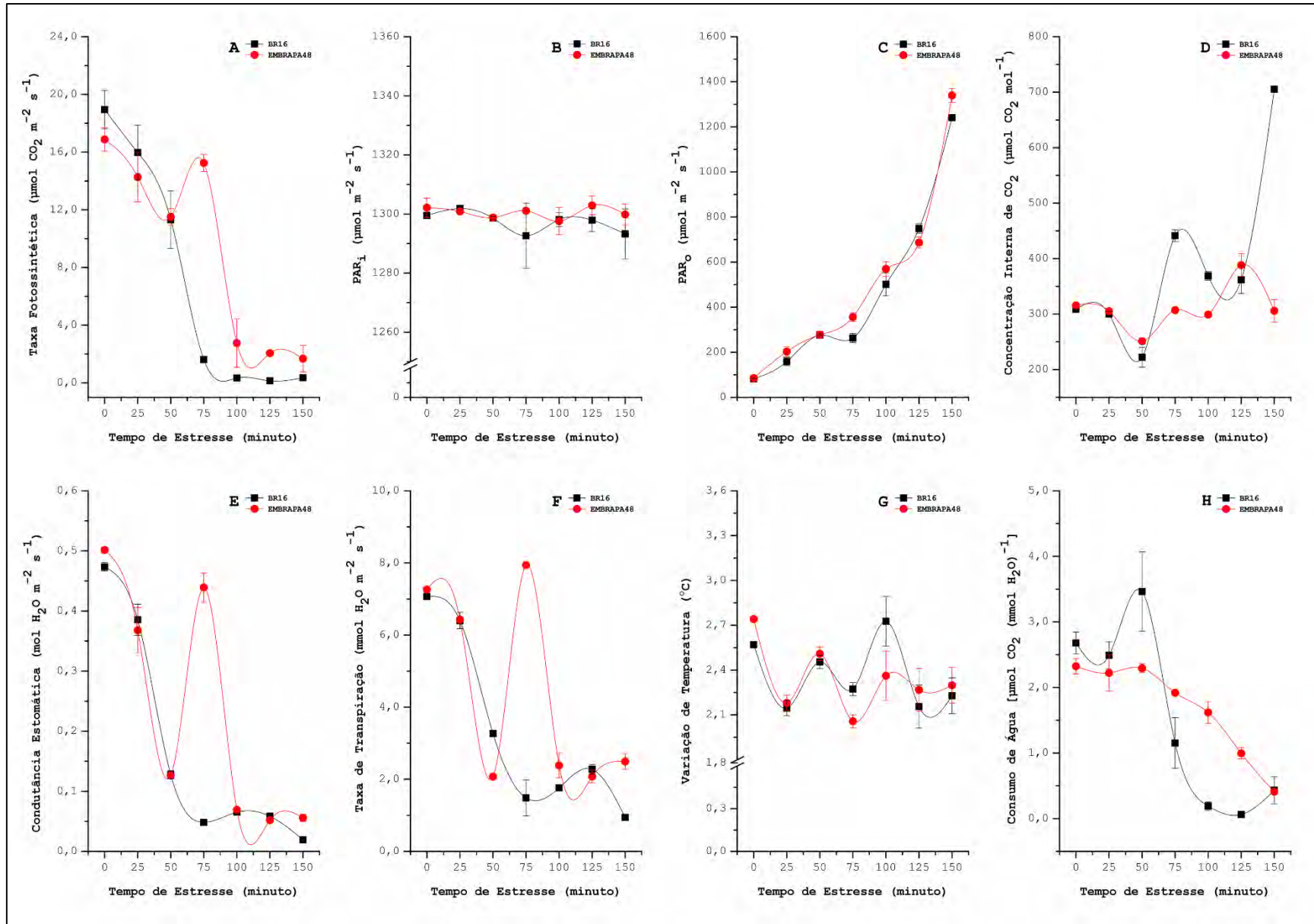
**Figura S7. Expressão Diferencial dos Genes dos Membros da Superfamília de Fatores de Transcrição AP2/ERF em Transcriptomas de Soja Submetida à Déficit Hídrico.** Os gráficos representam os níveis de expressão dos genes da superfamília AP2/ERF em transcriptomas de raiz e folha de duas cultivares de soja: BR16 e EMBRAPA48, sensível e tolerante à seca, respectivamente. Cada gene é representado por um nome genérico (ver Tabelas S13 - S15). As siglas correspondem a: **A1** - raiz/25 à 50 minutos sob condições de seca; **A2** - folha/25 à 50 minutos sob condições de seca; **B1** - raiz/75 à 100 minutos sob condições de seca; **B2** - folha/75 à 100 minutos sob condições de seca; **C1** - raiz/125 à 150 minutos sob condições de seca; **C2** - folha/125 à 150 minutos sob condições de seca; **FPKM** - fragmentos por quilobase de transcrito por milhão de fragmentos mapeados.





**Figura S8. Comparação Tecido-Específica da Expressão Diferencial dos Genes de Biossíntese e Sinalização por Etileno em Transcriptomas de Soja Submetida à Déficit Hídrico.** Os gráficos de dispersão comparam os níveis de expressão dos genes de biossíntese e sinalização por etileno (transdução de sinal e superfamília de fatores de transcrição AP2/ERF) entre tecidos semelhantes das cultivares de soja BR16 e EMBRAPA48, sensível e tolerante à seca, respectivamente. Cada gene expresso identificado (ver Figuras S5 - S7) é representado em cada gráfico por um ponto cujas coordenadas correspondem aos níveis de expressão em tecidos semelhantes nas duas cultivares. As siglas correspondem a: **A1** - raiz/25 à 50 minutos sob condições de seca; **A2** - folha/25 à 50 minutos sob condições de seca; **B1** - raiz/75 à 100 minutos sob condições de seca; **B2** - folha/75 à 100 minutos sob condições de seca; **C1** - raiz/125 à 150 minutos sob condições de seca; **C2** - folha/125 à 150 minutos sob condições de seca; **FPKM** - fragmentos por quilobase de transcrito por milhão de fragmentos mapeados.





**Figura S9. Avaliação de Parâmetros Fisiológicos Relevantes em Cultivares de Soja Sob Condições de Déficit Hídrico.** Durante o experimento de indução de déficit hídrico nas cultivares de soja BR16 e EMBRAPA48, cultivadas sob condições hidropônicas, foram determinados alguns parâmetros fisiológicos relevantes: **A** - taxa fotossintética ( $A$ ); **B** - radiação fotossinteticamente ativa (interna à câmara de leitura -  $PAR_i$ ); **C** - radiação fotossinteticamente ativa (externa à câmara de leitura -  $PAR_o$ ); **D** - concentração interna de  $\text{CO}_2$  ( $C_i$ ); **E** - condutância estomática ( $g_s$ ); **F** - taxa de transpiração ( $E$ ); **G** - variação de temperatura ( $\Delta T$ ), onde  $[\Delta T = T_{ar}$  (interna à câmara de leitura) -  $T_{folha}$ ; temperatura média do ar =  $29,7 \pm 1,9$   $^{\circ}\text{C}$ ]; **H** - eficiência do uso de água ( $WUE$ , relação entre a taxa fotossintética e a taxa de transpiração -  $A/E$ ). Cada ponto representa o valor médio ( $\pm$  desvio padrão) de três repetições, em plantas diferentes. A ausência de representação gráfica do desvio padrão em alguns pontos ocorre pelo fato de seus valores absolutos serem inferiores à escala.

# *CURRICULUM VITAE*

## **CURRICULUM VITAE RESUMIDO**

**ARRAES, F. B. M.**

### **1. DADOS PESSOAIS**

**Nome Completo:** Fabrício Barbosa Monteiro Arraes  
**Local e Data de Nascimento:** Brasília, DF, Brasil, 03/09/1981  
**Endereço Profissional:** EMBRAPA - Recursos Genéticos e Biotecnologia  
Laboratório de Interação Molecular Planta-  
Patógeno - Parque Estação Biológica - PqEB s/nº.  
Brasília, DF - Brasil - CEP 70770-901  
**Telefone Profissional:** (61) 3448-4705  
**E-mail:** fabricio.arraes@gmail.com

### **2. FORMAÇÃO ACADÊMICA**

**2001 - 2005** – Graduação em Ciências Biológicas  
Universidade de Brasília – UnB

### **3. EXPERIÊNCIA PROFISSIONAL**

**04/2003 - 10/2003** – Estágio de Iniciação Tecnológica - ITI  
Universidade de Brasília - UnB  
*Vínculo:* Laboratório de Biologia Molecular – Bolsista CNPq  
– Projeto Genoma  
*Orientadora:* Dra. Maria Sueli Soares Felipe  
*Projeto:* Sequenciamento do Genoma de Espécies de  
Mycoplasmas.  
**11/2003 - 08/2004** – Estágio de Iniciação Científica - PIBIC  
Universidade de Brasília - UnB  
*Vínculo:* Laboratório de Biologia Molecular  
*Orientadora:* Dra. Maria Sueli Soares Felipe

*Projeto:* Identificação, Isolamento e Caracterização do Gene do Repressor Transcricional *Pbtup1* do Fungo Dimórfico e Patogênico *Paracoccidioides brasiliensis*.

**09/2004 - 02/2005** – Estágio de Iniciação Tecnológica - ITI

Universidade de Brasília - UnB

*Vínculo:* Laboratório de Biologia Molecular

*Orientadora:* Dra. Maria Sueli Soares Felipe

*Projeto:* Transcriptoma Funcional do Fungo Dimórfico e Patogênico *Paracoccidioides brasiliensis*.

**01/2006 - 12/2006** – Estágio Remunerado/Estagiário Graduado

*Vínculo:* Laboratório Tecnogene Diagnósticos Moleculares LTDA – Brasília - DF

*Função:* Diagnósticos Moleculares

**01/2007 - 12/2008** – Estágio Remunerado/Estagiário Graduado

Universidade de Brasília - UnB

*Vínculo:* Laboratório de Virologia Molecular

*Orientador:* Dr. Enrique Roberto Argañaraz

*Projeto:* Identificação de Mutações em Proteínas Presentes em HIV.

**01/2009 - 12/2010** – Laboratório Tecnogene Diagnósticos Moleculares LTDA – Brasília - DF

*Vínculo:* Celetista - Biólogo

*Função:* Diagnósticos Moleculares

**01/2011 - 01/2012** – Estágio Remunerado/Estagiário Graduado – FAP-DF

EMBRAPA - CENARGEN

*Vínculo:* Laboratório de Interação Molecular Planta-Patógeno

*Orientadora:* Dra. Maria Fátima Grossi de Sá

*Projeto:* Estratégias Moleculares Aplicadas a Prospecção de Genes Visando o Controle de Insetos-Praga.

#### 4. PRODUÇÃO BIBLIOGRÁFICA

1. **Neiva M, Arraes FBM, De Souza JV, Radis-Baptista G, Prieto da Silva ARB, Walter MEMT, Brigido MdM, Yamane T, Lopez-Lozano JL, Astolfi-Filho S** (2009) Transcriptome analysis of the Amazonian viper *Bothrops atrox* venom gland using expressed sequence tags (ESTs). *Toxicon* **53**: 427-436.
2. **Nicola A, Andrade R, Dantas A, Andrade P, Arraes FBM, Fernandes L, Silva-Pereira I, Felipe M** (2008) The stress responsive and morphologically regulated *hsp90* gene from *Paracoccidioides brasiliensis* is essential to cell viability. *BMC Microbiology* **8**: 158.
3. **Arraes FBM, Carvalho MJA, Maranhao AQ, Brigido MM, Pedrosa FO, Felipe MSS** (2007) Differential metabolism of Mycoplasma species as revealed by their genomes. *Genetics and Molecular Biology* **30**: 182-189.
4. **Felipe MSS, Torres FAG, Maranhao AQ, Silva-Pereira I, Poças-Fonseca MJ, Campos EG, Moraes LMP, Arraes FBM, Carvalho MJA, Andrade RV, Nicola AM, Teixeira MM, Jesuíno RSA, Pereira M, Soares CMA, Brígido MM** (2005) Functional genome of the human pathogenic fungus *Paracoccidioides brasiliensis*. *FEMS Immunology & Medical Microbiology* **45**: 369-381.
5. **Arraes FBM, Benoliel B, Burtet RT, Costa PL, Galdino AS, Lima LH, Marinho-Silva C, Oliveira-Pereira L, Pfrimer P, Procopio-Silva L, Reis VC, Felipe MS** (2005) General metabolism of the dimorphic and pathogenic fungus *Paracoccidioides brasiliensis*. *Genetics and Molecular Research* **4**: 290-308.
6. **Albuquerque P, Baptista AJ, Derengowsky Lda S, Procopio L, Nicola AM, Arraes FBM, Souza DP, Kyaw CM, Silva-Pereira I** (2005) *Paracoccidioides brasiliensis* RNA biogenesis apparatus revealed by functional genome analysis. *Genetics and Molecular Research* **4**: 251-272.
7. **Benoliel B, Arraes FBM, Reis VC, Siqueira SJ, Parachin NS, Torres FA** (2005) Hydrolytic enzymes in *Paracoccidioides brasiliensis* - ecological aspects. *Genetics and Molecular Research* **4**: 450-461.

8. Felipe MSS, Andrade RV, Arraes FBM, Nicola AM, Maranhao AQ, Torres FAG, Silva-Pereira I, Poças-Fonseca MJ, Campos EG, Moraes LMP, Andrade PA, Tavares AHFP, Silva SS, Kyaw CM, Souza DP, Network P, Pereira M, Jesuino RSA, Andrade EV, Parente JA, Oliveira GS, Barbosa MS, Martins NF, Fachin AL, Cardoso RS, Passos GAS, Almeida NF, Walter MEMT, Soares CMA, Carvalho MJA, Brígido MM (2005) Transcriptional profiles of the human pathogenic fungus *Paracoccidioides brasiliensis* in mycelium and yeast cells. Journal of Biological Chemistry **280**: 24706-24714.
9. Felipe MSS, Andrade RV, Petrofeza SS, Maranhão AQ, Torres FAG, Albuquerque P, Arraes FBM, Arruda M, Azevedo MO, Baptista AJ, Bataus LAM, Borges CL, Campos EG, Cruz MR, Daher BS, Dantas A, Ferreira MASV, Ghil GV, Jesuino RSA, Kyaw CM, Leitão L, Martins CR, Moraes LMP, Neves EO, Nicola AM, Alves ES, Parente JA, Pereira M, Poças-Fonseca MJ, Resende R, Ribeiro BM, Saldanha RR, Santos SC, Silva-Pereira I, Silva MAS, Silveira E, Simoes IC, Soares RBA, Souza DP, De-Souza MT, Andrade EV, Xavier MAS, Veiga HP, Venancio EJ, Carvalho MJA, Oliveira AG, Inoue MK, Almeida NF, Walter MEMT, Soares CMA, Brígido MM (2003) Transcriptome characterization of the dimorphic and pathogenic fungus *Paracoccidioides brasiliensis* by EST analysis. Yeast **20**: 263-271.

## 5. TRABALHOS PUBLICADOS EM ANAIS DE CONGRESSOS

1. Grossi-de-Sá MF, Arraes FBM, Amorim RM, Beneventi MA, Albuquerque EVS, Oliveira RS, Nepomuceno A, Abdelnoor R, Sá MEL. Soybean ethylene biosynthesis genes potentially involved in drought tolerance. *In*: Plant and Animal Genome XXI, San Diego - Livro de Resumos, 2013.
2. Arraes FBM, Sá MEL, Amorim RM, Beneventi MA, Albuquerque EVS, Oliveira RS, Nepomuceno A, Abdelnoor R, Grossi-de-Sá MF. *In silico* analysis of ACC Synthase (ACS) genes identified in Brazil Soybean

Genome Consortium Database. *In*: IX World Soybean Research Conference, África do Sul - Livro de Resumos, 2013.

3. **Arraes FBM, Sá MEL, Amorim RM, Beneventi MA, Albuquerque EVS, Oliveira RS, Nepomuceno A, Abdelnoor R, Grossi-de-Sá MF.** *In silico* analysis of ACC Synthase (ACS) genes identified in Brazil Soybean Genome Consortium Database. *In*: 4º Congresso Brasileiro de Biotecnologia, Guarujá - Livro de Resumos, 2012.
4. **Arraes FBM, Pereira M, Andrade EV, Jesuino RSA, Simões IC, Teixeira MM, Bailão AM, Silva G, Parente J, Camargo J, Barbosa MS, Castro NS, Soares CMA, Brígido MM, Felipe MSS.** Metabolic analysis of the transcriptome from the fungal pathogen *Paracoccidioides brasiliensis*. *In*: XXIV Reunião de Genética de Microrganismos, Gramado - Livro de Resumos, 2004.
5. **Arraes FBM, Nicola AM, Reis VCB, Neves E, Derengowski LS, Pereira IS, Fonseca MJP, Felipe MSS.** Identificação, isolamento e caracterização do gene do repressor transcricional *Pbtup1* do fungo dimórfico e patogênico *Paracoccidioides brasiliensis*. *In*: XXIV Reunião de Genética de Microrganismos, Gramado - Livro de Resumos, 2004.
6. **Derengowski LS, Albuquerque P, Baptista AJ, Machado LS, Nicola AM, Arraes FBM, Neves EO, Fonseca MJP, Felipe MSS, Pereira IS.** Isolamento e caracterização de sequências promotoras de genes expressos diferencialmente no fungo dimórfico *Paracoccidioides brasiliensis*. *In*: XXIV Reunião de Genética de Microrganismos, Gramado - Livro de Resumos, 2004.
7. **Felipe MSS, Andrade RV, Petrofeza SS, Carvalho MJA, Albuquerque P, Alves ES, Arraes FBM, Arruda M, Azevedo MO, Baptista AJ, Bataus LA, Borges C, Campos EG, Daher BS, Dantas A, Ferreira M, Ghil GV, Inove MK, Kyaw C, Leitão L, Maranhão AQ, Martins CR, Moraes LMP, Nicola AM.** Initial transcriptome characterization of the dimorphic and pathogenic fungus *Paracoccidioides brasiliensis*. *In*: XXXII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, 2003, Caxambu - Livro de Resumos, 2003.

8. Nicola AM, Arraes FBM, Dantas A, Pereira IS, Felipe MSS. The HSP90 gene in *Paracoccidioides brasiliensis*. In: XXII Congresso Brasileiro de Microbiologia, Florianópolis - Livro de Resumos, 2003.
9. Arraes FBM, Andrade RV, Petrofeza SS, Azevedo MO, Baptista AJ, Maranhão AQ, Moraes LMP, Nicola AM, Saldanha R, Pereira IS, Torres FAG, Andrade EV, Xavier MAS, Walter MEMT, Soares CMA, Brígido MM, Felipe MSS. Ubiquitin genes in the partial transcriptome characterization of *Paracoccidioides brasiliensis*. In: VIII Encontro Internacional Sobre Paracoccidioidomicose, 2002, Pirenópolis - GO. ARBS - Annual Review of Biomedical Sciences - Special Issue, 2002.
10. Andrade RV, Petrofeza SS, Albuquerque P, Alves ES, Arraes FBM, Arruda M, Azevedo MO, Baptista AJ, Bataus LA, Borges C, Campos EG, Daher BS, Dantas A, Ferreira M, Ghil GV, Inove MK, Kyaw C, Leitão L, Maranhão AQ, Martins CR, Moraes LMP, Nicola AM, Oliveira AG, Parente J, Pereira M. Partial transcriptome characterization of the dimorphic and pathogenic fungus *Paracoccidioides brasiliensis*. In: VIII Encontro Internacional Sobre Paracoccidioidomicose, 2002, Pirenópolis - GO. ARBS - Annual Review of Biomedical Sciences - Special Issue, 2002.

## 6. PRÊMIOS E DISTINÇÕES

- 2004 – XXIV Reunião de Genética de Microrganismos - Gramado/RS – **Prêmio dos 10 Melhores Trabalhos** – Identificação, Isolamento e Caracterização do Gene do Repressor Transcricional *Pbtup1* do Fungo Dimórfico e Patogênico *Paracoccidioides brasiliensis*.