



UNIVERSIDADE
ESTADUAL DE LONDRINA

SILVIA GRACIELE HÜLSE DE SOUZA

**FATORES DE TRANSCRIÇÃO DA FAMÍLIA ERF EM *Coffea*
arabica: IDENTIFICAÇÃO E AVALIAÇÃO
TRANSCRICIONAL EM ESTRESSES ABIÓTICOS**

Londrina
2011

SILVIA GRACIELE HÜLSE DE SOUZA

**FATORES DE TRANSCRIÇÃO DA FAMÍLIA ERF EM *Coffea*
arabica: IDENTIFICAÇÃO E AVALIAÇÃO
TRANSCRICIONAL EM ESTRESSES ABIÓTICOS**

Tese apresentada ao Programa de Pós-Graduação em Agronomia, da Universidade Estadual de Londrina.

Orientadora: Profa. Dra Valéria Carpentieri-Pípolo

Co-orientador: Dr. Luiz Filipe Protasio Pereira

Londrina
2011

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

Dados Internacionais de Catalogação-na-Publicação (CIP)

S729f Souza, Silvia Graciele Hülse de.

Fatores de transcrição da família ERF em *Coffea arabica*: identificação e avaliação transcricional em estresses abióticos / Silvia Graciele Hülse de Souza. – Londrina, 2011. 174 f. : il.

Orientador: Valéria Carpentieri-Pipolo.

Co-orientador: Luiz Filipe Protasio Pereira.

Tese (Doutorado em Agronomia) – Universidade Estadual de Londrina, Centro de Ciências Agrárias, Programa de Pós-Graduação em Agronomia, 2011. Inclui bibliografia.

1. Café – Melhoramento genético – Teses. 2. Genética – Expressão – Teses. 4. Plantas – Genética molecular – Teses. I. Carpentieri-Pipolo, Valéria. II. Pereira, Luiz Filipe Protasio. III. Universidade Estadual de Londrina. Centro de Ciências Agrárias. Programa de Pós-Graduação em Agronomia. IV. Título.

CDU 631.52:633.73

SILVIA GRACIELE HÜLSE DE SOUZA

**FATORES DE TRANSCRIÇÃO DA FAMÍLIA ERF EM *Coffea arabica*:
IDENTIFICAÇÃO E AVALIAÇÃO TRANSCRICIONAL EM ESTRESSES
ABIÓTICOS**

Tese apresentada ao Programa de Pós-Graduação em Agronomia, da Universidade Estadual de Londrina.

BANCA EXAMINADORA

Dr. Douglas Silva Domingues
IAPAR – Londrina - Pr

Prof. Dr. Laurival A. Vilas Bôas
UEL – Londrina - Pr

Dr. Luiz Gonzaga Esteves Vieira
IAPAR – Londrina - Pr

Prof. Dr. Ricardo de Tadeu Faria
UEL – Londrina - Pr

Profa. Dra Inês C. B. Fonseca
(Suplente)
UEL – Londrina - Pr

Prof. Dr. Cássio Egídeo C. Prete
UEL (Suplente)

Profa. Dra. Valéria Carpentieri-Pípolo
Orientadora
Universidade Estadual de Londrina

Londrina, 30 de junho de 2011.

*À minha mãe Terezinha,
pelo exemplo de vida, dedicação e amor incondicional.
Por muitas vezes ter abdicado mão de seus sonhos pela realização dos meus,
Com amor...*

*À minha orientadora, co-orientadores e as pessoas
que tornaram possível a realização desse trabalho,
Com carinho...*

AGRADECIMENTOS

À Deus, pela vida, por servir de apoio e consolo nos momentos de dificuldade, por me proporcionar conhecer pessoas empenhadas em contribuir para a Ciência, e pelas amizades verdadeiras construídas ao longo do doutorado que levarei por toda vida.

À minha orientadora, Prof. Dra. Valéria Carpentieri-Pípolo, pela orientação, confiança e oportunidades oferecidas. Obrigada pela dedicação, incentivo e amizade durante quase 10 anos de convívio.

Ao meu co-orientador, Dr. Luiz Filipe Protasio Pereira, pela possibilidade de realizar o trabalho no Laboratório de Biotecnologia do IAPAR, pela disponibilidade e orientação, pelo apoio dado durante toda a realização desse trabalho, pelos ensinamentos e amizade.

Ao meu co-orientador no exterior, Dr. Mondher Bouzayen, pela possibilidade de realizar parte desse trabalho no Laboratoire de Génomique et Biotechnologie des Fruits, pela orientação e confiança em mim depositada.

À Universidade Estadual de Londrina, em especial aos professores do Programa de Pós-Graduação em Agronomia que tanto contribuíram para a minha formação acadêmica e que contribuem para a qualidade do programa.

À Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Pesquisa e Desenvolvimento (CNPq), pela concessão das bolsas de estudos no exterior e Brasil, respectivamente.

Às queridas secretárias Weda (UEL) e Brigitte (GBF) pela amizade e por sempre resolverem todas as questões burocráticas.

À Prof. Dra. Anne Bernadac e PhD Julien Pirrelo pela disponibilidade e valiosas orientações, pelos ensinamentos, apoio e confiança em mim depositada.

Ao Dr. Douglas Silva Domingues por toda disponibilidade em tirar minhas dúvidas, por me ajudar a buscar as melhores saídas, pela amizade e por prontamente aceitar fazer parte da banca examinadora deste trabalho

Ao Dr. Getúlio Takahashi pelo auxílio na condução nas análises fisiológicas e pelo comprometimento com o trabalho.

Ao amigo Dr. Deoclécio Domingos Garbúglio pela amizade e auxílio nas análises estatísticas.

Ao Dr. Prof. Dr. Ricardo de Tadeu Faria, Prof. Dr Laurival A. Vilas Boas, Prof. Dr. Cássio Egídeo C. Prete, Profa. Dra Inês C. B. Fonseca por tão prontamente aceitarem fazer parte da banca examinadora deste trabalho.

Ao Tiago, Sandra e Viviane por toda ajuda prestada, incentivo, dúvidas, risadas e pela amizade verdadeira.

Aos colegas do "LBI" Gislaine, Suzana, Kênia, Priscila, Júlia, Carol, Lucas, Karina, Renata, Eduardo e aos "companheiros do puxadinho" Léo, Giselly e Rafa por toda ajuda, pelos ótimos momentos vividos juntos e pelas amizades verdadeiras que levarei por toda vida.

Aos colegas do "GBF" Maha, Hanane, Emna, Donia, Li Ji, Winping, Xian, Laure, Isabele Mila, Simone, SiMo, Eduardo Purgartto, Cesar Rombaldi, Cristina, Isa Egea e Charo pela amizade e por tornarem o ambiente de trabalho agradável.

Aos colegas de curso de Pós-Graduação pela amizade e companheirismo e pelo muito que pude aprender do contato com seus projetos e experiências profissionais, tão diversas da minha.

Ao querido Carneiro, por todo amor, carinho e companheirismo, por acreditar em mim mais do que eu mesma acredito, por toda ajuda e principalmente

por estar ao meu lado sempre, mesmo quando estávamos a um oceano de distância.

À minha mãe, por ter me encorajado em todos os momentos e por ter se mantido ao meu lado em todas as minhas escolhas, por sempre confiar que tudo ia dar certo e por me ajudar a acreditar nisso.

Aos meus pais, no qual o doutorado não teria nenhum valor para mim se eu não pudesse ver a satisfação de vocês com essa conquista.

E, por fim a todos que de alguma forma colaboraram para o desenvolvimento deste trabalho e para a minha formação.

*"A mente que se abre a uma nova idéia
jamais voltará ao seu tamanho original".*

(Albert Einstein)

SOUZA, S.G.H. **Fatores de transcrição da família ERF em *Coffea arabica*: Identificação e avaliação transcricional em estresses abióticos.** 2011. 174 f. Tese (Doutorado em Agronomia) – Universidade Estadual de Londrina, Londrina, 2011.

RESUMO

Os fatores de transcrição *ethylene response factor* (ERF) desempenham importante papel no desenvolvimento e na expressão dos genes que regulam a resposta de defesa em plantas. Sendo assim, o objetivo desse estudo foi identificar os fatores de transcrição ERFs presentes no banco de dados de ESTs em *Coffea arabica* e avaliar a atividade transcricional de 13 ERFs no cafeeiro, sob diversas condições de estresses abióticos. Para isso, foram realizadas buscas dentro do banco de dados do Projeto Genoma Café para a identificação e caracterização dos fatores de transcrição ERFs. Foram identificados 36 ERFs em *Coffea arabica* com o domínio completo AP2/ERF. A identificação dos domínios AP2/ERF foram conservados entre *C. arabica* e *Arabidopsis*. A análise *in silico* do perfil de expressão gênica mostrou que níveis elevados de expressão foram observados em bibliotecas derivadas de tecidos de frutos, folhas, flores e em bibliotecas submetidas à estresse hídrico. A partir da análise *in silico*, 13 ERFs pertencentes a oito grupos da família ERF foram selecionados e as análises de expressão sob estresses hídrico, salino e térmico foram realizadas. No que se refere ao estresse hídrico foi observado que, através da análise do perfil transcricional dentro da subfamília DREB, *CaERF01* teve um aumento transcricional acentuado em plantas submetidas a déficit hídrico severo, enquanto que atividade transcricional de *CaERF09* decresceu durante o mesmo estresse. Dentro da subfamília ERF, em todos os *CaERFs*, o máximo acúmulo de transcritos ocorreu no déficit hídrico severo. Foi também possível observar que os grupos VIII e IX apresentaram padrões distintos de expressão no estresse hídrico. No estresse salino, dentre os membros da subfamília DREB, o máximo acúmulo de transcritos ocorreu seis dias após o início do estresse para *CaERF01* e *CaERF09*. Dentre os membros da subfamília ERF, os genes *CaERF30* e *CaERF22* apresentaram o máximo acúmulo de transcritos 2 e 4 dias após o início do tratamento, respectivamente. Em *CaERF19* e *CaERF29*, foi observado um aumento da atividade transcricional no sexto dia de estresse. Entretanto, para os genes *CaERF27* e *CaERF31* o padrão de transcrição se manteve alto em plantas com 12 dias de estresse salino. Durante o estresse térmico, dentro da subfamília DREB, nos genes *CaERF01* e *CaERF06* o acúmulo de transcritos se deu de forma gradual no início do estresse, atingindo o máximo cinco dias após o início tratamento. Dentro os membros da família ERF, os genes *CaERF30* e *CaERF31* apresentaram acentuado acúmulo no quinto dia de tratamento. A expressão do gene *CaERF19* também foi analisada por q-PCR durante os estresses térmico, salino e hídrico, onde foi observada a indução da expressão nos três estresses. A indução da expressão da família ERF, em resposta a mais de um estresse, sugere que existe uma resposta cruzada (crosstalk) entre as diferentes vias de sinalização dos diferentes estresses. Os dados gerados neste trabalho fornecem subsídios para selecionar o melhores candidatos para futuras análises funcionais da família ERF em *C. arabica*, os quais ajudarão a compreender os determinantes genéticos da tolerância aos estresses abióticos, o que constitui um passo importante nos programas de melhoramento genético do cafeeiro.

Palavras-chave: AP2/ERF. Café. Expressão gênica. Tolerância aos estresses.

SOUZA, S.G.H. **Transcription factors of the ERF family in *Coffea arabica*: identification and expression patterns under abiotic stresses.** 2011.174 f. Tese (Doutorado em Agronomia) – Universidade Estadual de Londrina, Londrina, 2011.

ABSTRACT

ERF (ethylene response factors) transcription factors play an important role in plant development and on the gene expression that regulate the plant defense response. Therefore, the objective of this study was to identify the *Coffea arabica* ERF transcription factors and evaluate the transcriptional activity of 13 ERFs coffee trees, under three abiotic stresses conditions. Searches were performed within the data bank of the Brazilian Coffee Genome Project for the identification and characterization of ERFs transcription factors. In this study, 36 ERFs that contained a complete AP2/ERF domain were identified in *C. arabica*. The identification of the AP2/ERF conserved domains was conserved between *C. arabica* and *Arabidopsis*. The *in silico* analysis of gene expression profile have shown that high levels of expression were observed in fruits, leaves and flowers and on libraries subjected to waters stresses. Starting from the *in silico* analysis, 13 ERFs belonging to eight different groups of the ERF family were selected to transcriptional evaluation under drought, saline, and heat stresses. Concerning the drought stress it was observed that when plants were subjected to severe hydric deficit the gene *CaERF01* had an intense transcriptional increase while the transcriptional activity of the gene *CaERF09* had a decrease during that same stress. Within the DREB subfamily, in all CaERFs, the maximum accumulation of transcripts occurred during severe hydric deficit. It was also possible to observe that the groups VIII and IX presented distinct expression patterns under water stress. Under saline stress, among the members of the DREB subfamily, the maximum accumulation of transcripts occurred six days after the beginning of the stress for the genes *CaERF01* and *CaERF09*. Among the members of the ERF family, the genes *CaERF30* and *CaERF22* presented the maximum accumulation of transcripts two and four days after the beginning of treatment, respectively. For the *CaERF19* and *CaERF29* genes, the increase on the transcriptional activity was observed at the sixth day of the stress. For the *CaERF27* and *CaERF31*, however, the transcription pattern remained high in plants with 12 days under saline stress. During heat stress, DREB subfamily, the upregulation occurred gradually at the beginning of the stress for the genes *CaERF01* and *CaERF06*, reaching its maximum five days after starting the treatment. Within the members of the family ERF, *CaERF30* and *CaERF31* presented marked accumulation of transcripts at the fifth day after treatment. The expression of the gene *CaERF19* was also analyzed by q-PCR during the water, saline and heat stresses, and the induction of expression was observed under the three stresses. The upregulation transcriptional of the ERF family, in response to more than one stress, suggests that there is a crosstalk response among distinct signalization pathways of the different stresses. This work provide support in selecting candidates for future functional analyses in the *C. arabica* ERF family and will help in the understanding of the genetic determinants of the tolerance to abiotic stresses, which constitutes an important step in coffee breeding programs.

Key – words: AP2/ERF. Coffee. Genic expression. Stresses tolerance.

LISTA DE FIGURAS

- Figura 1** – Genes induzidos durante o estresse hídrico e ao frio e suas possíveis funções na tolerância e resposta ao estresse (adaptado de SHINOZAKI;YAMAGUCHI-SHINOZAKI, 2006).....24
- Figura 2** – Vias de transdução de sinal via ABA dependente e independente. ABRE é o principal elemento em resposta a ABA e os fatores de transcrição desta via são os AREB/ABFs. MYB2 e MYC2 e NAC também são induzidos por ABA, sensíveis ao JA e podem estar relacionados à resposta a estresses bióticos. DRE é o principal elemento das vias independentes de ABA, os fatores DREB1/CBFs estão envolvidos na resposta ao frio, DREB2s na resposta a desidratação e alta salinidade e NAC e HD-ZIP à seca e salinidade, mas não ao frio. (SHINOZAKI;YAMAGUCHI-SHINOZAKI, 2007).....25
- Figura 3** – Organização estrutural da superfamília AP2 (adaptado de SHARONI et al, 2011).....27
- Figura 4** – Esquema das funções fisiológicas dos ERFs em resposta aos diferentes estresses (SINGH et al, 2002).....29

4 Artigo A

- Figura 1** – Árvore filogenética dos ERFs de *C. arabica* e *arabidopsis*. As sequências de aminoácidos do domínio AP2 de 36 genes ERF de *C. arabica* e *arabidopsis* foram alinhadas pelo Clustal W e a árvore filogenética foi construída com MEGA 4.0 utilizando o método NJ. (Subfamília DREB: ○ - grupo I, □ - grupo II, ▽ - grupo III, ◇ - grupo IV; Subfamília ERF: ▼ - grupo VI, ● - grupo VII, ■ - grupo 8, ▲ - grupo IX, ◆ - grupo X).....51
- Figura 2** – Alinhamento dos domínios AP2 a partir das proteínas ERF de *C. arabica*. As colunas de cores, azul escuro e azul claro, indicam resíduos de aminoácidos idênticos e conservados, respectivamente. A barra preta e as setas representam a α -hélice e as regiões β -sheets, respectivamente. *CaERF01* a 12

	pertencem a subfamília DREB; <i>CaERF13</i> a 36 pertencem a subfamília ERF	54
Figura 3	– Distribuição dos motivos conservados na família ERF em <i>Coffea arabica</i> . Os Motivos foram identificados em <i>C. arabica</i> e classificados de acordo com a classificação proposta por NAKANO et al, (2006). A posição e a ordem dos símbolos correspondem a real região conservada na seqüência da proteína	57
Figura 4	– Perfil de expressão <i>in silico</i> dos ERFs em <i>Coffea arabica</i> . O número de <i>reads</i> foi normalizado para cada biblioteca, e os valores representados pela escala de expressão relativa	59
5 Artigo B		
Figura 1	– Alinhamento dos domínios AP2/ERF a partir das proteínas ERF de <i>C. arabica</i> . As colunas de cores, azul escuro e azul claro, indicam resíduos de aminoácidos idênticos e conservados, respectivamente	86
Figura 2	– Estrutura das proteínas <i>CaERFs</i> (<i>CaERF01</i> , <i>CaERF04</i> , <i>CaERF06</i> , <i>CaERF09</i> , <i>CaERF13</i> , <i>CaERF19</i> , <i>CaERF22</i> , <i>CaERF23</i> , <i>CaERF27</i> , <i>CaERF29</i> , <i>CaERF30</i> , <i>CaERF31</i> e <i>CaERF36</i>). As regiões em preto e cinza representam o domínio AP2/ERF e o domínio ácido, respectivamente. Os putativos sinais de localização nuclear (NLS) estão marcados com as setas. A região em cinza claro e 'M' indica o motivo MCGGAIL/L na extremidade N-terminal, enquanto que a região hachurada e 'EAR' corresponde ao motivo conservado repressor EAR (ERF-associado repressão anfífilicos) na extremidade C-terminal. Os números abaixo das caixas indicam as posições dos resíduos de aminoácidos	87
Figura 3	– Análise filogenética entre os ERFs de <i>Coffea arabica</i> e dos organismos que apresentaram maior similaridade em <i>Arabidopsis thaliana</i> AtERF059 (AT1G78080), AtERF010 (AT5G67190), AtERF034 (AT2G44940), AtERF066	

(AT3G14230),	AtERF075	(AT3G14230),	AtERF078
(AT3G15210),	AtERF100	(AT4G17500),	AtERF112
(AT2G33710),	<i>Vitis vinifera</i> VvERF004	(GSVIVP00015253001),	
VvERF007	(GSVIVP00019541001),	VvERF029	
(GIDVvP00033798001),	VvERF054	(GSVIVP00006483001),	
VvERF059	(GSVIVP00034010001),	VvERF064	
(GSVIVP00003062001),	VvERF077	(GSVIVP00014229001),	
VvERF113	(GSVIVP00019350001),	<i>Solanum lycopersicum</i>	
SIERF6, SIERF13, SIERF21, SIERF31, SIERF76, SIERF35,			
SIERF70, SIERF83 e <i>Populus trichocarpa</i> PtDREB19			
(XP_002301042.1), PtDREB34 (XP_002307277.1), PtDREB82			
(XP_002334238.1), PtERF82 (XP_002319978.1), PtERF38			
(XP_002331217.1), PtERF48 (XP_002305121.1), PtERF13			
(XP_002304592.1), PtERF76 (XP_002306712.1).....			90

Figura 4A – Expressão relativa dos genes ERFs de *Coffea arabica* submetidos ao estresse hídrico (C - controle, EM - estresse moderado, ES-estresse severo e R - recuperado). Os níveis de expressão foram determinados por RT-PCR semi-quantitativa usando oligonucleotídeos específicos e normalizadas com o gene GAPDH. As imagens fotográficas da eletroforese em gel são mostradas na parte inferior de cada gráfico. Entre parênteses está o número de ciclos. Os dados representam as médias \pm desvio padrão de três repetições biológicas. Médias seguidas de mesma letra não diferem entre si por Tukey 5%..... 95

Figura 4B – Expressão relativa dos genes ERFs de *Coffea arabica* submetidos ao estresse salino (C-controle, D1- dia 1, D2- dia 2, D4-dia 4, D6 - dia 6 e D12 - dia 12). Os níveis de expressão foram determinados por RT-PCR semi-quantitativa usando oligonucleotídeos específicos e normalizadas com o gene GAPDH. As imagens fotográficas da eletroforese em gel são mostradas na parte inferior de cada gráfico. Entre parênteses está o número de ciclos. Os dados representam as médias \pm desvio padrão de três repetições biológicas. Médias seguidas de mesma letra não diferem entre si por Tukey 5%..... 96

- Figura 4C** – Expressão relativa dos genes ERFs de *Coffea arabica* submetidos ao estresse térmico (C-controle, D1- dia 1, D3- dia 3 e D5-dia 5). Os níveis de expressão foram determinados por RT-PCR semi-quantitativa usando oligonucleotídeos específicos e normalizadas com o gene GAPDH. As imagens fotográficas da eletroforese em gel são mostradas na parte inferior de cada gráfico. Entre parênteses está o número de ciclos. Os dados representam as médias \pm desvio padrão de três repetições biológicas. Médias seguidas de mesma letra não diferem entre si por Tukey 5%97
- Figura 5** – Padrão de expressão exibido pela família ERF em *Coffea arabica* durante o estresse hídrico. Foram observados 5 padrões de expressão. O eixo x representa: C-controle, EM-estresse moderado, ES-estresse severo e R-recuperado e nas caixas em preto está o número de cada CaERF.....97
- Figura 6** - A análise da expressão de *CaERF19* em resposta aos tratamentos de estresses abióticos por quantitativa RT-PCR. Figura 6A: Estresse hídrico (C-controle, EM-estresse moderado, ES-estresse severo e R-recuperado). Figura 6B: Estresse térmico (C-controle, D1 - dia 1, D3 - dia 3 e D5 - dia 5) e Figura 6C: Estresse salino (C-controle, D1 - dia 1, D2 - dia 2, D4 - dia 4, D6 - dia 6 e D12 - dia 12). Os dados representam as médias \pm desvio padrão de três repetições biológicas. Médias seguidas de mesma letra não diferem entre si por Tukey 5%.....99

LISTA DE TABELAS

4 Artigo A

- Tabela 1** – Sequências de *Coffea arabica* com homologia com os genes da família ERF em *Arabidopsis thaliana*50
- Tabela 2** – Número de genes presentes em cada grupo da família ERF em *C. arabica* e nas espécies cujos genomas foram completamente sequenciados, e tamanho dos respectivos genomas.....52

5 Artigo B

- Tabela 1** – Sequências dos oligonucleotídeos de *CaERF* utilizados para a semi-quantitativa RT-PCR.....82
- Tabela 2** – Comparação entre as seqüências completas preditas de aminoácidos dos fatores de resposta ao etileno (*CaERF*) cDNAs. Acima da diagonal estão os valores da identidade enquanto que abaixo da diagonal estão os valores da similaridade entre os ERFs de *Coffea arábica*85
- Tabela 3** - Organismos que apresentaram maior similaridade em *Arabidopsis thaliana*, *Vitis vinifera*, *Solanum lycopersicum* e *Populus trichocarpa* genes ERFs de *Coffea arábica*.....91

SUMÁRIO

1 INTRODUÇÃO	19
2 REVISÃO DE LITERATURA	21
2.1 IMPORTÂNCIA ECONÔMICA DA CAFEICULTURA	21
2.2 REPOSTA AOS ESTRESSES ABIÓTICOS	22
2.3. FATORES DE TRANSCRIÇÃO	26
2.3.1 Superfamília AP2/ERF	26
2.3.2 Diferenças e Classificação da Família ERF	27
2.3.2.1 Subfamília DREB	29
2.3.2.2 Subfamília ERF	30
3 REFERÊNCIAS	34
4 ARTIGO A – ANÁLISE <i>IN SILICO</i> DA FAMÍLIA ERF EM <i>Coffea arabica</i>: IDENTIFICAÇÃO E PADRÃO DE EXPRESSÃO	43
4.1 RESUMO	44
4.2 ABSTRACT	44
4.3 INTRODUÇÃO	45
4.4 MATERIAL E MÉTODOS	47
4.5 RESULTADOS	49
4.6 DISCUSSÃO	60
4.7 AGRADECIMENTOS	66
4.8 REFERÊNCIAS	66
5 ARTIGO B – ANÁLISE DA EXPRESSÃO GÊNICA DE TREZE ERFS DE <i>Coffea arabica</i> SUBMETIDOS A DIFERENTES ESTRESSES ABIÓTICOS	74
5.1 RESUMO	75
5.2 ABSTRACT	75
5.3 INTRODUÇÃO	76
5.4 MATERIAL E MÉTODOS	78
5.5 RESULTADOS	83

5.6 DISCUSSÃO	100
5.7 AGRADECIMENTOS	106
5.8 REFERÊNCIAS	107
6 CONCLUSÕES GERAIS	114
APÊNDICES	115
APÊNDICE A – Alinhamento dos domínios AP2/ERF de <i>Coffea arabica</i> e <i>Arabidopsis thaliana</i>	116
APÊNDICE B – Motivos conservados da família ERF de <i>Coffea arabica</i> e <i>Arabidopsis thaliana</i>	159
APÊNDICE C – Alinhamento entre as sequencias de aminoácidos de <i>CaERF17</i> de <i>Coffea arabica</i> e <i>CoERF</i> de <i>Coffea canephora</i>	172
APÊNDICE D – Dados fisiológicos referentes aos estresses hídrico e térmico	173
APÊNDICE E – Curvas de dissociação dos genes <i>RPL39</i> e <i>CaERF19</i>	174

1 INTRODUÇÃO

Com uma produção mundial de 134 milhões de sacas de grãos em 2010 (ICO, 2011), o café é um importante produto agrícola cultivado em mais de 80 países. O Brasil é o maior produtor mundial e junto com a Colômbia e o Vietnã são responsáveis por mais de 50% da produção mundial (MAPA, 2011). O *Coffea arabica*, um dos maiores commodities agrícolas do mundo, representa uma importante fonte de renda para os países tropicais em desenvolvimento (PAY, 2009). No gênero *Coffea*, duas espécies são responsáveis pela quase totalidade da produção de grãos de café: *C. canephora* e *C. arabica*, que contribuem aproximadamente com 30 e 70% da produção mundial, respectivamente. A produção de café está sujeita a oscilações regulares que é pode ser explicado, principalmente, pelo ciclo bienal da cultura, mas também pelos efeitos adversos de condições climáticas. Entre eles, a seca e as altas temperaturas são fatores que afetam o desenvolvimento e produção do cafeeiro (ASSAD et al, 2004; DaMATTA, 2004; DaMATTA, 2006).

O hormônio etileno é importante em vários processos fisiológicos que ocorrem nas plantas, sendo indutor de respostas relacionadas aos diferentes estresses. Como a taxa de transcrição de um gene é importante determinante de sua expressão, os mecanismos moleculares pelos quais a transcrição gênica é regulada têm grande interesse crescente e se tornado um dos tópicos principais de estudos em genética molecular envolvendo o etileno. A cascata transcriconal que ele regula termina pela ativação dos fatores de transcrição específicos que controlam as diferentes respostas relacionadas a este hormônio (SOLANO et al, 1998).

Ethylene Response Factors (ERF) são, portanto, os últimos elementos da cascata transcriconal que constituem uma grande família multigênica. Os fatores de transcrição pertencentes a família ERF fazem parte da superfamília AP2/ERF. Assim como as famílias AP2 e RAV, são caracterizadas pela presença do domínio AP2, caracterizado por 60-70 aminoácidos implicados na ligação ao DNA (RIECHMANN et al, 2000). As proteínas da família ERF contém um único domínio AP2, enquanto que as proteínas da família AP2 contém dois domínios repetidos AP2 (Jofuku et al, 1994). As proteínas da família RAV, além do domínio AP2, contém um domínio B3, que é um domínio de ligação ao DNA (KAGAYA et al, 1999; NAKANO et al, 2006). A família ERF por sua vez, apresenta uma subdivisão em duas

subfamílias, ERF e CBF/DREB (JOFUKU et al, 1994; SAKUMA et al, 2002; NAKANO et al, 2006). De modo geral, os genes da subfamília ERF estão envolvidos principalmente em respostas a estresses bióticos, reconhecendo o elemento *cis-acting* AGCCGCC denominado GCC-box (HAO et al, 1998). Genes da subfamília CBF/DREB desempenham um papel crucial na resposta das plantas a estresses abióticos, reconhecendo o elemento responsivo a desidratação (*Dehydration Responsive Element* - DRE) com um motivo central A/GCCGAC (SHINOZAKI; SHINOZAKI,1994; THOMASHOW, 1999).

Fatores de transcrição das subfamílias ERFs e CBF/DREB foram identificados em várias espécies vegetais, incluindo *Arabidopsis* (LIU et al, 1998; SAKUMA et al, 2002; NAKANO et al, 2006), arroz (NAKANO et al, 2006; Sharone et al, 2011), algodão (CHAMPION et al, 2009), soja (ZHANG et al, 2008), álamo (ZHUANG et al, 2008), uva (ZHUANG et al, 2009; LICAUSI et al, 2010), milho (ZHUANG et al, 2010), tomate (SHARMA et al, 2010), trigo (ZHUANG et al, 2011), e maçã (ZHUANG et al, 2011). Entretanto, poucos trabalhos com objetivo de caracterizar os membros da família ERF em *Coffea* ssp foram realizados. Bustamante-Porras et al, (2005) isolou o primeiro ERF em *C. canephora* cuja expressão está envolvida nos processos de diferenciação celular e na maturação de frutos. Em *C. arabica* nenhum membro dessa família foi caracterizado até o momento.

Considerando a importância dos ERFs na resposta aos diferentes estresses e a falta de informações da família ERF em *C. arabica*, esse trabalho teve como objetivo a identificação e caracterização dos fatores de transcrição ERF presentes em *Coffea arabica* a partir do banco de dados de ESTs do Genoma Café Brasileiro e estudar a sua expressão transcricional em plantas de café submetidas aos diferentes estresses abióticos.

2 REVISÃO DE LITERATURA

2.1 IMPORTÂNCIA ECONÔMICA DA CAFEICULTURA

O café é um importante produto agrícola que há décadas tem se destacado mundialmente. Entre as commodities, o café ocupa a segunda posição em importância no comércio internacional perdendo somente para o petróleo (VEGA et al., 2003). A cafeicultura movimenta anualmente algo em torno de 12 bilhões de dólares e emprega cerca de meio milhão de pessoas, o que representa uma importante fonte de renda para os países em desenvolvimento como África, Ásia e América Latina (VIEIRA et al, 2006; DaMATTA et al, 2007). Mundialmente estima-se que cerca de 15 bilhões de árvores de café são cultivados numa área de 100.000 km² (SAMSON et al, 2007).

O cafeeiro pertence à família Rubiaceae, a quarta maior família das angiospermas (STEVENS, 2006). O gênero *Coffea* é composto por 103 espécies entre elas, duas são cultivadas para fins comerciais: *Coffea arabica* que representa 70% da produção total e *C. canephora* com cerca de 30%. O *C. arabica* L. é nativo do continente Africano, tendo como seu local de origem o sudoeste da Etiópia, o Sudão e norte do Quênia, sendo preferencialmente cultivado em áreas acima de 600m de altitude, em zonas temperadas (BERTHAUD; CHARRIER, 1988). A espécie *C. arabica* produz um café de melhor qualidade, difundido em todo o mundo, sendo esta bebida consumida na forma de café expresso ou torrado, puro ou em misturas com outras espécies. Por sua vez, *C. canephora* é uma espécie originária da República do Guiné, Libéria, Sudão e Uganda (LASHERMES et al, 1999). É cultivado em áreas de baixa altitude, sob temperaturas mais elevadas, produzindo uma bebida com sabor e aroma neutro.

Dentre as espécies do gênero *Coffea*, *C. arabica* é a única espécie tetraplóide ($2n = 4x = 44$) autógama com 90% de autopolinização das flores, enquanto que as outras espécies são diplóides autoincompatíveis e geralmente se multiplicam exclusivamente por fecundação cruzada (CHARRIER; BERTHAUD, 1985). Houve controvérsias sobre a origem tetraplóide dessa espécie. Os primeiros trabalhos sugeriram que essa espécie poderia ter sido produzida através de autoploidia ou aloploidia. Baseada inicialmente no comportamento meiótico e seu centro de diversidade genética, Carvalho (1952) sugeriu pela primeira vez uma

origem autotetraplóide. Posteriormente, Carvalho (1969), estudando o comportamento de alguns híbridos de *C. arabica*, mostrou que esta espécie possuía estrutura alopoliplóide. Dentro da hipótese da alopoliploidia, possíveis ancestrais foram sugeridos. Thomas (1942) registrou a presença de *C. eugeniodes* e *C. canephora* nas mesmas regiões das florestas de Uganda, o possível centro de origem de *C. arabica*, fazendo com que o autor sugerisse o cruzamento entre essas duas espécies. Pesquisas recentes com base na análise da sequência de DNA sugerem que *C. arabica* resultou da combinação de duas espécies diplóides de *Coffea* (LASHERMES et al, 1993; OROZCO-CASTILLO et al, 1996; Lashermes et al, 1999). Lashermes et al (1999) através de marcadores moleculares e técnicas de citogenética molecular, concluíram que o genoma da *C. arabica* resultou da combinação dos cromossomos de dois genomas distintos, sendo um anfidiplóide formado pelo cruzamento entre *C. eugeniodes* e *C. canephora*, sendo esta a hipótese aceita atualmente.

2.2 RESPOSTA AOS ESTRESSES ABIÓTICOS

As plantas são expostas a condições ambientais adversas como seca, salinidade e temperaturas extremas durante o ciclo de vida. Esses fatores ambientais limitam severamente o crescimento e a produtividade agrícola. Os estresses abióticos têm influenciado de maneira negativa a cafeicultura. A produção de café está sujeita a oscilações regulares que é explicada principalmente pelo ciclo bienal natural, mas também pelos efeitos adversos de condições climáticas. Entre eles, a seca e as altas temperaturas são os principais fatores que afetam o desenvolvimento e a produção do cafeeiro (ASSAD et al, 2004; DAMATTA, 2004; DaMATTA, 2006). Se períodos de seca severa pode levar à morte da planta, os períodos de seca moderada também são muito prejudiciais aos cafeicultores, pois afetam o florescimento, o desenvolvimento do fruto e, conseqüentemente, a produção de café. Além disso, grandes variações na precipitação e temperatura também aumenta os defeitos no fruto, modificando a composição bioquímica e a qualidade final da bebida (SILVA et al, 2005; MAZZAFERA, 2005).

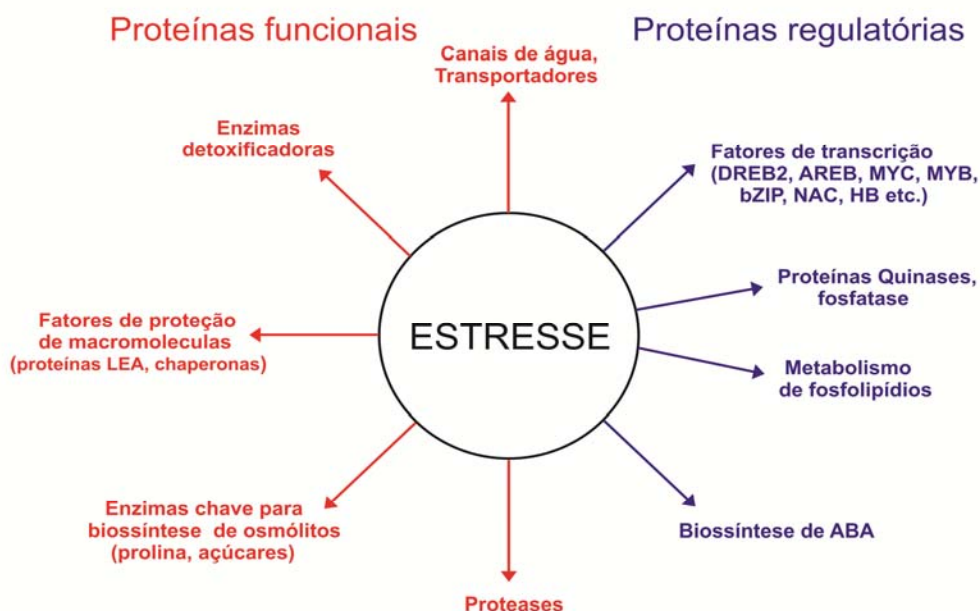
Entre todos os tipos de estresses abióticos, a seca e os extremos de temperatura são os que mais afetam o desenvolvimento vegetal. As plantas respondem ao estresse hídrico e às altas e baixas temperaturas com diversas

mudanças relacionadas ao seu desenvolvimento e a sua fisiologia (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2000). Uma das respostas das plantas aos estresses é a expressão de um grande número de genes, cujos produtos podem estar envolvidos em diversas funções adaptativas em condições de estresse (THOMASHOW, 1999). Genes induzidos durante os estresses abióticos codificam proteínas que não funcionam apenas na proteção das células contra o estresse, mas também na expressão gênica e transdução de sinal em resposta ao estresse (THOMASHOW, 1999; SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2000; Zhu, 2002 ; SHINOZAKI et al, 2003).

Para compreender a natureza e a função dos genes envolvidos na tolerância aos estresses, uma intensa pesquisa foi dedicada à identificação e caracterização de genes estresse-induzidos. Segundo Shinozaki; Yamaguchi-Shinozaki (2006) os produtos dos genes induzidos pelo estresse podem ser classificados em dois grupos (Figura 1): no primeiro está aqueles que protegem a planta contra os estresses ambientais; e no segundo aqueles que regulam a expressão gênica e a transdução de sinais de resposta a estresses.

O primeiro grupo inclui as proteínas que, provavelmente funcionam na tolerância ao estresse abiótico, e inclui as chaperonas, proteínas LEA (“late embryogenesis abundant”), osmotinas, proteínas anticongelantes, proteínas de ligação de mRNA, enzimas chave para a biossíntese de osmólitos, aquaporinas, transportadores de açúcar e prolina, enzimas detoxificadoras e diferentes proteases. O segundo grupo, envolvido na transdução de sinais, estão os fatores de transcrição que desempenham papel fundamental na resposta da planta ao estresse, e incluem ainda as proteínas quinases, fosfatases, enzimas envolvidas no metabolismo de fosfolípidos, e outras moléculas sinalizadoras (SEKI et al., 2003; MARUYAMA et al, 2004). Muitos fatores de transcrição são estresse induzido, sugerindo que diferentes mecanismos regulatórios transcricionais podem funcionar regulando a sinalização das vias de transdução aos estresses a seca, o frio ou salinidade (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2007).

Figura 1 – Genes induzidos durante o estresse hídrico e ao frio e suas possíveis funções na tolerância e resposta ao estresse (adaptado de SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2006).



Condições de seca desencadeiam a produção do ácido abscísico (ABA), que causa o fechamento dos estômatos e induz a expressão de genes relacionados ao estresse. Muitos genes são induzidos por tratamento com ABA exógeno, enquanto outros não. Isto sugere a existência de duas vias de sinalização de estresse: uma dependente de ABA e outra ABA independente. Segundo Shinozaki; Yamaguchi-Shinozaki (2007) existem pelo menos seis vias de transdução de sinais para desencadear a modulação de genes induzidos por seca, alta salinidade e frio, sendo três dependentes de ABA e três independentes (Figura 2). Nas vias dependentes de ABA, ABRE (*ABA-responsive element*) é o principal elemento em resposta ao ABA e os fatores de transcrição envolvidos nesta via são os AREB/ABFs (*ABRE-binding protein/ABRE-binding factor*). MYB2 e MYC2 e NAC são outros fatores que também atuam na expressão gênica induzida por ABA, também são sensíveis ao ácido jasmônico (JA) e podem estar relacionados à resposta a estresses bióticos. Enquanto nas vias independentes de ABA, DRE (*dehydration-responsive element*) é o principal elemento envolvido na regulação de genes induzidos por seca, salinidade e frio. Os fatores DREB1/CBFs (*DRE-binding protein/C-repeat-binding factor*) estão envolvidos na expressão gênica em resposta

ao frio e os DREB2s na resposta a desidratação e alta salinidade. A terceira via independente de ABA é controlada por seca e salinidade, mas não por frio e envolve os fatores de transcrição NAC e HD-ZIP (*homeodomain Leu zipper*).

Figura 2 – Vias de transdução de sinal via ABA dependente e independente. ABRE é o principal elemento em resposta a ABA e os fatores de transcrição desta via são os AREB/ABFs. MYB2 e MYC2 e NAC também são induzidos por ABA, sensíveis ao JA e podem estar relacionados à resposta a estresses bióticos. DRE é o principal elemento das vias independentes de ABA, os fatores DREB1/CBFs estão envolvidos na resposta ao frio, DREB2s na resposta a desidratação e alta salinidade e NAC e HD-ZIP à seca e salinidade, mas não ao frio. (SHINOZAKI;YAMAGUCHI-SHINOZAKI, 2007).



Muitos genes reguladores chave, capazes de controlar um subgrupo de genes, fatores de transcrição e sinalizadores celulares envolvidos na tolerância aos estresses em plantas já foram descritos e seu uso pode ser interessante no melhoramento de várias espécies via engenharia genética já que eles podem regular a expressão de muitos genes (UMEZAWA et al, 2006).

2.3 FATORES DE TRANSCRIÇÃO

Os fatores de transcrição são importantes reguladores da expressão gênica. Em geral, um fator de transcrição é composto por pelo menos dois domínios distintos, um domínio de ligação ao DNA e um domínio de ativação/repressão, que operam em conjunto para regular os diferentes processos fisiológicos e bioquímicos, modulando a taxa de iniciação da transcrição dos genes alvo (PTASHNE, 1988). As plantas atribuem grande parte do seu genoma aos fatores de transcrição. Mais de 1600 genes (7% do genoma) foram identificados em *Arabidopsis* (RIECHMANN et al 2000), revelando a complexidade da regulação em nível transcricional. De acordo com os tipos de domínios de ligação ao DNA, os fatores de transcrição em eucarióticos foram classificados em diversas famílias. Durante as duas últimas décadas, foi mostrado que os fatores de transcrição da superfamília AP2/ERF desempenham um papel importante no controle genético do desenvolvimento floral e na resposta a estresses bióticos e abióticos.

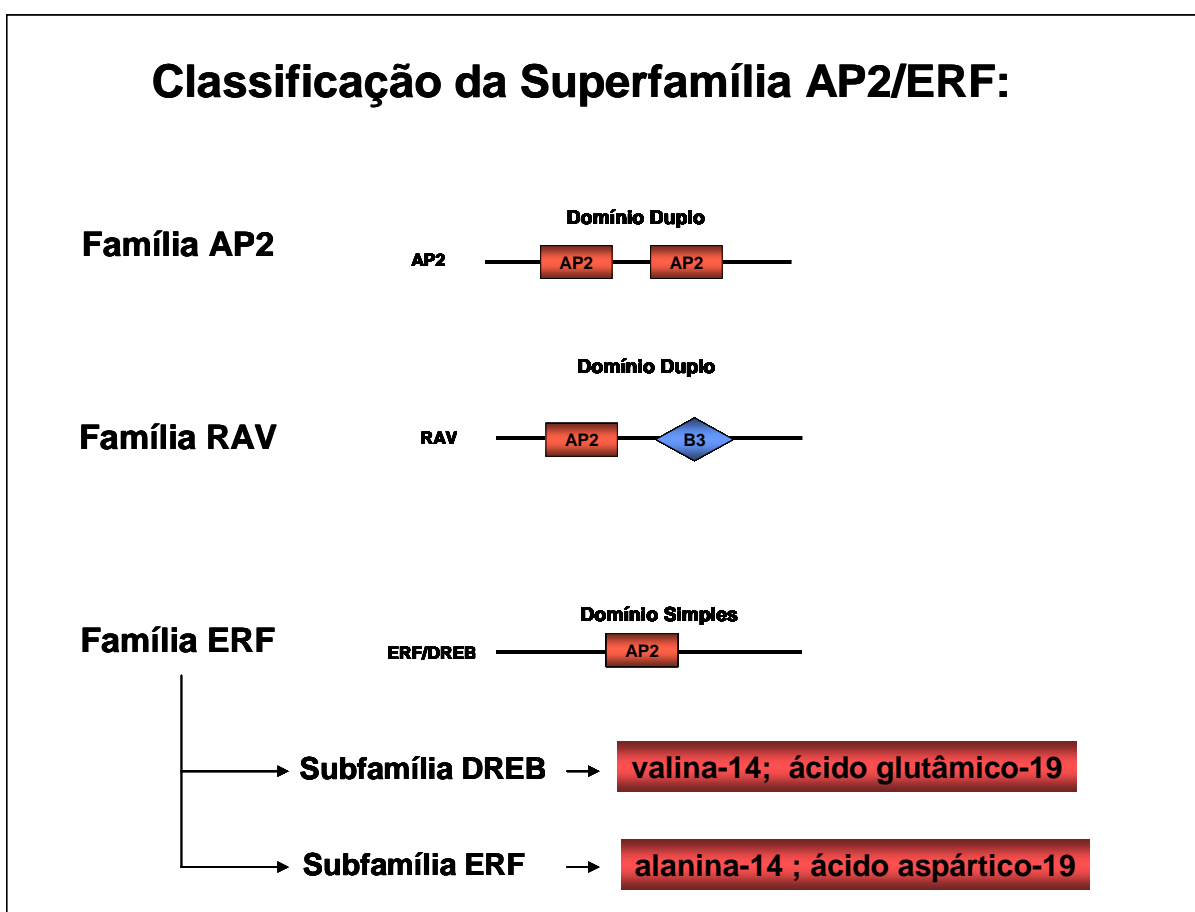
2.3.1 Superfamília AP2/ERF

A superfamília AP2/ERF é caracterizada pela presença do domínio AP2/ERF, composto por 60-70 aminoácidos implicados na ligação ao DNA (JOFUKU et al, 1994; HAO et al, 1998; RIECHMANN et al, 2000). O domínio de ligação ao DNA AP2/ERF é específico de plantas, apesar de haver relatos em ciliados, bactérias e bacteriófagos contendo endonucleases desse domínio (MAGNANI et al, 2004). O domínio AP2 foi identificado pela primeira vez em *Arabidopsis thaliana* como um motivo repetido da proteína AP2 (JOFUKU et al 1994). Mais tarde foram identificadas quatro proteínas de ligação ao DNA em *Nicotiana tabacum* chamados de Proteínas de Ligação ao Elemento Etileno-Responsivo (*Ethylene-Responsive Element Binding Proteins* -EREBPs) 1, 2, 3 e 4 (OHME-TAKAGI; SHINSHI, 1995), atualmente denominadas de Fatores de Resposta ao Etileno (ERFs) 1, 2, 3 e 4. A estrutura do domínio ERF de AtERF1 mostrou que esse domínio consiste em uma fita de três β -sheets anti-paralelas e uma α -helix paralelo às β -sheets (ALLEN et al, 1998).

A superfamília AP2/ERF foi dividida em três famílias com base no número de domínios AP2 presentes (SAKUMA et al, 2002; NAKANO et al, 2006). As

proteínas da família AP2 contém dois domínios repetidos AP2/ERF (JOFUKU e AI, 1994); as proteínas da família RAV além de um único domínio AP2/ERF, contém um domínio B3, que é um domínio de ligação ao DNA, conservado em outros fatores de transcrição, incluindo VP1/ABI3 (KAGAYA et al, 1999; SAKUMA et al, 2002; NAKANO et al, 2006) e as proteínas da família ERF contém um único domínio AP2/ERF. A família ERF, por sua vez, é dividida em duas subfamílias principais, a subfamília ERF e a subfamília CBF/DREB (Figura 3) (SAKUMA et al, 2002; NAKANO et al, 2006).

Figura 3 – Organização estrutural da superfamília AP2 (adaptado de SHARONI et al, 2011).



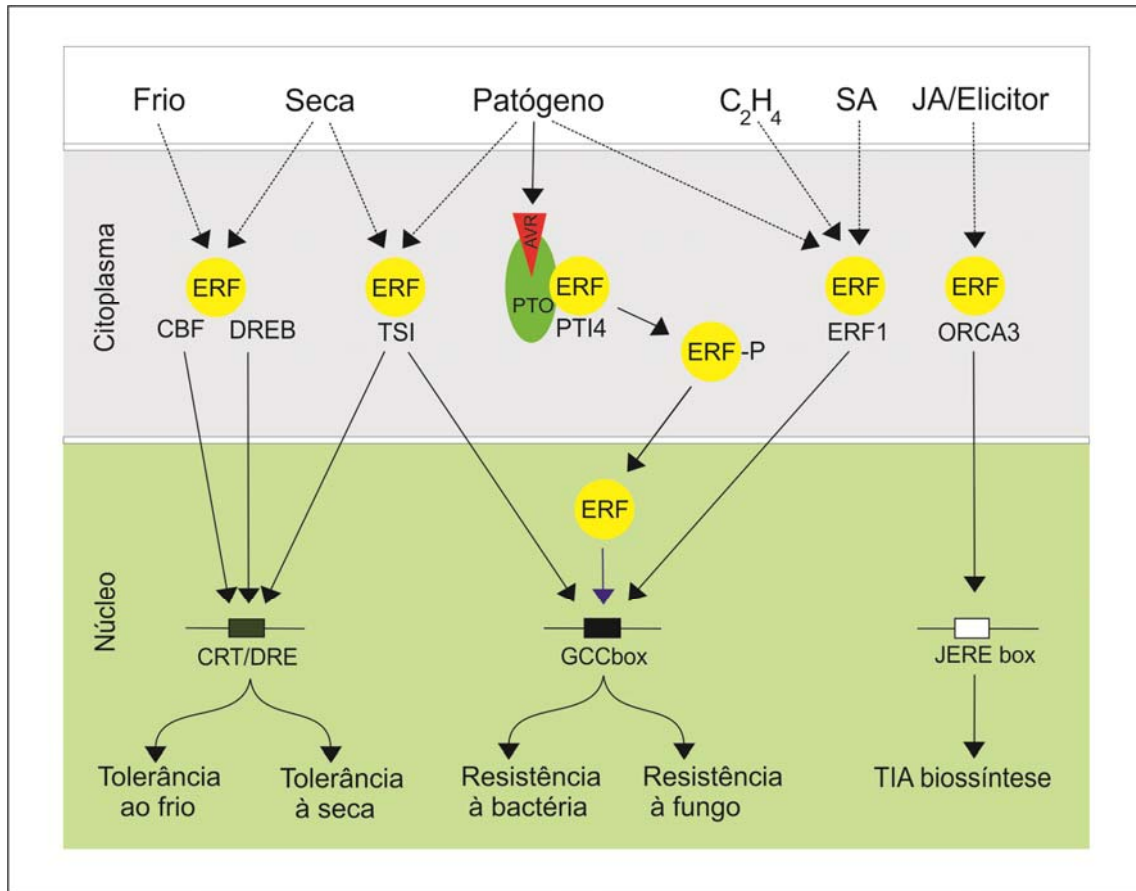
2.3.2 Diferenças e classificação da família ERF

Os fatores de transcrição da família ERFs foram identificados em várias espécies vegetais, incluindo arábida (LIU et al, 1998; SAKUMA et al, 2002; NAKANO et al, 2006), arroz (NAKANO et al, 2006; SHARONE et al, 2011),

algodão (CHAMPION et al, 2009), soja (ZHANG et al, 2008), Populus (ZHUANG et al, 2008), uva (ZHUANG et al, 2009; LICAUSI et al, 2010), milho (ZHUANG et al, 2010), tomate (SHARMA et al, 2010) , trigo (ZHUANG et al, 2011), e maçã (ZHUANG et al, 2011). Nakano et al (2006) classificou os ERFs em arabidopsis em 10 grupos principais. Os membros dos grupos I, II, III, IV são caracterizadas pela presença de uma valina na posição 14 e um ácido glutâmico na posição 19 do domínio AP2. Estes membros são classificados como DREBs (Dehydration Response Element Binding). Enquanto que os membros dos grupos V, VI, VII, VIII, IX e X são denominados de ERFs e são caracterizados por uma alanina e um ácido aspártico, respectivamente, nas posições 14 e 19 do domínio AP2 (SAKUMA et al., 2002).

De modo geral, as proteínas da subfamília CBF/DREB desempenham um papel crucial na resposta das plantas a estresses abióticos, reconhecendo o elemento responsivo à desidratação (DRE) com um motivo central A/GCCGAC (Figura 4) (SHINOZAKI; SHINOZAKI,1994; THOMASHOW, 1999). Os membros da subfamília ERF estão envolvidos principalmente em respostas a estresses bióticos, reconhecendo o elemento cis-acting AGCCGCC denominado GCC-box (HAO et al., 1998). No entanto, além de sua participação nas respostas bióticas, algumas proteínas da família ERF podem integrar diferentes vias de sinalização. Por exemplo, Tsi1 de tabaco (PARK et al, 2001), TERF1 de tomate (ZHANG et al, 2005) e CaFP1 de pimenta (YI et al, 2004) são integradores de respostas aos estresses biótico e abiótico através da interação com GCC-box e DRE. Assim, as proteínas ERF desempenham um papel importante, não apenas em respostas de defesa ao ataque a patógenos, mas também em tolerância a diferentes estresses abióticos (Figura 4).

Figura 4 – Esquema das funções fisiológicas dos ERFs em resposta aos diferentes estresses (SINGH et al, 2002).



2.3.2.1 Subfamília DREB

Segundo Nakano et al, (2006), os fatores de transcrição DREBs podem ser divididos em 4 grupos (I, II, III e IV) com base em semelhanças no domínio de ligação. Estes fatores de transcrição reconhecem o elemento CRT/DRE responsivo ao frio e a desidratação. Estes elementos estão presentes nas regiões dos promotores de muitos genes responsivos ao frio e a desidratação em *Arabidopsis*, incluindo COR (cold-regulated) (SHINOZAKI; SHINOZAKI, 1994; THOMASHOW, 1999). A expressão constitutiva dos genes CBF em plantas transgênicas de *Arabidopsis* resultou na indução da expressão do gene COR e um aumento da tolerância ao congelamento, sem estímulo de baixa temperatura (KASUGA et al, 1999; GILMOUR et al, 2000). Significativamente, várias alterações bioquímicas são associadas com a aclimação ao frio. Gilmour et al., (2000) mostraram que em plantas transgênicas de *arabidopsis* que expressam constitutivamente CBF3 houve acúmulo de açúcares e de prolina em plantas não

aclimatadas, o que pode contribuir para o aumento da tolerância ao congelamento. Além disso, Gilmour et al (2000) propuseram que os genes CBF atuam na ativação de múltiplos componentes atuando na resposta a aclimação ao frio.

Além da indução ao frio, Huang et al, (2008) demonstrou que a expressão de *GhDBP2* de algodão é induzido pela seca, salinidade e ABA. Trabalhos tem mostrado que alguns DREBs , como é o caso de *GhDBP2* podem estar envolvidos na regulação de alguns genes do tipo LEA (Late Embryogenesis Abundant) (KIZIS;PAGES 2002; HUANG et al, 2008). Essas proteínas acumulam-se durante os estágios avançados da embriogênese e também em tecidos expostos a estresses como desidratação, estresse osmótico e baixa temperatura (HONG-BO et al, 2005). *GhDBP3* está envolvido no caminho de sinalização do estresse osmótico e o ABA participa da regulação da resposta a esses estresses ambientais (HUANG; LIU, 2006). Resultados semelhantes foram encontrados em *TINY2* em arábido, membro do grupo III (NAKANO et al, 2006). Este que foi expresso em baixos níveis em plantas não tratadas, entretanto foi significativamente induzida após tratamento com ABA, NaCl e seca (WEI et al, 2005).

Poucos trabalhos foram realizados a fim de estudar o papel dos DREBs em resposta ao estresse térmico, embora análises *in silico* tenham mostrado que os genes contendo DRE/CRT podem ser induzidos por choque térmico (GEISLER et al, 2006). Sakuma et al, (2006) mostrou que a indução da expressão ao estresse térmico de HsfA3 em arábido depende de *DREB2A*. Mais tarde, Chen et al (2010) demonstraram que HsfA3 também é transcricionalmente ativado durante o choque térmico por *DREB2C*. Em plantas transgênicas de arábido, a superexpressão constitutivamente de *DREB2A* de arábido e *ZmDREB2A* de milho ativou a expressão de muitos genes induzidos por estresse ao calor, tais como proteínas de choque térmico (Hsps) e HsfA3 cuja indução resultou em uma maior termotolerância (QIN et al, 2007; SCHRAMM et al, 2008).

2.3.2.2 Subfamília ERF

Os membros da subfamília ERF ligam-se ao elemento cis-regulatório, denominado GCC-box nas regiões promotoras dos genes relacionadas à patogênese regulando sua expressão em resposta ao etileno. Análise detalhada *in vitro* da ligação de três ERFs de tabaco indicaram que os contatos de

aminoácidos específicos dos ERFs limitam-se a 6 pb (GCCGCC), onde G1, G4 e C6 são essenciais para que haja a ligação (HAO et al, 1998). Essas seqüências podem também modular diferentes ligações por diferentes ERFs (SHINSHI, 2008).

O gene Tsi1 (Tobacco stress-induced gene 1) confere resistência tanto ao ataque de patógenos como ao estresse osmótico quando superexpresso. Desta forma, demonstrou-se que o domínio AP2/ERF pode se ligar tanto a GCC-box bem como com o elemento DRE (PARK et al, 2001). Mais tarde, diversas proteínas ERFs capazes de se ligar a GCC-box e DRE foram identificados e incluem DREB2A (SAKUMA et al, 2002), CBF1 (HAO et al, 2002), JERF1 (ZHANG et al, 2004), e BnDREBIII-1 (LIU et al, 2006), sugerindo que as vias de sinalização mediada por DREB e ERF não são exclusivas e que existe uma comunicação cruzada (*crosstalk*) entre elas (LEE et al, 2004, ZHANG et al, 2008). A superexpressão de genes da família do ERF em arábido ou tabaco tem demonstrado conferir maior resistência a estresses bióticos e abióticos (PARK et al, 2001; BERROCAL-LOBO et al, 2002; XU et al, 2007; PAN et al, 2010).

Combinando abordagens genético-moleculares, uma série de genes reguladores da subfamília ERF envolvidos em diferentes vias metabólicas tem sido analisados, incluindo aqueles relacionados à seca (YAMAGUCHI-SHINOZAKI; SHINOZAKI, 2006, DUBOUZET et al, 2003), à salinidade (DUBOUZET et al, 2003, FU et al, 2007), à baixa temperatura (YANG et al, 2005; QIN et al, 2007) e a doenças (GUTTERSON; REUBER, 2004). Além da resposta a estresses abióticos, a subfamília ERF também está envolvida na formação de raízes (BANNO et al, 2001), germinação (PIRRELLO et al, 2006) e desenvolvimento de frutos e maturação (WANG et al, 2007; EL-SHARKAWY et al, 2009).

Além dessas funções alguns ERFs podem estar envolvidos na nodulação de algumas espécies leguminosas. A interação simbiótica responsável pela nodulação necessita da expressão dos genes-alvos, os fatores Nod. Middleton et al, (2007) identificaram um ERF, necessário para a nodulação (ERN – ERF Required for nodulation), identificado como um trans-ativador dos fatores Nod. Asamizu et al (2008) propuseram que *LjERF1* funciona como um regulador chave da infecção bem sucedida de *L. japonicus* por *Mesorhizobium loti*. Andriankaja et al, (2007) demonstraram que ERN1 e ERN2 mediam a expressão de *MtENOD11*, um gene marcador pré-infecção e posterior estágio da associação simbiótica (JOURNET et al, 2001).

Trabalhos têm mostrado que pode existir uma expressão grupo-específica. As proteínas do grupo VII estão associados à maturação em frutos. Os membros do grupo VII possuem o motivo conservado MCGGAIL/L identificado por Tournier et al, (2003). Este autor demonstrou que *SlERF2* se liga a *GCC-box* desempenhando um importante papel durante o processo de maturação em tomate. O mesmo foi observado por Yin et al, (2010) em diferentes ERFs expressos durante o processo de maturação dos frutos de kiwi. *MdERF1* em maçã (WANG et al, 2007), *PsERF2a* e *PsERF2b* em ameixa (EL-SHARKAWY et al, 2009) e *AdERF10* e *AdERF14* em kiwi (YIN et al, 2010) são proteínas que se expressam durante a maturação nos frutos pertencentes ao grupo VII. Bustamante-Porras et al, (2005) isolou em *C. canephora* um gene ERF (CoERF) cuja expressão demonstrou estar envolvida nos processos de diferenciação celular e na maturação de frutos. Entretanto trabalhos têm mostrado que este grupo também pode estar envolvido nas respostas aos diferentes estresses. Jin et al, (2010) observou que quando as plantas de algodão eram tratados com etileno, ácido abscísico, sal, frio e seca, os transcritos de *GhERF2*, *GhERF3* e *GhERF6* foram rapidamente levados a níveis elevados. O mesmo foi observado por Wang et al (2010) em *GoDREB* de *Galegae orientalis*. A superexpressão de *GoDREB* em plantas transgênicas de tabaco resultou num aumento da tolerância à salinidade. A expressão ectópica do gene *GmERF3* em plantas transgênicas de tabaco induziu a expressão de alguns genes PR aumentando a resistência contra a infecção por *Ralstonia solanacearum*, *Alternaria alternata* e vírus do mosaico do tabaco (TMV), assim como aumentou a tolerância à salinidade e seca (ZHANG et al, 2009). Além disso, a superexpressão de *GmERF3* em tabaco transgênico levou a níveis mais elevados de prolina livre e carboidratos solúveis em plantas sob condições de seca.

O motivo EAR (ERF-associated amphiphilic repression EAR) é encontrado em membros do grupo VIII. Ohta et al, (2001) identificaram um motivo EAR, que tem se mostrado funcionar como um domínio de repressão. O motivo EAR de seqüência conservada, (L/F)DLN(L/F)xP, encontra-se nas regiões C-terminal do grupo VIII. Esse motivo foi identificado em várias proteínas repressoras incluindo ZAT7, 10, 12, ERF3, AUX/IAA, NIMIN1, HSI2, SUPERMAN (arabidopsis), NRR (arroz), ZFT1 (tabaco) e ZPT2-3 (petunia), que desempenham diferentes funções que vão desde o desenvolvimento das plantas até a tolerância aos diferentes estresses (TIWARI et al, 2004; HIRATSU et al, 2004; KAZAN, 2006). Dois membros

do grupo VIII, AtERF4 e AtERF7, foram bem caracterizados. Ambas as proteínas regulam negativamente o gene *PDF1.2*, no qual contem *GCC-box* na região promotora. AtERF4 demonstrou ser um regulador negativo da expressão de etileno, JA e genes ABA-responsivos (OHTA et al, 2001; McGRATH et al, 2005). Além disso, foi mostrado que AtERF7 atua como um regulador negativo nas respostas ao ABA (YANG et al, 2005). Pan et al, (2010) mostrou que *SlERF3* foi induzido pelo frio, seca e sal. A superexpressão de plantas transgênicas que tiveram o motivo EAR deletado induziu a expressão de genes como *PR1*, *PR2* e *PR5* aumentando a tolerância a *Ralstonia solanacearum*. Por outro lado, Zhang et al, (2010) demonstrou que a expressão de *GmERF4* em soja foi induzido pelo etileno, JA, o frio, o sal, a seca, e o vírus do mosaico da soja. As plantas transgênicas de tabaco *GmERF4* mostraram reprimir a expressão da *GCC-box* contendo os genes *PR1*, *PR2*, *PR4*, osmotina e SAR8.2. A expressão constitutiva dessas plantas promoveu um aumento da tolerância aos estresses salino e a seca, porém se mostraram suscetíveis a infecção bacteriana. Desta forma, os membros das proteínas ERFs que contêm o motivo EAR são capazes de se expressar diferencialmente e alguns deles podem ser induzidos por estresses bióticos e abióticos (FUJIMOTO et al, 2000).

Os membros do grupo IX apresentam indução da expressão quando são submetidos ao ataque de vários patógenos. A superexpressão constitutiva de *AtERF2* do subgrupo IXa mostrou induzir a expressão gênica dos gene *PDF1.2* (BROWN et al, 2003). Da mesma forma, a superexpressão do gene *AtERF1* de *Arabidopsis*, um homólogo próximo a *AtERF2*, confere resistência a *Botrytis cinerea*, *Sclerotinia sclerotiorum* e *Erysiphe orontii* em *arabidopsis* (GUTTERSON; REUBER, 2004). Recentemente Anderson et al, (2010) demonstrou que a superexpressão do gene *MtERF1-1* em raízes de *Medicago truncatula* aumentou a resistência a *Rhizoctonia solani*, assim como a *Phytophthora medicaginis*. Krishnaswamy et al, (2011) caracterizou *RAP2.6* e *RAP2.6L* de *Arabidopsis*. A expressão de *RAP2.6* foi induzida pelos estresses salinos, calor e seca, enquanto que *RAP2.6L* foi induzida pelos estresses salinos e seca. A superexpressão de *RAP2.6L* em transgênicos de *arabidopsis* aumentou a tolerância aos estresse salinos e a seca.

3 REFERÊNCIAS

- ASSAD, E. D.; PINTO, H. S.; ZULLO JUNIOR J.; ÁVILA, A. M. H. Impacto das mudanças climáticas no zoneamento agroclimático no café no Brasil. **Pesq. Agropec. Bras.**, v.39, n.11, p. 1057-1064. 2004.
- ALLEN, M. D.; YAMASAKI, K.; OHME-TAKAGI, M.; TATENO, M.; SUZUKI, M. 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. **EMBO J.**, v.15, p. 5484-5496. 1998.
- ANDERSON, J. P.; LICHTENZVEIG, J.; GLEASON, C.; OLIVER, R. P.; SINGH, K. B. The B-3 Ethylene Response Factor MtERF1-1 Mediates Resistance to a Subset of Root Pathogens in *Medicago truncatula* without Adversely Affecting Symbiosis with Rhizobia. **Plant Physiology**, v.154, p. 861–873. 2010.
- ANDRIANKAJA, A.; BOISSON-DERNIER, A.; FRANCES, L.; SAUVIAC, L.; JAUNEAU, A.; BARKER, D. G.; de CARVALHO-NIEBEL, F. AP2-ERF transcription factors mediate Nod factor dependent Mt ENOD11 activation in root hairs via a novel cis-regulatory motif. **Plant Cell**, v.19, p. 2866-2885. 2007.
- ASAMIZU, E.; SHIMODA, Y.; KOUCHI, HI.; TABATA, S.; SATO, S. A Positive Regulatory Role for LjERF1 in the Nodulation Process Is Revealed by Systematic Analysis of Nodule-Associated Transcription Factors of *Lotus japonicus*. **Plant Physiology**, v.147, p. 2030–2040. 2008.
- BANNO, H.; IKEDA, Y.; NIU, Q. W., CHUA, N. H. Overexpression of Arabidopsis ESR1 induces initiation of shoot regeneration. **Plant Cell**, v.13, n.12, 2609-2618. 2001.
- BERROCAL-LOBO, M.; MOLINA, A.; SOLANO, R. Constitutive expression of ETHYLENE-RESPONSE-FACTOR1 in Arabidopsis confers resistance to several necrotrophic fungi. **Plant J.**, v.29, p. 23-32. 2002.
- BERTHAUD, J.; CHARRIER, A. Genetics resources of Coffea. In: CLARKE, R. J., MACRAE, R. **London Elsevier Applied Science**, v.4, p 1-42. 1988.
- BROWN, R. L.; KAZAN, K.; MCGRATH, K.C.; MACLEAN, D.J.; MANNERS, J. M. A role for the GCC-box in jasmonate-mediated activation of the *PDF1.2* gene of *Arabidopsis*. **Plant Physiol.**, v. 132, p.1020–1032. 2003.
- BUSTAMANTE-PORRAS, J.; NOIROT, M.; CAMPA, C.; HAMON, S.; KOCHKO. Isolation and characterization of a *Coffea canephora* ERF-like cDNA. **African Journal of Biotechnology**, v.4, p. 157-159, 2005.
- CARVALHO, A. Taxonomia de *Coffea arabica* L. Caracteres morfológicos dos haploides. **Bragantia**, v.12, p. 201–212. 1952.

CARVALHO, A.; FERWERDA, F. P.; FRAHM-LELIVELD J. A.; MEDINA P. M.; MENDES, A. J. T.; MÔNACO, L. C. Coffee. In: FERWERDA, F. P, WIT, F. (Ed.). **Outlines of perennial crop breeding in the tropics**. Veenman & Zonen, Wageningen, p. 189-241. 1969.

CHAMPION, A.; HEBRARD, E.; PARRA, B.; BOURNAUD, C.; MARMEY, P.; TRANCHANT, C.; NICOLE, M. Molecular diversity and gene expression of cotton ERF transcription factors reveal that group IXa members are responsive to jasmonate, ethylene and Xanthomonas. **Molecular Plant Pathology**, v.10, p. 471-485. 2009.

CHARRIER, A.; BERTHAUD, J. Botanical classification of coffee. In: CLIFFORD, M. N, WILSON, K. C (Ed.) **Coffee: botany, biochemistry and production of beans and beverage**. AVI Publishing Company, Westport, Connecticut, p. 13-47.1985.

CHEN, Y. F.; ETHERIDGE, N.; SCHALLER, G. E. Ethylene signal transduction. **Ann Bot.**, v.95, p. 901–915. 2005.

CLARK, K. L.; LARSEN, P. B.; WANG, X.; CHANG, C. Association of the Arabidopsis CTR1 Raf-like kinase with the ETR1 and ERS ethylene receptors. **Proc. Natl. Acad. Sci.** v.95, p. 5401-5406. 1998.

DaMATTA, F. M.; RONCHI C. P.; MAESTRI, M.; BARROS, R. S. Ecophysiology of coffee growth and production. **Braz. J. Plant Physiol**, v.19, n.4, p. 485-510. 2007.

DaMatta, F. M. Exploring drought tolerance in coffee: a physiological approach with some insights for plant breeding. **Braz. J. Plant. Physiol**, v.16, n.1, p. 1-6. 2004.

DaMATTA, F. M.; RAMALHO, J. C. Impact of drought and temperature stress on coffee physiology and production: a review. **Braz J Plant Physiol**, v.18, n.1, 55-81. 2006.

DUBOUZET, J.G.; SAKUMA, Y.; ITO, Y.; KASUGA, M.; DUBOUZET, E. G.; MIURA, S.; SEKI, M.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. *OsDREB* genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. **Plant. J.**, v.33, p.751-763, 2003.

EL-SHARKAWY, I.; SHERIF, S.; MILA, I.; BOUZAYEN, M.; JAYASANKAR, S. Molecular characterization of seven genes encoding ethylene-responsive transcriptional factors during plum fruit development and ripening. **J. Exp. Bot.**, v.60, p. 907–922. 2009.

FUJIMOTO, S. Y.; OHTA, M.; USUI, A.; SHINSHI, H.; OHME-TAKAGI, M. Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. **Plant Cell**, v.12, p. 393-404. 2000.

GEISLER, M.; KLECZKOWSKI, L. A.; KARPINSKI, S. A universal algorithm for genome-wide in silico identification of biologically significant gene promoter putative cis-regulatory-elements; identification of new elements for reactive oxygen species and sucrose signaling in Arabidopsis, **Plant J.**, v.45, p. 384–398. 2006.

GILMOUR, S. J.; SEBOLT, A. M.; SALAZAR, M. P.; EVERARD, L. D.; THOMASHOW, M. F. Overexpression of *Arabidopsis* CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation. **Plant Physiol.**, v.124, p. 1854-1865. 2000.

GUTTERSON, N.; REUBER, T. L. Regulation of disease resistance pathways by AP2/ERF transcription factors. **Curr. Opin. Plant Biol.**, v.7, p. 465-471. 2004.

HAO, D.; OHME-TAKAGI, M.; SARAI, A. Unique mode of GCC box recognition by the DNA-binding domain of ethylene responsive element-binding factor (ERF domain) in plants. 1. **Biol.Chem.** v.273, p. 26857-26861. 1998.

HAO, D.; YAMASAKI, K.; SARAI, A.; OHME-TAKAGI, M. Determinants in the sequence specific binding of two plant transcription factors, CBF1 and NtERF2, to the DRE and GCC motifs. **Biochemistry**, v.41, p. 4202-4208. 2002.

HIRATSU, K.; MITSUDA, N.; MATSUI, K.; OHME-TAKAGI, M. Identification of the minimal repression domain of SUPERMAN shows that the DLELRL hexapeptide is both necessary and sufficient for repression of transcription in *Arabidopsis*. **Biochem Biophys Res Commun.**, v.321, p. 172-178. 2004.

HUANG, B.; LIU, J. Y. Cloning and functional analysis of the novel gene GhDBP3 encoding a DRE-binding transcription factor from *Gossypium hirsutum*. **Biochimica et Biophysica Acta**, v.1759, p. 263-269. 2006.

HONG-BO, S.; ZONG-SUO, L.; MING-AN, S. LEA proteins in higher plants: structure, function, gene expression and regulation. *Colloids Surf B: Biointerfaces*, v.45, p. 131-5. 2005.

ICO. **International Coffee Organization**. Disponível em:
<<http://www.ico.org/prices/po.htm>> Acesso em: 20 maio 2011.

JIN, L. G.; LI, H.; LIU, J. Y. Molecular Characterization of Three Ethylene Responsive Element Binding Factor Genes from Cotton. **Journal of Integrative Plant Biology**, v.52, n.5, p. 485-495. 2010.

JOFUKU, K. D.; DEN BOER, B. G.; VAN MONTAGU, M.; OKAMURO, J. K. Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2. **Plant Cell**, v.6, p.1211-1225. 1994.

JOURNET, E. P.; EL-GACHTOULI, N.; VERNOUD, V.; DE BILLY, F.; PICHON, M.; DEDIEU, A.; ARNOULD, C.; MORANDI, D.; BARKER, D. G.; GIANINAZZI-PEARSON, V. *Medicago truncatula* ENOD11: a novel RPRP-encoding early nodulin gene expressed during mycorrhization in arbuscule-containing cells. **Mol. Plant Microbe Interact.**, v.14, p. 737-748. 2001.

KAGAYA, Y.; OHMIYA, K.; HATTORI, T. RAVI, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. **Nucleic Acids Res.**, v.27, p. 470-478. 1999.

KASUGA, M.; LIU, Q.; MIURA, S.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Improving plant drought, salt and freezing tolerance by gene transfer of a single stress-inducible transcription factor, **Nature Biotechnol.** v.17, p. 287-291. 1999.

KAZAN, K. Negative regulation of defence and stress genes by EAR-motif-containing repressors. **Trends Plant Sci.**, v.11, p. 109–112. 2006.

KIZIS, D.; PAGES, M. Maize DRE-binding proteins DBF1 and DBF2 are involved in rab17 regulation through the drought-responsive element in an ABA-dependent pathway. **Plant J.**, v.30, p. 679–689. 2002.

KRISHNASWAMY, S.; VERMA, S.; RAHMAN, M. H.; KAV, N. N. V. Functional characterization of four APETALA2-family genes (RAP2.6, RAP2.6L, DREB19 and DREB26) in Arabidopsis. **Plant Mol. Biol.**, v.75, p. 107–127. 2011.

LASHERMES, P.; COMBES, M. C.; ROBERT J, TROUSLOT, P.; D'HONT, A.; ANTHONY, F.; CHARRIER, A. Molecular characterisation and origin of the *Coffea arabica* L. genome. **Mol. Gen. Genet.**, v.261, p. 259-266. 1999.

LASHERMES, P.; CROS, J.; MARMEY, P.; CHARRIER, A. Use of random amplified DNA markers to analyse genetic variability and relationships of *Coffea* species. **Genetic Resources and Crop Evolution**, v.40, p. 91–99. 1993.

LEE, J. H.; HONG, J. P.; OH, S. K, LEE, S.; CHOI, D.; KIM, W. T. The ethylene-responsive factor like protein 1 (CaERFLP1) of hot pepper (*Capsicum annuum* L.) interacts in vitro with both GCC and DRE/CRT sequences with different binding affinities: possible biological roles of CaERFLP1 in response to pathogen infection and high salinity conditions in transgenic tobacco plants. **Plant Mol Biol.**, v.55, p. 61–81. 2004.

LICAUSI, F.; GIORGI, F. M.; ZENON, S.; OSTI, F.; PEZZOTTI, M.; PERATA, P. Genomic and transcriptomic analysis of the AP2/ERF superfamily in *Vitis vinifera*. **BMC Genomics.**, v.11, p. 719. 2010.

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

LIU, Y.; ZHAO, T. J.; LIU, J. M.; LIU, W. Q.; LIU, Q, YAN, Y. B.; ZHOU H. M. The conserved Ala37 in the ERF/AP2 domain is essential for binding with the DRE element and the GCC box. **FEBS Lett.**, v.580, p. 1303–1308. 2006.

MAGNANI, E.; SJOLANDER, K.; HAKE, S. From endonucleases to transcription factors: evolution of the AP2 DNA binding domain in plants. **Plant Cell**, v.16, p. 2265-2277. 2004.

MAPA. **Ministério da Agricultura, Pecuária e Abastecimento**. Disponível em: <http://www.agricultura.gov.br/arq_editor/file/3894_relatorio_do_mercado_cafeeiro_-_marco_2011.pdf>. Acesso em: 5 abr. 2011.

- MARUYAMA, K.; SAKUMA, Y.; KASUGA, M.; ITO, Y.; SEKI, M.; GODA, H.; SHIMADA, Y.; YOSHIDA, S.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Identification of cold-inducible downstream genes of the *Arabidopsis* DREB/CBF3 transcriptional factor using two microarray systems. **The Plant Journal**, v.38, p. 982-993, 2004.
- McGRATH, K. C.; DOMBRECHT, B, MANNERS, J. M.; SCHENK, P. M.; EDGAR, C. I.; MACLEAN, D. J.; SCHEIBLE, W. R.; UDVARDI, M. K.; KAZAN, K. Repressor- and activator-type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genome-wide screen of *Arabidopsis* transcription factor gene expression. **Plant Physiol.**, v.139, p. 949–959. 2005.
- MAZZAFERA, P. Estresse hídrico e conseqüências na composição das sementes de café e qualidade da bebida In: SALVA, T. J, GUERREIRO-Filho, O, THOMAZIELLO, R. A.; FAZUOLI, L. C. **Cafés de qualidade: aspectos tecnológicos, científicos e comerciais**. Instituto Agronômico Campinas; 2007, p. 73-90.
- MIDDLETON, P. H.; JAKAB, J.; PENMETSA, R. V.; STARKER, C. G.; DOLL, J.; KALO, P.; PRABHU, R.; MARSH, J. F.; MITRA, R. M.; KERESZT, A.; DUDAS, B.; VANDENBOSCH, K.; LONG, S. R. COOK, D. R.; KISS, G. B.; OLDROYD, G. E. An ERF transcription factor in *Medicago truncatula* that is essential for Nod factor signal transduction. **Plant Cell**, v.19, p. 1221-1234. 2007.
- NAKANO, T.; SUZUKI, K.; FUJIMURA, T.; SHINSHI, H. Genome-Wide Analysis of the ERF Gene Family in *Arabidopsis* and Rice. **Plant Physiol.**, v.140, p. 411-432. 2006.
- OHME-TAKAGI, M.; SHINSHI, H. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. **Plant Cell**, v.7, p. 173-182. 1995.
- OHTA, M.; MATSUI, K.; HIRATSU, K.; SHINSHI, H.; OHME-TAKAGI, M. Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. **The Plant Cell**, v.13, p. 1959–1968. 2001.
- OROZCO-CASTILLO, C.; CHALMERS, K. J.; WAUGH, R.; POWELL, W. RAPD and organelle specific PCR re-affirms taxonomic relationships within the genus *Coffea*. **PI Cell Rep.** v.15, p. 337–341. 1996.
- PAN, I. C.; LI, C. W.; SU, R. C.; CHENG, C. P, LIN, C. S.; CHAN, M. T. Ectopic expression of an EAR motif deletion mutant of SIERF3 enhances tolerance to salt stress and *Ralstonia solanacearum* in tomato. **Planta**, v.232, p. 1075–1086. 2010.
- PARK, J. M.; PARK, C. J.; LEE, S. B.; HAM, B. K.; SHIN, R.; PAEK, K. H. Overexpression of the tobacco Tsi1 gene encoding an EREBP/ AP2-type transcription factor enhances resistance against pathogen attack and osmotic stress in tobacco. **Plant Cell**, v.13, p. 1035-1046. 2001.
- PAY, E: **The market for organic and fair-trade coffee**. *FAO Rome* 2009.

- PIRRELLO, J.; JAIMES-MIRANDA, F.; SANCHEZ-BALLESTA, M. T.; TOURNIER, B.; KHALIL-AHMAD, Q.; REGAD, F.; LATSHE, A.; PECH, J. C.; BOUZAYEN, M. SIERF2, a tomato ethylene response factor involved in ethylene response and seed germination. **Plant Cell Physiol.**, v.47, n. 9, p. 1195–1205. 2006.
- PTASHNE, M. How eukaryotic transcriptional activators work. **Nature**, v.335, p. 683–689. 1988.
- QIN F.; KAKIMOTO, M.; SAKUMA, Y.; MARUYAMA, K.; OSAKABE, Y.; TRAN L. S.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Regulation and functional analysis of ZmDREB2A in response to drought and heat stresses in *Zea mays* L, **Plant J.**, v.50, p. 54–69. 2007.
- RIECHMANN, J. L.; HEARD, J.; MARTIN, G.; REUBER, L.; JIANG, C.; KEDDIE, J.; ADAM, L.; PINEDA, O.; RATCLIFFE, O. J.; SAMAHA, R. R. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. **Science**, v.290, n.5499, p. 2105–2110. 2000.
- SAKUMA, Y.; LIU, Q.; DUBOUZET, J.G.; ABE, H.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration and cold-inducible gene expression. **Biochem. Biophys. Res. Commun.**, v.290, p. 998–1009. 2002.
- SAKUMA, Y.; MARUYAMA, K.; QIN, F.; OSAKABE, Y.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Dual function of an Arabidopsis transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression. **Proc. Natl. Acad. Sci. USA**, v.103, p. 18822–18827. 2006.
- SAMSON, N.; BAUSHER, M. G.; LEE, S. B.; JANSEN, R. K.; DANIELL, H. The complete nucleotide sequence of the coffee (*Coffea arabica* L.) chloroplast genome: organization and implications for biotechnology and phylogenetic relationships amongst angiosperms. **Plant Biotechnology Journal**. v. 5, p. 339–353. 2007.
- SCHRAMM, F.; LARKINDALE, J.; KIEHLMANN, E.; GANGULI, A.; ENGLISH, G.; VIERLING, E.; VON KOSKULL-DÖRING, P. A cascade of transcription factor DREB2A and heat stress transcription factor HsfA3 regulates the heat stress response of Arabidopsis. **The Plant Journal**, v.53, p. 264–274. 2008.
- SHARONI, A. M.; NURUZZAMAN, M.; SATOH, K.; SHIMIZU, T.; KONDOH, H.; SASAYA, T.; CHOI, I.R.; OMURA, T.; KIKUCHI, S. Gene Structures, Classification and Expression Models of the AP2/EREBP Transcription Factor Family in Rice. **Plant Cell Physiol.**, v.52, n. 2, p. 344–360. 2011.
- SHARMA, M. K., KUMAR, R.; SOLANKE, A. U.; SHARMA, R.; TYAGI, A. K.; SHARMA, A. K. Identification, phylogeny, and transcript profiling of ERF family genes during development and abiotic stress treatments in tomato. **Mol Genet Genomics**. v.284, n. 6, p. 455–475. 2010.
- SEKI, M.; KAMEI, A.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Molecular responses to drought, salinity and frost: common and different paths for plant protection. **Current Opinion of Biotechnology**, v.14, p.194–199, 2003.

SINGH, K.; FOLEY, R. C.; ONATE-SANCHEZ, L. Transcription factors in plant defense and stress responses. **Curr. Opin. Plant Biol.** v.5, p. 430-436. 2002.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K.; SEKI, M. Regulatory network of gene expression in the drought and cold stress responses. **Curr Opin in Plant Biol.**, v.6, p. 410-417. 2003.

HSHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Global Analysis of Networks to Solve Complex Abiotic Stress Responses. In: CHEN, T. H.; UEMURA, M.; FUJIKAWA, S. (Ed.). **Cold Hardiness in Plants Molecular Genetic, Cell Biology and Physiology**. 7thed. ND-CABI PUBLISHING: CAB International. 2004. p.1-10.

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

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

SHINOZAKI, Y. K.; SHINOZAKI, K. A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress, **Plant Cell**, v.6, p. 251-264. 1994.

SILVA, E. A.; MAZZAFERA, P.; BRUNINI, O.; SAKAI, E.; ARRUDA, F. B.; MATTOSO, L. H. C.; CARVALHO, C. R. L.; PIRES, R. C. M. The influence of water management and environmental conditions on the chemical composition and beverage quality of coffee beans. **Braz. J. Plant. Physio.**, v.17, n.2, p. 229-238. 2005.

SOLANO R.; STEPANOVA A.; CHAO Q.; ECKER J.R.: Nuclear events in ethylene signaling: A transcriptional cascade mediated by ethylene-insensitive3 and ETHYLENERESPONSIVE-FACTOR1. **Genes Dev.**, v.12, p. 3703–3714. 1998.

STEVENS, P. F. **Angiosperm Phylogeny Website**, Version 7. 2006. Disponível em: < <http://www.mobot.org/MOBOT/research/APweb/>>. Acesso em: 27 mar. 2008.

THOMAS, A. S. The wild arabica coffee on the Boma Plateau, Anglo-Egyptian Sudan. **Emp. Exp. Agric.**, v.10, p. 207–212. 1942.

THOMASHOW, M. F. Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. **Ann Rev Plant Physiol Plant Mol Biol.**, v.50, p. 571-599. 1999.

TIWARI, S. B.; HAGEN G.; GUILFOYLE, T. J. Aux/IAA proteins contain a potent transcriptional repression domain. **Plant Cell**, v.16, p. 533–543. 2004.

TOURNIER, B.; SANCHEZ-BALLESTA, M. T.; JONES, B.; PESQUET, E.; REGAD, F.; LATCHE, A.; PECH, J. C.; BOUZAYEN, M. New members of the tomato ERF family show specific expression pattern and diverse DNA-binding capacity to the GCC box element. **FEBS Lett**, v.550, p. 149–154. 2003.

UMEZAWA, T.; FUJITA, M.; FUJITA, Y.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future. **Curr Opin Biotech**, v.17, p.1-10, 2006.

VEGA, F. E.; ROSENQUIST, E.; COLLINS, W. Global project needed to tackle coffee crisis. **Nature**, v.425, p. 343. 2003.

VIEIRA, L. G. E.; ANDRADE, A. C.; COLOMBO, C. A.; MORAES, A. H. A.; METHA A, OLIVEIRA, A. C.; LABATE, C. A.; MARINO, C. L.; MONTEIRO-VITORELLO, C. B.; MONTE, D. C. et al: Brazilian coffee genome project: an EST-based genomic resource. **Brazilian Journal of Plant Physiology**, 2006, v.18, n.1, p. 95-108. 2006.

WANG, A.; TAN, D. M.; TAKAHASHI, A.; LI, T. Z.; HARADA, T. MdERFs, two ethylene-response factors involved in apple fruit ripening. **J Exp. Bot.**, v.58, p. 3743–3748. 2007.

WANG, X.; CHEN, X.; WANG, Z.; NIKOLAY, D.; VLADIMIR, C.; GAO, H. Isolation and characterization of GoDREB encoding an ERF-type protein in forage legume *Galegae orientalis*. **Genes Genet. Sust.**,v.85, p.157-166. 2010.

WEI, G.; PAN, Y.; LEI, J.; ZHU, Y. X. Molecular cloning, phylogenetic analysis, expressional profiling and in vitro studies of TINY2 from *Arabidopsis thaliana*. **J. Biochem. Mol. Biol.**, v.38, p. 440–446. 2005.

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. Isolation and molecular characterization of the *Triticum aestivum* L. ethylene-responsive factor 1 (TaERF1) that increases multiple stress tolerance. **Plant Mol Biol.**, v.65, p. 719–732. 2007.

YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Transcriptional Regulatory Networks in Cellular Responses and Tolerance to Dehydration and Cold Stresses. **Annu Rev Plant Biol**, v.57, p.781-803. 2006.

YANG T. W.; ZHANG L. J.; ZHANG, T. G.; ZHANG, H.; XU, S. J. Transcriptional regulation network of cold-responsive genes in higher plants. **Plant Sci.**, v.169, p. 987–995. 2005.

YI, S. Y.; KIM, J. H.; JOUNG, Y. H.; LEE, S.; KIM, W. T.; YU, S. H.; CHOI, D. The pepper transcription factor CaPF1 confers pathogen and freezing tolerance in *Arabidopsis*. **Plant Physiol.**, v.136, p. 2862–2874. 2004.

YIN, XUE-REN, ALLAN A.C., CHEN KUN-SONG, FERGUSON I.B. Kiwifruit EIL and ERF Genes Involved in Regulating Fruit Ripening. **Plant Physiology**, v.153, p. 1280–1292. 2010.

ZHANG, G.; CHEN, M.; CHEN, X.; XU, Z.; GUAN, S.; LI, L. C.; LI, A.; GUO, J.; MAO, L, MA Y. Phylogeny, gene structures, and expression patterns of the ERF gene family in soybean (*Glycine max* L). **J Exp Bot.**, v.59, p. 4095–4107. 2008.

ZHANG, G.; CHEN, M.; CHEN, X.; XU, Z.; LI, L.; GUO, J.; MA, Y. I. Isolation and characterization of a novel EAR-motif-containing gene GmERF4 from soybean (*Glycine max* L.) **Mol Biol. Rep.**,v.37, p. 809–818. 2010.

ZHANG, G.; CHEN, M.; LI, L.; XU, Z.; CHEN, X.; GUO, J.; MA, Y. Overexpression of the soybean GmERF3 gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. **Journal of Experimental Botany**, v.60, n.13, p. 3781–3796, 2009.

ZHANG, H.; HUANG, Z.; XIE, B.; CHEN, Q.; TIAN, X.; ZHANG, X.; ZHANG, H.; LU, X.; HUANG, D.; HUANG, R. The ethylene-, jasmonate-, abscisic acid- and NaCl-responsive tomato transcription factor JERF1 modulates expression of GCC box-containing genes and salt tolerance in tobacco. **Planta**, v.220, p. 262-270. 2004.

ZHANG, X.; ZHANG, Z.; CHEN, J.; CHEN, Q.; WANG, X. C.; HUANG, R.; Expressing TERF1 in tobacco enhances drought tolerance and abscisic acid sensitivity during seedling development. **Planta**, v.222, p. 494-501. 2005.

ZHU, J-K. Salt and drought stress signal transduction 10 plants. **Annual Review of Plant Biology**., v.53, p. 247-273. 2002.

ZHUANG, J.; CAI, B.; PENG, R. H.; ZHU, B.; JIN, X. F.; XUE, Y.; GAO, F.; FU, X. Y, TIAN, Y. S.; ZHAO, W.; QIAO, Y. S.; ZHANG, Z.; XIONG, A. S.; YAO, Q. H. Genome-wide analysis of the AP2/ERF gene family in *Populus trichocarpa*. **Biochem Biophys Res Commun.**, v.371, p. 468–474. 2008.

ZHUANG, J.; CHEN, J. M.; YAO, Q. H.; XIONG, F.; SUN, C. C, ZHOU, X. R.; ZHANG, J. A.; XIONG A. S. Discovery and expression profile analysis of AP2/ERF family genes from *Triticum aestivum*. **Mol. Biol. Rep.**, v.38, p. 745–753. 2011.

ZHUANG, J.; DENG, D. X.; YAO, Q. H.; ZHANG, J., XIONG, F.; CHEN, J. M.; XIONG, A. S. Discovery, phylogeny and expression patterns of AP2-like genes in maize. **Plant Growth Regul.**, v.62, p. 51–58. 2010.

ZHUANG, J.; PENG, R. H.; CHENG, Z. M.; ZHANG, J.; CAI, B.; ZHANG, Z.; GAO, F.; ZHU, B.; FU, X. Y.; JIN, X. F.; CHEN, J. M.; QIAO, Y. S.; XIONG, A. S.; YAO, Q. H. Genome-wide analysis of the putative AP2/ERF family genes in *Vitis vinifera*. **Sci Hortic.**, v.123, p. 73–81. 2009.

ZHUANG, J.; YAO, Q. H.; XIONG, A. S.; ZHANG, J. Isolation, Phylogeny and Expression Patterns of AP2-Like Genes in Apple (*Malus × domestica* Borkh). **Plant Mol. Biol. Rep.**, v.29, p. 209–216. 2011.

4 ARTIGO A

**ANÁLISE *IN SILICO* DA FAMÍLIA ERF EM *Coffea arabica*: IDENTIFICAÇÃO E
PADRÃO DE EXPRESSÃO**

ANÁLISE *IN SILICO* DA FAMÍLIA ERF EM *Coffea arabica*: IDENTIFICAÇÃO E PADRÃO DE EXPRESSÃO

4.1 RESUMO

Os membros da família de fatores de transcrição ERF possuem um papel importante no desenvolvimento das plantas e na expressão dos genes que regulam a resposta a estresses bióticos e abióticos. Neste trabalho foram identificados 36 ERFs em *Coffea arabica* com o domínio AP2/ERF completo utilizando o banco de ESTs do Projeto Genoma Café Brasileiro. Os resíduos conservados de aminoácidos e a reconstrução filogenética utilizando o domínio conservado AP2 sugeriu que a distribuição em nove dos 10 grupos principais utilizados em *Arabidopsis* é aplicável à família ERF em *C. arabica*. Além do domínio AP2, a análise dos motivos MEME resultou na identificação de domínios conservados, característicos dos membros de cada grupo. A análise *in silico* do perfil de expressão por *northern eletrônico* mostrou que, níveis elevados de expressão foram observados em bibliotecas derivadas de tecidos de frutos, folhas e flores, assim como em bibliotecas submetidas à estresse hídrico. Esses resultados sugerem o envolvimento dos ERFs de *C. arabica* em distintas funções biológicas mediada por etileno, tais como controle de desenvolvimento, maturação e respostas ao estresse hídrico. Os resultados desse trabalho será útil na seleção de genes candidatos para futuras caracterizações funcionais que ajudarão na compreensão das complexas redes reguladoras durante o desenvolvimento e respostas aos estresses, abrindo oportunidades para aplicações no melhoramento do cafeeiro.

Palavras-chave: AP2/ERF. Café. Etileno. Fator de transcrição.

4.2 ABSTRACT

The members of the ERF transcription factors family play an important role in plant development and on the expression of the genes that regulates the responses to biotic and abiotic stresses. In this research work, 36 ERF that contained a complete AP2/ERF domain were identified in *Coffea arabica* using the bank of ESTs of the Brazilian Coffee Genome Project. The conserved amino acids residues and the phylogenetic reconstruction using the conserved domain AP2 suggested that the distribution in nine of the 10 main groups utilized in *Arabidopsis* is applicable to the ERF family in *Coffea arabica*. Besides the AP2 domain, the MEME detection motifs analysis resulted in the identification of conserved domains, which are characteristic of the members of each group. The *in silico* analysis of the expression profile using the northern electronic technique has shown that high levels of expression were observed in libraries derived from fruits, leaves and flowers tissues as well as in libraries subjected to drought stress. These results suggest the involvement of the *C. arabica* ERFs in distinct biological functions mediated by ethylene such as control of development, maturation and response to drought stress. Results of this work will be useful in selecting genes for future functional characterization that will aid the understanding of the complex regulatory network during the development and responses to the stresses, creating opportunities for applications in coffee breeding programs.

Keywords: AP2/ERF. Coffee. Ethylene. Transcription factor.

4.3 INTRODUÇÃO

A cafeicultura atualmente é difundida em mais de 80 países distribuídos em quatro continentes. O *C. arabica*, um dos maiores commodities agrícolas do mundo, representa uma importante fonte de renda para os países tropicais em desenvolvimento (PAY, 2009). O Brasil é o maior produtor mundial e junto com a Colômbia e o Vietnã são responsáveis por mais de 50% da produção mundial (MAPA, 2011). Entre as 103 espécies identificadas, somente duas são importantes economicamente: *Coffea arabica* L. e *Coffea canephora* P. (CHARRIER; BERTHAUD, 1985) que representam cerca de 70% e 30% do mercado total de café, respectivamente.

Durante o ciclo de vida, as plantas são expostas a diferentes estresses bióticos e abióticos, que acabam limitando o crescimento, o desenvolvimento e a produtividade da cultura (WANG et al, 2003). Para sobreviver em tais condições, as plantas desenvolveram uma complexa rede de sinalização em nível molecular, celular e bioquímico (FUJITA et al, 2006). A regulação gênica em nível de transcrição é um dos principais pontos no controle dos processos biológicos e os fatores de transcrição (TFs) desempenham um papel fundamental nesse processo (SHARMA et al, 2010).

Os fatores de transcrição pertencentes a família *Ethylene Responsive Factor* (ERF) fazem parte da superfamília AP2/ERF. Assim como as famílias AP2 e RAV, são caracterizadas pela presença do domínio de ligação ao DNA AP2, caracterizado por 60-70 aminoácidos implicados na ligação ao DNA (RIECHMANN et al, 2000). As proteínas da família ERF contém um único domínio AP2, enquanto que as proteínas da família AP2 contém dois domínios repetidos AP2 (JOFUKU et al, 1994). As proteínas da família RAV, além do domínio AP2, contém um domínio B3, que é um domínio de ligação ao DNA (KAGAYA et al, 1999; NAKANO et al, 2006). A família ERF por sua vez, apresenta uma subdivisão em duas subfamílias, ERF e CBF/DREB (JOFUKU et al, 1994; SAKUMA et al, 2002; NAKANO et al, 2006).

De modo geral, os genes da subfamília ERF estão envolvidos principalmente em respostas a estresses bióticos, reconhecendo o elemento *cis-acting* AGCCGCC denominado GCC-box (HAO et al, 1998). Genes da subfamília CBF/DREB desempenham um papel crucial na resposta das plantas a estresses

abióticos, reconhecendo o elemento responsivo a desidratação (DRE) com um motivo central A/GCCGAC (SHINOZAKI; SHINOZAKI, 1994; THOMASHOW, 1999). Os papéis das proteínas ERF e CBF/DREB no desenvolvimento e na resposta a estresses bióticos e abióticos em diferentes espécies têm sido amplamente estudados. Combinando abordagens genético-moleculares, uma série de genes reguladores da família ERF envolvidos em diferentes vias metabólicas tem sido analisados, incluindo aqueles relacionados à seca (YAMAGUCHI; SHINOZAKI, 2006, DUBOUZET et al, 2003), a salinidade (DUBOUZET et al, 2003, FU et al, 2007), a baixa temperatura (YANG et al, 2005; QIN et al, 2007), a doenças (GUTTERSON; REUBER, 2004). Além da resposta a estresses abióticos, a família ERF também está envolvida na formação de raízes (BANNO et al, 2001), germinação (PIRRELLO et al, 2006) e desenvolvimento de frutos e maturação (WANG et al, 2007; EL-SHARKAWY et al, 2009).

Fatores de transcrição das subfamílias ERFs e CBF/DREB foram identificados em várias espécies vegetais, incluindo *Arabidopsis* (LIU et al, 1998; SAKUMA et al, 2002; NAKANO et al, 2006), arroz (NAKANO et al, 2006; SHARONE et al, 2011), algodão (JIN; LIU, 2008; CHAMPION et al, 2009), soja (ZHANG et al, 2008), *Populus* (ZHUANG et al, 2008), uva (ZHUANG et al, 2009; LICAUSI et al, 2010), milho (ZHUANG et al, 2010), tomate (SHARMA et al, 2010), trigo (ZHUANG et al, 2011a), e maçã (ZHUANG et al, 2011b). Entretanto, poucos trabalhos com objetivo de caracterizar os membros da família ERF em *Coffea* ssp foram realizados. Bustamante-Porras et al, (2005) isolou o primeiro ERF em *C. canephora* cuja expressão está envolvida nos processos de diferenciação celular e na maturação de frutos. Em *C. arabica* nenhum membro dessa família foi descrito até o momento.

A pesquisa sobre genômica e transcriptômica em cafeeiros tem ganhado cada vez mais destaque. O Projeto Genoma Café Brasileiro (VIEIRA et al, 2006), foi desenvolvido para investigar as características de cafeeiros através do seqüenciamento de cDNAs. Esse banco de dados possui um conjunto de 265.889 seqüências expressas (ESTs) provenientes de diversos tecidos, em diferentes etapas de desenvolvimento e diversas condições ambientais derivados das espécies *C. arabica*, *C. canephora* e *C. racemosa*. Portanto, esse trabalho tem como objetivo a identificação e caracterização dos possíveis fatores de transcrição ERF presentes em *Coffea arabica* a partir do banco de dados de ESTs do Genoma Café Brasileiro.

4.4 MATERIAL E MÉTODOS

Identificação de genes da família ERF em *Coffea arabica*

Para isolar os genes da família ERF em *Coffea arabica*, o domínio AP2 do ERF4 (GENEBANK: AAO34706) de *Solanum lycopersicum* foi utilizado para pesquisas no Banco de Dados Genoma Café Brasileiro (www.lge.ibi.unicamp.br./café), usando o programa BlastP (ALTSCHUL et al, 1997). Nesse banco de dados estão disponibilizados mais de 265.889 sequências ESTs obtidos de quarenta e três bibliotecas de cDNA, principalmente de *C. arabica*. O cDNA foi obtido a partir de diferentes tecidos do cafeeiro (folhas, raízes, flores, sementes, frutos, entre outros) em diferentes estádios de desenvolvimento e submetidas a diversas condições de estresse (VIEIRA et al, 2006; MONDEGO et al, 2011). Buscas também foram realizadas utilizando as palavras chaves: ERF, Fator de Resposta ao Etileno e EREBP, visando aumentar as chances de identificação de novas ESTs. As sequências foram confrontadas com outras sequências depositadas no banco de dados do GenBank, utilizando os programas BlastP e BlastX (National Center for Biotechnology Information - NCBI - <http://www.ncbi.nlm.nih.gov>), para confirmar a identidade das mesmas. A sequência deduzida de aminoácidos de cada *contig* formado foi obtida através do programa ORF Finder (Open Reading Frame Finder - NCBI - <http://www.ncbi.nlm.nih.gov/gorf/gorf.html>). As sequências inicialmente coletadas, que apresentaram o domínio AP2 incompleto ou que os ORFs se mostraram incorretos, foram excluídos da análise. As sequências proteicas contendo o domínio AP2 foram alinhadas utilizando o algoritmo ClustalW versão 2.0 (LARKIN et al, 2007) e as entradas redundantes foram removidas.

Análise filogenética

A análise filogenética foi realizada alinhando os domínios AP2 de *C. arabica* e *A. thaliana* usando o algoritmo ClustalW versão 2.0 (LARKIN et al, 2007). A árvore filogenética foi construída utilizando-se o método neighbor-joining (NJ) usando a opção *pair-wise deletion* com a ajuda do programa MEGA versão 4.0 (TAMURA et al, 2007). Para testar a confiabilidade das análises foram utilizadas 1.000 réplicas de *bootstrap*.

Determinação dos motivos conservados

A identificação dos motivos conservados fora do domínio altamente conservado AP2 foi realizado em seqüências proteicas de *C. arabica* usando o algoritmo ClustalW versão 2.0 (LARKIN et al, 2007) e MEME Suite versão 4.0.0 (BAILEY; ELKAN, 1995) (<http://meme.sdsc.edu/meme4/cgi-bin/meme.cgi>), com os seguintes parâmetros: tamanho ideal: 6-80 aminoácidos; qualquer número de repetições dos motivos e número máximo de 25 motivos. Os motivos resultantes foram verificados nas bases de dados do NCBI (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) e ExPaSy - Expert Protein Analysis System (<http://ca.expasy.org/>) para verificar sua significância.

Padrão de expressão por northern eletrônico

A frequência de *reads* de cada biblioteca foi normalizada dividindo-se o número de *reads* encontrado pelo número total de *reads* seqüenciados disponíveis no Banco de Dados e o resultado multiplicado por 10.000. O Northern Eletrônico foi realizado utilizando-se o Programa Genesis versão 1.7.5.

4.5 RESULTADOS

Identificação de genes da família ERF e relação filogenética entre a família ERF em *C. arabica* e *Arabidopsis*

A análise *in silico* a fim de identificar os membros da família ERF em *C. arabica*, compreendeu o processo de mineração de dados dentro do banco de dados do Projeto Genoma Café no qual foram selecionadas as sequências (*reads*) que possivelmente codificam o domínio AP2 em *Coffea arabica*. A procura por palavras chave e a busca através do domínio AP2 das ESTs resultou na identificação de 38 fatores de transcrição codificando as proteínas ERFs. Destes 38 possíveis ERFs, verificou-se que 36 proteínas ERFs contém um domínio AP2 completo, enquanto que os outros 2 ERFs possuem somente parte do domínio AP2 e portanto não foram incluídos na análise. A identidade das proteínas ERFs de *C. arabica* com os seus homólogos em *Arabidopsis* variou de 63 a 90% (Tabela 1). O grau de conservação da sequência em comparação com *Arabidopsis* foi maior para os membros do grupo X, onde variou de 75 a 90%, enquanto que os menores valores foram observados para os grupos III (65%) e VI (63%).

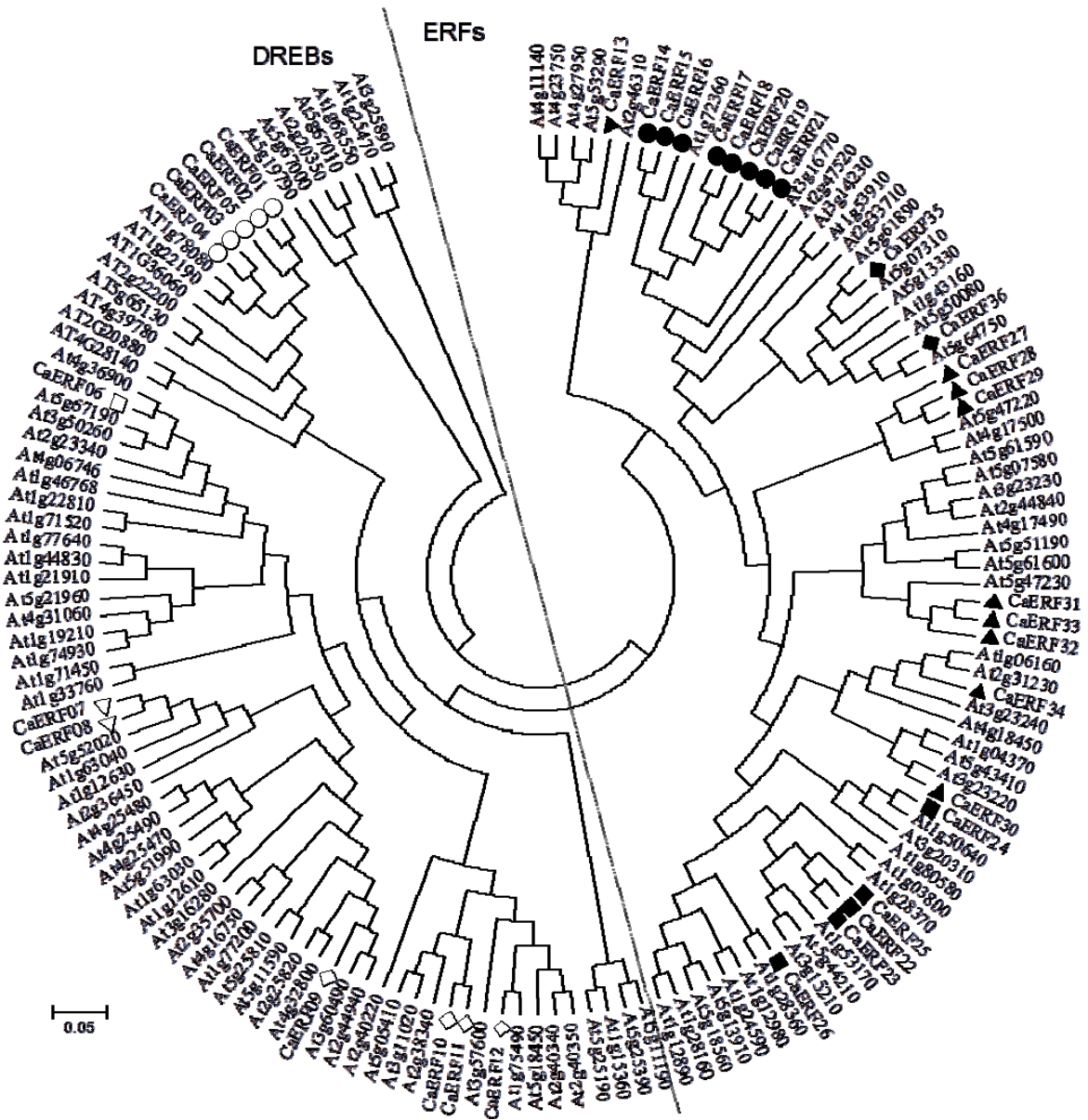
Com a identificação dos 122 ERFs descrito anteriormente por Sakuma et al, (2002) na planta modelo *Arabidopsis*, foi possível realizar uma reconstrução filogenética (Figura 1). A árvore foi construída com o método *neighbor-joining* usando as 58-59 seqüências de aminoácidos do domínio AP2 de *C. arabica* e *Arabidopsis*. As sequencias ERFs são bastante conservadas entre as espécies o que permite distinguir os 10 principais grupos, nomeados como grupos I-X por Nakano et al, (2006). Como mostrado na Figura 1, a análise comparativa da árvore filogenética entre *Arabidopsis* e *C. arabica* agrupou um grande número de seqüências identificadas em *C. arabica* (22,22%, 8 seqüências do total de 36) com as seqüências de *Arabidopsis* pertencentes aos grupos VII e IX. Em *C. arabica*, uma quantidade menor de proteínas foi agrupada com os ERFs de *Arabidopsis* pertencente aos grupos I e VIII (13,89%, 5 sequencias), seguido pelos grupos III e IV (8,33%, 3 seqüências). Os grupos dos ERF com menor número de membros em *C. arabica* foram X, II e VI (5,56%, 2,78% e 2,78%, correspondendo a 2, 1 e 1 seqüências, respectivamente). Nenhum membro de *C. arabica* foi atribuído ao grupo V e aos subgrupos VI-L e Xb-L. Esses dois subgrupos são caracterizados em

Arabidopsis por uma baixa homologia na região C-terminal do domínio AP2/ERF (CHAMPION et al, 2009).

Tabela 1 – Sequências de *Coffea arabica* com homologia com os genes da família ERF em *Arabidopsis thaliana*.

<i>Arabidopsis thaliana</i>		<i>Coffea arabica</i>			
Nome	Gene	ERF	Cobertura (%)	e-value	Identidade (%)
AT1G78080	RAP2.4	CaERF01	99	1,00E-68	81
AT1G78080	RAP2.4	CaERF02	97	2,00E-63	81
AT1G78080	RAP2.4	CaERF03	100	2,00E-58	80
AT1G78080	RAP2.4	CaERF04	100	5,00E-62	80
AT1G78080	RAP2.4	CaERF05	100	2,00E-34	81
AT5G67190	AtERF10	CaERF06	100	3,00E-39	74
AT5G52020	AtERF25	CaERF07	63	8,00E-30	71
AT5G52020	AtERF25	CaERF08	85	6,00E-45	65
AT244940	AtERF34	CaERF09	65	6,00E-48	77
AT240340	DREB2C	CaERF10	84	9,00E-49	68
AT240340	DREB2C	CaERF11	78	8,00E-38	70
AT1G75490	DREB2D	CaERF12	87	3,00E-37	69
AT4G27950	CRF4	CaERF13	71	3,00E-30	63
AT3G16770	ATEBP/RAP2.3	CaERF14	96	8,00E-34	72
AT3G16770	ATEBP/RAP2.3	CaERF15	93	2,00E-29	73
AT3G16770	ATEBP/RAP2.3	CaERF16	96	2,00E-29	68
AT3G16770	ATEBP/RAP2.3	CaERF17	62	1,00E-28	81
AT3G16770	ATEBP/RAP2.3	CaERF18	90	6,00E-27	81
AT3G14230	RAP2.2	CaERF19	99	8,00E-33	77
AT3G16770	ATEBP/RAP2.3	CaERF20	85	8,00E-28	77
AT3G16770	ATEBP/RAP2.3	CaERF21	96	1,00E-27	84
AT3G15210	AtERF4/RAP2.5	CaERF22	42	3,00E-30	74
AT3G15210	AtERF4/RAP2.5	CaERF23	41	5,00E-30	74
AT1G50640	AtERF3	CaERF24	82	2,00E-42	76
AT5G44210	ATERF9	CaERF25	97	4,00E-30	76
AT1G28360	ATERF12	CaERF26	68	1,00E-26	85
AT4G17500	AtERF1	CaERF27	75	9,00E-53	75
AT4G17500	AtERF1	CaERF28	86	1,00E-44	73
AT4G17500	AtERF1	CaERF29	79	5,00E-52	74
AT3G23240	ERF1	CaERF30	82	4,00E-31	73
AT4G17490	AtERF-6	CaERF31	97	5,00E-42	77
AT4G17490	AtERF-6	CaERF32	84	4,00E-31	78
AT4G17490	AtERF-6	CaERF33	93	7,00E-34	74
AT3G23240	ERF1	CaERF34	54	2,00E-15	73
AT5G61890	ABR1	CaERF35	94	2,00E-33	90
At2G33710	AtERF112	CaERF36	49	4,00E-28	75

Figura 1 – Árvore filogenética dos ERFs de *C. arabica* e *arabidopsis*. As seqüências de aminoácidos do domínio AP2 de 36 genes ERF de *C. arabica* e *arabidopsis* foram alinhadas pelo Clustal W e a árvore filogenética foi construída com MEGA 4.0 utilizando o método NJ. (Subfamília DREB: ○ - grupo I, □ - grupo II, ▽ - grupo III, ◇ - grupo IV; Subfamília ERF: ▼ - grupo VI, ● - grupo VII, ■ - grupo 8, ▲ - grupo IX, ◆ - grupo X)



Com base na similaridade das seqüências de aminoácidos os ERFs também foram classificados em duas subfamílias. Doze putativos ERFs foram identificados como sendo membros da subfamília DREB (Grupos I, II, III e IV), em comparação com 57, 40, 66 e 58 em arabidopsis, uva, álamo e arroz, respectivamente. Foram identificados 24 genes codificando para a subfamília ERF (Grupos VI, VII, VIII, IX, X), comparando com 65, 82, 103 e 87 em arabidopsis, uva, álamo e arroz, respectivamente. A organização dos genes da família ERF em *C. arabica* é mostrada na Tabela 2 juntamente com a distribuição comparativa entre arabidopsis, uva, álamo e arroz.

Tabela 2 – Número de genes presentes em cada grupo da família ERF em *C. arabica* e nas espécies cujos genomas foram completamente sequenciados, e tamanho dos respectivos genomas.

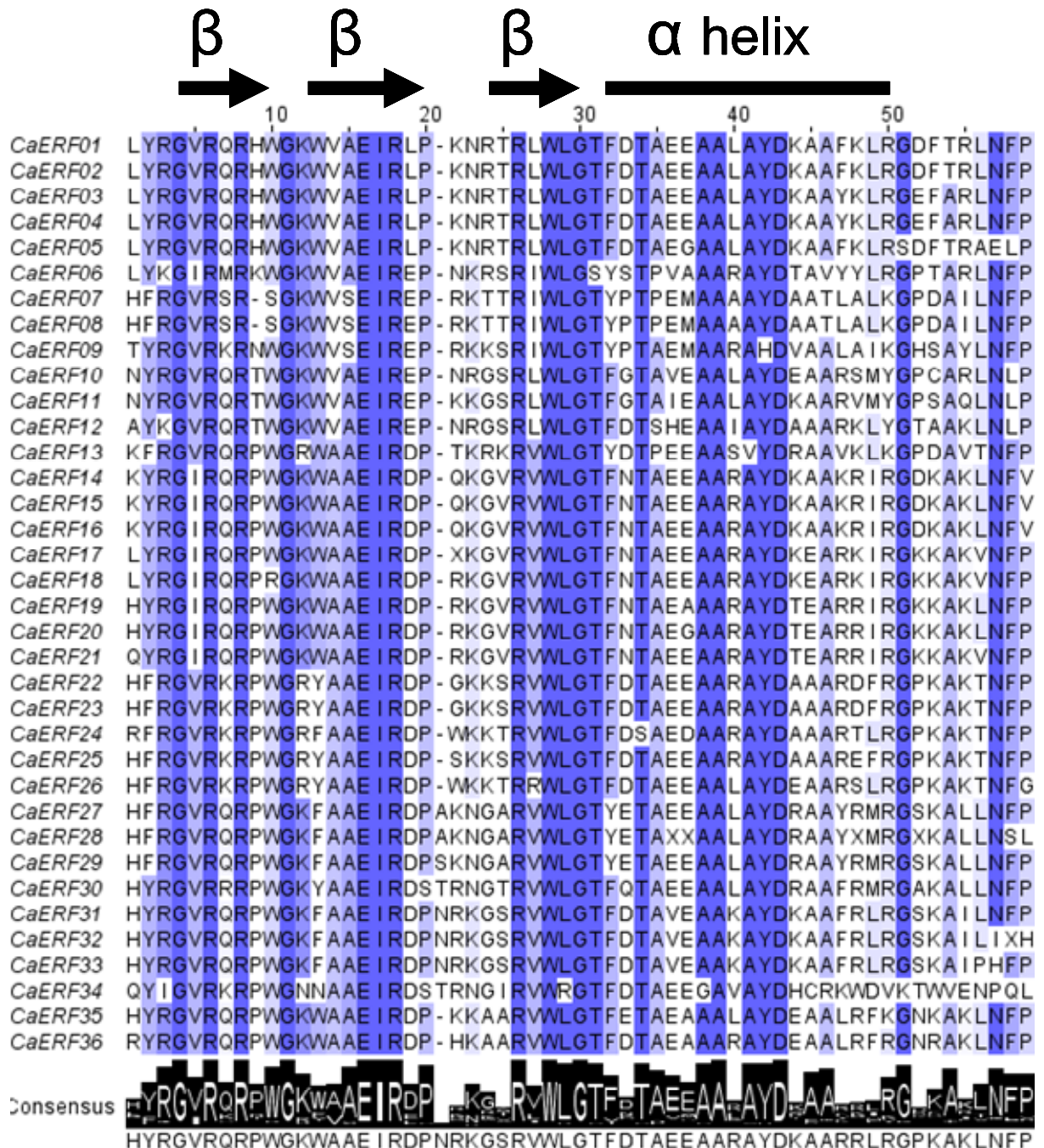
Família	Subfamília	Grupo	Arabidopsis ^a	Uva ^b	Álamo ^c	Arroz ^a	Café
ERF	DREB	I	10	5	5	9	5
		II	15	8	20	16	1
		III	23	22	35	27	3
		IV	9	5	6	6	3
	ERF	V	5	11	10	8	0
		VI	8	2	11	6	1
		VI-L	4	5	4	10	0
		VII	5	3	6	15	8
		VIII	15	11	17	15	5
		IX	17	40	42	18	8
		X	8	10	9	12	2
		Xb-L	3	0	4	3	0
	Total de genes ERFs			122	122	169	145
Tamanho do genoma (Mb)			125	487	465	430	1.300

^aNAKANO et al, (2006), ^bLICAUSI et al, (2010) e ^cZHUANG et al, (2008).

A fim de se estudar as relações filogenéticas entre os genes da família ERF em *C. arabica* e Arabidopsis, múltiplas análises foram realizadas com o alinhamento das seqüências deduzidas dos aminoácidos do domínio AP2. O alinhamento indicou que os resíduos Gly-4, Arg-6, Glu-16, Trp-28 e Gly-30 estão completamente conservados entre todas as proteínas da família ERF em *C. arabica* e Arabidopsis. Além disso, mais de 95% dos membros da família ERF contem os resíduos conservados Arg-8, Gly-11, Ile-17, Arg-18, Arg-26, Leu-29, Ala-38, Ala-39, Asp-43 e Asn-56. Como demonstrado anteriormente por Sakuma et al, (2002) a subfamília de genes ERF inclui dois resíduos de aminoácidos principais, a alanina

(A) na posição 14 e o aspartato (D) na posição 19, que acredita-se contribuir para uma atividade funcional de ligação à *GCC-box* em muitos ERFs. Por outro lado, a família de genes CBF/DREB possui uma valina (V) e um ácido glutâmico (E) nas posições 14 e 19, respectivamente. Na subfamília DREB, todos os genes apresentaram o resíduo conservado Valina14, já na posição 19 os genes *CaERF01* ao *05* contem uma Leucina (L), enquanto que os genes *CaERFs 06* ao *12* apresentam um ácido glutâmico (E). O domínio AP2 de todos os *CaERFs* apresentaram o resíduo conservado para a subfamília ERF. A região C-terminal do domínio de AP2/ERF da proteína *CaERF34* apresentou baixa homologia na região de consenso com os demais genes (Figura 2). Essa região corresponde à metade terminal da α -hélice (ALLEN et al, 1998), que inclui os resíduos altamente conservados Asp-43 e Asn-57. De modo geral a família ERF apresentou similaridade significativa com o restante do domínio

Figura 2 – Alinhamento dos domínios AP2 a partir das proteínas ERF de *C. arabica*. As colunas de cores, azul escuro e azul claro, indicam resíduos de aminoácidos idênticos e conservados, respectivamente. A barra preta e as setas representam a α -hélice e as regiões β -sheets, respectivamente. *CaERF01* a *12* pertencem a subfamília DREB; *CaERF13* a *36* pertencem a subfamília ERF.



Distribuição dos motifs conservados

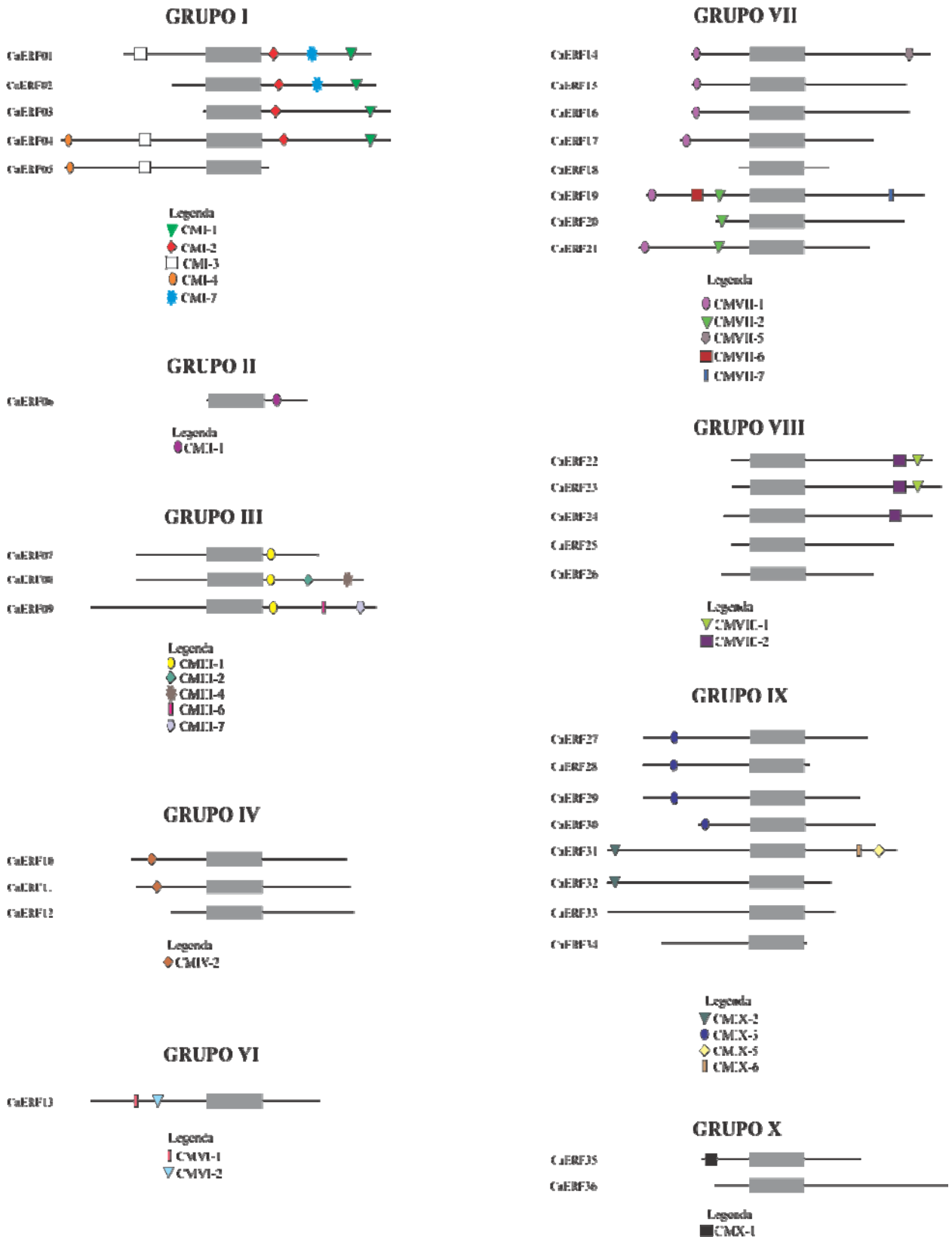
De maneira geral, as regiões fora do domínio de ligação ao DNA nos fatores de transcrição contêm importantes domínios funcionalmente envolvidos na atividade transcricional, como interações proteína-proteína, e podem estar envolvidos na localização nuclear (LIU et al, 1999). Esses domínios funcionais são freqüentemente conservados entre os membros de um subgrupo em grandes famílias de fatores de transcrição de plantas, e provavelmente esses motivos compartilham as mesmas funções (KRANZ et al, 1998; REYES et al, 2004; NAKANO et al, 2006).

A fim de associar os ERFs putativos de *C. arabica* a funções biológicas, outros motivos conservados, fora da região AP2/ERF, foram investigados nas sequências deduzidas de aminoácidos, empregando análises de sequências múltiplas e o conjunto de programas MEME v.4.0 (BAILEY; ELKAN, 1995). A maioria dos membros de um mesmo grupo compartilhou um ou mais motivos fora do domínio AP2 (Figura 3). Por exemplo, o grupo I é composto por 5 ERFs (*CaERF1* ao 5) e contém 5 motifs conservados (Figura 3). Com exceção do *CaERF05*, os demais membros desse grupo contém o motivo CMI-1 e CMI-2 na região C-terminal. O gene *CaERF06* único membro em *C. arabica*, pertencente ao grupo II contém o motivo CMII-1 na região C-terminal adjacente ao domínio AP2. Segundo Nakano et al, (2006) todos os genes desse grupo contém o motivo CMII-1 e, este é semelhante ao motivo CMIII-1 do grupo III. O fato desses motivos estarem presentes em outros relações filogenéticas sugerem uma forte semelhança entre os grupos II e III. No grupo III estão os *CaERF07* ao *09*, onde todos os genes desse grupo contem o motivo CMIII-1 na região C-terminal. Além do motif CMIII-1, o gene *CaERF08* possui os motivos CMIII-2 e CMIII-4, sendo que esse último foi identificado como o motif conservado LWSY em *OsDREB1A/B/C* e *AtCBF3/DREB1A* (DUBOUZET et al, 2003). O gene *CaERF09* apresenta os motivos CMIII-6 e CMIII-7. No grupo IV (*CaERF10* ao *12*) foi encontrado somente o motivo CMIV-2 nos genes *CaERF10* e *CaERF11*. O CMIV-2 possui o putativo sinal de localização nuclear (LIU et al, 1998).

O grupo VI (*CaERF13*) consiste de duas proteínas que compartilham os motivos conservados, CMVI-1 e CMVI-2, na região N-terminal. O grupo VII foi primeiramente descrita por Tournier et al, (2003) e contém 8 membros em *C. arabica*, caracterizada pela presença de um motivo altamente conservada na região

N-terminal (MCGGAIL/L). O motivo EAR (CMVIII-1)(ERF associado à repressão anfifílica) foi encontrado em membros do grupo VIII, nos ERFs *CaERF22* e *CaERF23*, que também contém o motivo CMVIII-2. O grupo IX é constituído por 8 genes (*CaERF 27 ao CaERF34*). O genes *CaERF27* ao *30* apresentam somente o motivo CMIX-3, enquanto que o *CaERF32* possui somente o motivo CMIX-2. Os motivos CMIX-2 e CMIX-3 são regiões putativas ácidas que poderiam funcionar como um domínio de ativação da transcrição (FUJIMOTO et al, 2000). O motivo CMIX-3 corresponde a uma seqüência conservada que anteriormente foi referido como um motivo DMLV (GUTTERSON; REUBER, 2004). Já o gene *CaERF32* além de possuir o motivo CMIX-3, possui os motivos CMIX-5 e CMIX-6 que podem ser um suposto sítio de fosforilação da MAP quinase na região C-terminal da proteína (FUJIMOTO et al, 2000). Os genes desse grupo estão relacionados à expressão de genes de defesa em resposta ao ataque de patógenos. O grupo X é representado por *CaERF35* e *CaERF36*. Os membros desse grupo normalmente contém um motivo conservado CMX-1 na região N-terminal como é o caso do *CaERF35*, enquanto que *CaERF36* não apresenta nenhum motivo conservado.

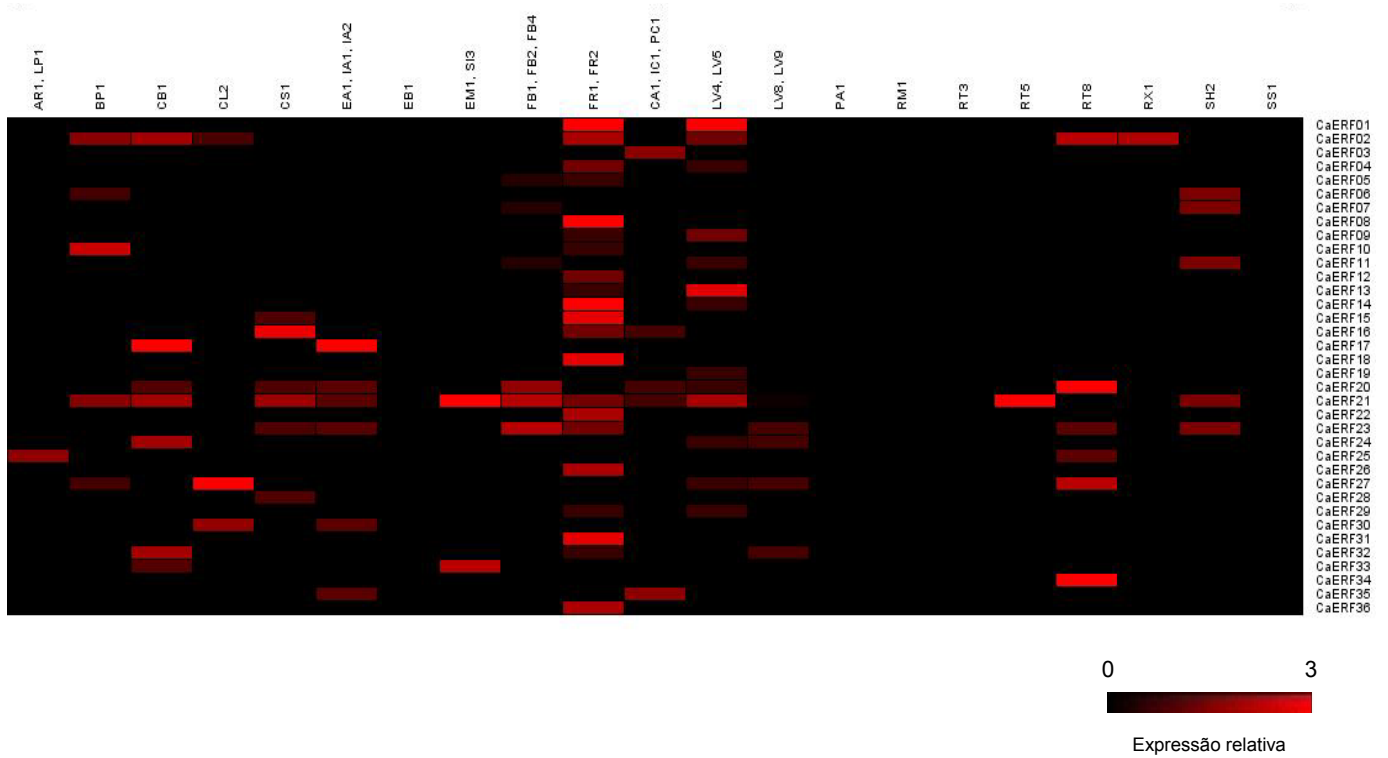
Figura 3 – Distribuição dos motivos conservados na família ERF em *Coffea arabica*. Os Motivos foram identificados em *C. arabica* e classificados de acordo com a classificação proposta por NAKANO et al, (2006). A posição e a ordem dos símbolos correspondem a real região conservada na seqüência da proteína.



Padrão de expressão gênica dos fatores de transcrição da família ERF em *Coffea arabica*

O padrão de expressão dos ERFs em *C. arabica* foi investigado *in silico* por Northern Eletrônico, nas bibliotecas de *C. arabica*, a fim de se avaliar as diferenças entre os transcriptomas dos diferentes tecidos ou órgãos. Níveis elevados de expressão foram observados em bibliotecas derivadas de tecidos de frutos, folhas e flores (Figura 4). Bibliotecas provenientes da análise diferencial de expressão em tecidos submetidos a diferentes estresses também foram detectadas, porém com uma menor frequência de transcritos quando comparados com os tecidos derivados dos tecidos de diversas partes da planta e de diferentes etapas de desenvolvimento. Para os ERFs *CaERF20*, *CaERF21* e *CaERF23* foram detectados transcritos na maioria das bibliotecas avaliadas. No entanto, grande parte dos 36 TF da família ERF foram detectadas em algumas bibliotecas mas não em outras, mostrando ser tecido-específicos. É o caso de *CaERF03*, *CaERF08*, *CaERF12*, *CaERF18*, *CaERF22*, *CaERF26*, *CaERF 31* e *CaERF36* que se expressaram somente em frutos. Já *CaERF07*, *CaERF19* e *CaERF34*, apresentaram expressão somente em flores, folhas e raiz, respectivamente. Expressão em bibliotecas submetidas a estresse hídrico foi observado em *CaERF06*, *CaERF07*, *CaERF11*, *CaERF21* e *CaERF23*, onde os três primeiros ERFs apresentaram uma maior expressão nessa biblioteca.

Figura 4 – Perfil de expressão *in silico* dos ERFs em *Coffea arabica*. O número de *reads* foi normalizado para cada biblioteca, e os valores representados pela escala de expressão relativa.



4.6 DISCUSSÃO

As proteínas dos fatores de transcrição são os principais reguladores dos processos biológicos e surgiram como uma poderosa ferramenta na manipulação de complexas vias metabólicas (GROTEWOLDA, 2008). A fim de utilizar essas proteínas no melhoramento de plantas, é preciso compreender seu papel nas redes reguladoras dos genes. A família de fatores de transcrição ERF apresentam um elemento altamente conservado que inclui um domínio AP2/ERF responsável pela atividade de ligação ao DNA, sendo essenciais para o desenvolvimento das plantas (RIECHMANN; MEYEROWITZ 1998; OHTA et al, 2001; SAKUMA et al, 2002, CAO et al, 2006). Nakano et al, (2006) estudaram sistematicamente a filogenia, as estruturas, e os motivos conservados da família ERF em *Arabidopsis* e arroz. A fim de se obter maiores informações sobre essa família em *C. arabica*, este trabalho identificou e analisou 36 possíveis proteínas ERFs a partir dos dados EST de *C. arabica* disponíveis no Projeto Genoma Café Brasileiro (VIEIRA et al, 2006; MONDEGO et al, 2011).

Os domínios conservados AP2 foram usados para agrupar seus ortólogos em *C. arabica*. A maioria das seqüências de *C. arabica* encontra-se agrupados nos grupos VII e IX, seguido pelos grupos I e VIII, enquanto que, apenas algumas seqüências se agruparam aos grupos III, VI, X, II e VI. No entanto, o grupo VII representa cerca de 4,1% da família em *Arabidopsis* (NAKANO et al, 2006), 2,46% em *Vitis* (LICAUSI et al, 2010), 3,55% em *Populus* (ZHUANG et al, 2008) e 10,34% em arroz (NAKANO et al, 2006). Em *C. arabica*, esse grupo representa mais de 22% das proteínas contendo o domínio único AP2 encontrado no banco de dados do Genoma Café; o que pode indicar um papel biológico significativo do grupo nessa espécie e uma evolução diferencial dessa família de genes em *C. arabica*. Neste trabalho, foi observado que dos 10 principais grupos identificados na família ERF em *Arabidopsis*, 9 grupos ocorreram em *C. arabica*, mostrando que a metodologia utilizada por Nakano et al, (2006) é aplicável nessa espécie. A presença da maioria dos grupos e subgrupos nas 2 espécies dicotiledôneas, bem como nos sistemas de monocotiledôneas sugere que muitos dos genes são anteriores a divergência de monocotiledôneas e dicotiledôneas (SHARMA et al, 2010). Da mesma forma, alguns grupos e subgrupos estão presentes em apenas uma espécie, por exemplo, os grupos XI-XIV ocorrem somente na família ERF em arroz, mas não em *Arabidopsis*.

ou outras espécies dicotiledôneas, sugerindo que esses grupos evoluíram ou foram perdidos em uma espécie após a divergência (ZHANG et al, 2008).

Apesar do grande número de ESTs analisados, o número de ERFs encontrados em *C. arabica* foi baixo comparado com *Arabidopsis* e arroz, que contém 122 e 139 membros na família ERF, respectivamente (NAKANO et al, 2006). Zhang et al, (2008) mostrou que embora a soja tenha um genoma grande (1.115Mb), a estrutura e a filogenia da superfamília AP2/ERF foi semelhante quando o autor comparou com *Arabidopsis* e arroz. O mesmo foi observado em genomas já completos, como *populus* (ZHUANG et al, 2008) e *uva* (ZHUANG et al, 2009). O *Coffea arabica* é uma espécie alotetraplóide ($2n=4x=44$) formada por uma recente (aproximadamente 1 milhão de anos) hibridação natural entre as espécies diplóides: *C. eugenioides* e *C. canephora* (LASHERMES et al, 1999; CLARINDO; CARVALHO, 2009). Como consequência de sua natureza poliplóide, o fato do genoma de *C. arabica* encontrar-se ainda incompleto, e da exclusão das seqüências que continham somente parte do domínio AP2, é difícil de se prever o tamanho da família de genes ERF nessa espécie, mas tende a ser de tamanho similar a de outros organismos vegetais.

A análise estrutural revelou que todas as proteínas ERF possuem Ala-14 e Asp-19 conservados, enquanto que, as proteínas DREB apresentam a Val-14 e Glu/Leu-19. A análise comparativa dos resíduos de aminoácidos do domínio ERF/AP2 em *C. arabica* com as proteínas da família ERF de *Arabidopsis* sugeriu que os domínios AP2/ERF foram bem conservados entre as espécies. Estes aminoácidos conservados provavelmente desempenham um papel importante na família de genes ERF, onde estes podem estar envolvidos em diferentes formas de contato com o DNA. De acordo com Allen et al, (1998), o domínio AP2/ERF reconhece o DNA através dos resíduos conservados arginina e triptofano localizados nas β -sheets. A Ala-37 no domínio ERF desempenha um papel importante na estabilidade do domínio ERF ou na ligação ao DNA com o elemento DRE ou GCC-box (LIU et al, 2006).

Os fatores de transcrição geralmente contêm domínios conservados funcionalmente importantes fora do domínio de ligação ao DNA (De BODT et al, 2003; ARORA et al, 2007). A distribuição de motivos específicos em proteínas pertencentes a grupos específicas da árvore filogenética também foi observado em proteínas ERF de *C. arabica*, no qual demonstram semelhanças estruturais entre as

proteínas dentro do mesmo subgrupo. A grande maioria das sequências ERFs identificadas em *C. arabica* compartilha um ou mais motivos fora do domínio AP2 com os seus homólogos de *Arabidopsis*, como observado em arroz e soja (NAKANO et al, 2006; ZHANG et al, 2008). Por exemplo, Ohta et al, (2001) identificaram um motivo EAR (ERF associado à repressão anfifílica), que tem mostrado funcionar como um domínio de repressão. O motivo EAR de sequência conservada, (L/F)DLN(L/F)xP, identificado neste trabalho como CMVIII-1 encontra-se nas regiões C-terminal do grupo VIII. Esse motivo foi identificado em várias proteínas repressoras incluindo ZAT7, 10, 12, ERF3, AUX/IAA, NIMIN1, HSI2, SUPERMAN (*Arabidopsis*), NRR (arroz), ZFT1 (tabaco) e ZPT2-3 (petúnia), que desempenham diferentes funções que vão desde o desenvolvimento das plantas até a tolerância aos diferentes estresses (TIWARI et al, 2004; HIRATSU et al, 2004; KAZAN, 2006). Dois membros do grupo VIII, AtERF4 e AtERF7, foram bem caracterizados. Ambas as proteínas regulam negativamente o gene PDF1.2, no qual contem GCC-*box* na sua região promotora. AtERF4 demonstrou ser um regulador negativo da expressão de etileno, JA e genes ABA-responsivos (OHTA et al, 2001; MCGRATH et al, 2005). Além disso, foi mostrado que AtERF7 atua como um regulador negativo nas respostas ao ABA (YANG et al, 2005). Recentemente, uma proteína DREB contendo o motivo EAR, DEAR1 demonstrou funcionar como um repressor de proteínas de ligação ao elemento responsivo a desidratação que mediam respostas a estresses bióticos e abióticos (TSUTSUI et al, 2009). Kagale et al, (2010) identificaram 4 proteínas, DEAR2, DEAR3, DEAR4 e DEAR5, estruturalmente semelhantes ao DEAR1, que contém a sequência de DLNxxP um tipo de motivo EAR na região C-terminal. O motivo CMIV-2 na região N-terminal poderia funcionar como um sinal de localização nuclear (NLS) (EL KAYAL et al, 2006). Esse se assemelha ao sinal de transporte nuclear descrito anteriormente por Stockinger et al, (1997) indicando que esses ERFs podem estar envolvidos no tráfego de proteínas. Recentemente, este tem sido apontado ser essencial em *Arabidopsis*, no qual CBF1 se liga ao DNA, assim, é indispensável para a sua atividade transcricional (CANELLA et al, 2009). Um sinal de localização nuclear putativo (KRKRK) foi identificado nas proteínas ERF (VAN RAEMDONCK et al, 2005, EL-SHARKAWY et al, 2009). A análise comparativa dos motivos conservados em *C. arabica* e *Arabidopsis* sugeriu que, as funções da proteína foram conservadas e divergiu durante a evolução da família de genes ERF. Sharma et al, (2010) mostrou que alguns motivos foram específicos em

espermatófitas, enquanto que muitos motivos foram identificados desde plantas inferiores, bactérias, fungos e animais. A presença desses motivos conservados em organismos evolutivamente diferentes indica que eles desempenham um papel funcional importante, enquanto que os motivos específicos às espermatófitas podem ter evoluído mais tarde para atender funções específicas.

Nesse trabalho, 8 ERFs pertencentes aos grupos I, III, VI, VII, VIII, IX e X foram expressos somente em bibliotecas de frutos. Embora os fatores de transcrição ERFs sejam regulados por uma série de estímulos físico-químicos, muitos ERFs são responsivos ao etileno, sugerindo que podem estar envolvidos no processo de maturação em frutos climatérios. Tournier et al, (2003) demonstraram que *SIERF2*, um ERF que se liga a *GCC-box*, desempenha um importante papel durante o processo de maturação em tomate. O mesmo foi observado por Yin et al, (2010) em diferentes ERFs expressos durante o processo de maturação dos frutos de kiwi. Pereira et al, (2005) mostrou que a produção de etileno autocatalítico em frutos de *C. arabica* verde é muito baixa, no entanto, esta aumenta consideravelmente durante a fase inicial de maturação. Essas observações demonstram a natureza climatérica da maturação dos frutos de *C. arabica* e a importância do etileno nesse processo. Bustamante-Porras et al, (2005) isolou em *C. canephora* um gene ERF (*CoERF*) cuja expressão demonstrou estar envolvido nos processos de diferenciação celular e na maturação de frutos. Ao se comparar *CoERF* (GENEBANK: AY522505) com *C. arabica*, este compartilhou 97% de identidade e 98% de similaridade com *CaERF17*. Considerando que *C. canephora* é um dos ancestrais de *C. arabica* (LASHERMES et al, 1999) e a alta identidade encontrada entre os ERFs de *Coffea* ssp, era esperado que *CaERF17* também se expressasse em frutos. Entretanto, não foi encontrado *reads* em bibliotecas de frutos em *CaERF17*, embora *CoERF* esteja envolvido na maturação de frutos. Em alotetraplóides é esperado que genes estejam presentes em duas formas homólogas, altamente semelhantes, mas não idênticas (PETITOT et al, 2008). A redundância de genes pode levar ao silenciamento gênico ou a divergência funcional de genes duplicados (CHEN;Ni 2006). Vidal et al, (2010) encontrou em alguns casos que aparentemente um homólogo de *C. canephora* é recrutado para ser expresso em determinados tecidos, enquanto que os homólogos de *C. eugenioides* é silenciado. Desta forma, as diferenças na expressão em *C. arabica* pode ser atribuído às contribuições dos diferentes subgenomas dos ancestrais *C. canephora*

ou *C. eugenoides*. Esses genes podem ser bons candidatos para futuras caracterizações que ajudarão a compreender o processo de regulação desses genes durante os processos de desenvolvimento e maturação em frutos de *Coffea* ssp.

A maioria dos ERF têm mostrado aumentar a tolerância das plantas contra estresses bióticos e abióticos (PARK et al, 2001; GUO et al, 2004; ZHANG et al, 2009). Neste trabalho os ERFs *CaERF06*, *CaERF07*, *CaERF11*, apresentaram alta expressão em bibliotecas submetidas à estresse hídrico. Esses genes pertencem a subfamília DREB, que desempenha um papel importante na tolerância de plantas a estresses abióticos, reconhecendo o elemento responsivo a desidratação (Dehydration Responsive Element - DRE), o motivo *core* A/GCCGAC (LIU et al, 1998). Estudos tem mostrado que a superexpressão de genes DREB em *Arabidopsis* ativam a expressão de muitos genes relacionados ao estresse melhorando a tolerância à seca, sal e frio (LIU et al. 1998; SAKUMA et al, 2006). Por exemplo, a superexpressão do *AoDREB* de *Asparagus officinalis* L. em *Arabidopsis* transgênicas induziu a expressão dos genes *rd29A* e *COR15A*, o que resultou em maior tolerância das plantas transgênicas à seca e à salinidade (LIU et al, 2010). Desta forma, esse genes podem ser candidatos para futuros estudos que podem auxiliar na compreensão do mecanismo de regulação da família ERF nas respostas das plantas de *C. arabica* aos diferentes estresses.

Trabalhos anteriores sugerem a hipótese de que poderia haver um padrão de expressão grupo-específico. Por exemplo, dos 8 genes pertencentes ao grupo VII, 5 se expressam em frutos, onde os ERFs *CaERF14*, *CaERF15* e *CaERF18* apresentam uma alta expressão relativa em frutos. Esse grupo tem sido particularmente associado à maturação em frutos. *LeERF2* em tomate (TOURNIER et al, 2003), *MdERF1* em maçã (WANG et al, 2007), *PsERF2a* e *PsERF2b* em ameixa (EL-SHARKAWY et al, 2009) e *AdERF10* e *AdERF14* em kiwi (YIN et al, 2010) são proteínas que se expressam durante a maturação nos frutos pertencentes ao grupo VII. A indução do amadurecimento também foi relatado no grupo VIII em ameixa (EL-SHARKAWY et al, 2007) e uva (LICAUSI et al, 2010). Outros trabalhos tem mostrado que os membros do grupo IX apresentam indução da expressão quando são submetidos ao ataque de vários patógenos. A superexpressão constitutiva de *AtERF2* do subgrupo IXa mostrou induzir a expressão gênica dos gene PDF1.2 (BROWN et al, 2003). Da mesma forma, a superexpressão do gene *AtERF1* de *Arabidopsis*, um homólogo próximo a *AtERF2*, conferem resistência a

Botrytis cinerea, *Sclerotinia sclerotiorum* e *Erysiphe orontii* em *Arabidopsis* (GUTTERSON; REUBER, 2004). Recentemente Anderson et al, (2010) demonstrou que a superexpressão do gene MtERF1-1 em raízes de *Medicago truncatula* aumentou a resistência a *Rhizoctonia solani*, assim como *Phytophthora medicaginis*. Desta forma, a presença de distintos padrões de expressão dos ERFs observado em *C. arabica* pela análise *in silico* pode estar associada à distância filogenética entre as seqüências, ou seja, proteínas ERFs relacionadas filogeneticamente apresentaram mais padrões similares de expressão do que as sequências divergentes. Assim, a expressão dos padrões genéticos sugere a existência de algum grau de especialização funcional dos grupos de ERFs investigados, embora seja esperado um alto grau de sobreposição das funções nas grandes famílias de genes em plantas (SOLTIS et al, 2008; SEMON; WOLFE, 2007).

A família de genes ERFs desempenha um papel crucial na regulação do desenvolvimento, bem como nas respostas aos estresses ambientais. Com o genoma do *C. arabica* sendo seqüenciado por um consórcio brasileiro, neste trabalho 36 ERFs foram identificados em *Coffea arabica*, onde 12 ERFs pertencem a subfamília DREB e 24 à subfamília ERF. O padrão de expressão gênica mostrou que níveis elevados de expressão foram observados em bibliotecas derivadas de tecidos de frutos, folhas e flores, assim como em bibliotecas submetidas ao estresse hídrico. Na comparação dos homólogos com outras espécies, cujo genoma foi seqüenciado, juntamente com seus perfis de expressão sugerem o envolvimento dos ERFs de *C. arabica* em distintas funções biológicas mediada por etileno, tais como controle de desenvolvimento, maturação e respostas ao estresse hídrico. *Coffea arabica* é uma planta perene cujo valor comercial é devido à produção de frutos. O conhecimento do papel dos fatores de transcrição ERF nos processos de desenvolvimento e maturação dessa espécie abrirá oportunidades para aplicações no melhoramento aumentando a produção e a qualidade de frutos do cafeeiro.

4.7 AGRADECIMENTOS

O primeiro autor agradece às agências financiadoras, CAPES e CNPq pela concessão das bolsas de doutorado no exterior e no país, respectivamente.

4.8 REFERÊNCIAS

ANDERSON, J. P.; LICHTENZVEIG, J.; GLEASON, C.; OLIVER, R. P.; SINGH, K. B. The B-3 Ethylene Response Factor MtERF1-1 Mediates Resistance to a Subset of Root Pathogens in *Medicago truncatula* without Adversely Affecting Symbiosis with Rhizobia. **Plant Physiology**, v.154, p. 861-873, 2010.

ALLEN, M. D.; YAMASAKI, K.; OHME-TAKAGI, M.; TATENO, M.; SUZUKI, M. 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. **EMBO J**, v.17, p. 5484-5496. 1998.

ALTSCHUL, S. F., MADDEN, T. L., SCHÄFFER, A. A., ZHANG, J., ZHANG, Z., MILLER, W., LIPMAN, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, v.25, p. 3389–3402. 1997.

ARORA, R.; AGARWAL, P.; RAY, S.; SINGH, A. K.; SINGH, V. P.; TYAGI, A. K.; KAPOOR, S. MADS-box gene family in rice: genome-wide identification, organization and expression profiling during reproductive development and stress. **BMC Genomics**, v.8, p. 242, 2007.

BAILEY, T. L.; ELKAN, C. The value of prior knowledge in discovering motifs with MEME. **Proc Int Conf Intell Syst Mol Biol**, v.3, p. 21-29, 1995.

BANNO, H.; IKEDA, Y.; NIU, Q. W.; CHUA, N. H. Overexpression of Arabidopsis ESR1 induces initiation of shoot regeneration. **The Plant Cell**, v.13, p. 2609-2618. 2001.

BROWN, R. L.; KAZAN, K.; McGRATH, K. C.; MACLEAN, D. J.; MANNERS, J. M. A role for the GCC-box in jasmonate-mediated activation of the *PDF1.2* gene of *Arabidopsis*. **Plant Physiology** v.132, p. 1020-1032. 2003.

BUSTAMANTE-PORRAS, J.; NOIROT, M.; CAMPA, C.; HAMON, S.; KOCHKO, A. Isolation and characterization of a *Coffea canephora* ERF-like cDNA. **African Journal of Biotechnology**, v.4, n.2, p. 157-159, 2005.

CANELLA, D.; GILMOUR, S. J.; KUHN, L. A.; THOMASHOW, M. F. DNA binding by the Arabidopsis CBF1 transcription factor requires the PKKP/RAGRxKFXETRHP signature sequence. **Biochim Biophys Acta**, v.1799, p. 454-462. 2010.

- CLARINDO, W. R.; CARVALHO, C. R. Comparison of the *Coffea canephora* and *C. arabica* karyotype based on chromosomal DNA content. **Plant Cell Rep**, v.28, n.1, p. 73-81. 2009.
- CAO, Y.; SONG, F.; GOODMAN, R. M.; ZHENG, Z. Molecular characterization of four rice genes encoding ethylene-responsive transcriptional factors and their expressions in response to biotic and abiotic stress. **Journal of Plant Physiology**, v.163, n.11, p. 1167-1178. 2006.
- CHAMPION, A.; HEBRARD, E.; PARRA, B.; BOURNAUD, C.; MARMEY, P.; TRANCHANT, C.; NICOLE, M. Molecular diversity and gene expression of cotton ERF transcription factors reveal that group IXa members are responsive to jasmonate, ethylene and *Xanthomonas*. **Molecular Plant Pathology**, v.10, n.4, p. 471-485. 2009.
- CHARRIER, A.; BERTHAUD, J. Botanical classification of coffee. In: CLIFFORD, M. N.; WILLSON, K. C. **Coffee: botany, biochemistry, and production of beans and beverage**. Edited by Clifford MN, Willson KC. New York; 1985, p. 13-47.
- CHEN, Z. J.; NI, Z. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. **BioEssays**, v.28, p. 240-252. 2006.
- De BODT, S.; RAES, J.; FLORQUIN, K.; ROMBAUTS, S.; ROUZE, P.; THEISSEN, G.; Van de Peer, Y. Genome-wide structural annotation and evolutionary analysis of the type I MADS-box genes in plants. **J. Mol. Evol.**, v.56, p. 573-586. 2003.
- DUBOUZET, J. G.; SAKUMA, Y.; ITO, Y.; KASUGA, M.; DUBOUZET, E. G.; MIURA, S.; SEKI, M.; SHINOZAKI, K.; YAMAGUCHI, S. K. Os DREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. **Plant J.**, v. 33, p. 751-763. 2003.
- EL-SHARKAWY, I.; SHERIF, S.; MILA, I.; BOUZAYEN, M.; JAYASANKAR, S. Molecular characterization of seven genes encoding ethylene-responsive transcriptional factors during plum fruit development and ripening. **J. Exp. Bot.**, v.60, p. 907-922. 2009.
- EL-SHARKAWY, I.; KIM, W. S.; EI-KEREAMY, A.; JAYASANKAR, S.; SVIRCEV AMD.; BROWN, C. W. Isolation and characterization of four ethylene signal transduction elements in plums (*Prunus salicina* L.). **J. Exp. Bot.**, v.58, p. 3631-3643. 2007.
- EL KAYAL, W.; NAVARRO, M.; MARQUE, G.; KELLER, G.; MARQUE, C. TEULIERES, C. Expression profile of CBF-like transcriptional factor genes from *Eucalyptus* in response to cold. **J. Exp. Bot.**, v.57, p. 2455-2469. 2006.
- FU, X. Y.; ZHANG, Z.; PENG, R. H.; XIONG, A. S.; LIU, J. G.; WU, L. J.; GAO, F.; ZHU, H.; GUO, Z. K.; YAO, Q. H. Isolation and characterization of a novel cDNA encoding ERF/AP2-type transcription factor OsAP25 from *Oryza sativa* L. **Biotechnol Lett.**, v.29, p. 1293-1299. 2007.

- FUJIMOTO S. Y.; OHTA, M.; USUI, A.; SHINSHI, H.; OHME-TAKAGI, M. Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. **Plant Cell.**, v.12, p. 393-404. 2000.
- FUJITA, M.; FUJITA, Y.; NOUTOSHI, Y.; TAKAHASHI, F.; NARUSAKA, Y.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. **Current Opinion in Plant Biology.**, v.9, p. 436-442. 2006.
- GROTEWOLDA, E. Transcription factors for predictive plant metabolic engineering: are we there yet? **Curr. Opin. Biotechnol.**, v.19, p. 138-144, 2008.
- GUO, Z. J.; CHEN, X. J.; WU, X. L.; LING, J. Q.; XU, P. Overexpression of the AP2/EREBP transcription factor OPBP1 enhances disease resistance and salt tolerance in tobacco. **Plant Mol. Biol.**, v.55, p. 607-618. 2004.
- GUTTERSON, N.; REUBER, T. L. Regulation of disease resistance pathways by AP2/ERF transcription factors. **Curr. Opin. Plant Biol.** v.7, p. 465–471.2004.
- HAO, D.; OHME-TAKAGI, M.; SARAI, A. Unique mode of GCC box recognition by the DNA-binding domain of ethylene-responsive element-binding factor (ERF domain) in **plant. J. Biol. Chem.** v.273, p. 26857-26861. 1998.
- HIRATSU, K.; MITSUDA, N.; MATSUI, K.; OHME-TAKAGI, M. Identification of the minimal repression domain of SUPERMAN shows that the DLELRL hexapeptide is both necessary and sufficient for repression of transcription in Arabidopsis. **Biochem Biophys Res Commun.** v.321, p 172-178. 2004.
- JIN, L. G.; LIU, J. Y. Molecular cloning, expression profile and promoter analysis of a novel ethylene responsive transcription factor gene GhERF4 from cotton. **Plant Physiology and Biochemistry.** v. 46, p. 46–53. 2008.
- JOFUKU, K. D.; DEN BOER, B. G.; VAN MONTAGU, M.; OKAMURO, J. K. Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. **Plant Cell**, v.6, p. 1211–1225. 1994.
- KAZAN, K. Negative regulation of defence and stress genes by EAR-motif-containing repressors. **Trends Plant Sci.**, v.11, p. 109-112. 2006.
- KAGAYA, Y.; OHMIYA, K.; HATTORI, T. RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. **Nucleic Acids Res.**, v.27, p. 470-478. 1999.
- KAGALE, S.; LINKS, M. G.; ROZWADOWSKI, K. Genome-Wide Analysis of Ethylene-Responsive Element Binding Factor-Associated Amphiphilic Repression Motif-Containing Transcriptional Regulators in Arabidopsis. **Plant Physiology.** v.152, p. 1109-1134. 2010.

KRANZ, H. D.; DENEKAMP, M.; GRECO, R.; JIN, H.; LEYVA, A.; MEISSNER, R. C., PETRONI, K.; URZAINQUI, A.; BEVAN, M.; MARTIN, C. et al. Towards functional characterisation of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. **Plant J.**, v.16, p. 263-276. 1998.

LARKIN, M. A.; BLACKSHIELDS, G.; BROWN, N. P.; CHENNA, R.; MCGETTIGAN, P. A.; MCWILLIAM, H.; VALENTIN, F.; WALLACE, I. M.; WILM, A.; LOPEZ, R.; THOMPSON J. D.; GIBSON, T. J.; HIGGINS, D. G. Clustal W and Clustal X version 2.0. **Bioinformatic.** v.23, p. 2947-2948. 2007.

LASHERMES, P.; COMBES, M. C.; ROBERT, J.; TROUSLOT, P.; D'HONT, A.; ANTHONY, F.; CHARRIER, A. Molecular characterisation and origin of the *Coffea arabica* L. genome. **Mol. Gen. Genet.**, v.261, n.2, p. 259-266. 1999.

LICAUSI, F.; GIORGI, F. M.; ZENON, S.; OSTI, F.; PEZZOTTI, M.; PERATA, P. Genomic and transcriptomic analysis of the AP2/ERF superfamily in *Vitis vinifera*. **BMC Genomics.** v.11, p. 719. 2010.

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

LIU, Y.; ZHAO, T. J.; LIU, J. M.; LIU, W. Q.; LIU, Q.; YAN, Y. B.; ZHOU, H. M. The conserved Ala37 in the ERF/AP2 domain is essential for binding with the DRE element and the GCC box. **FEBS Lett.**, v.580, p. 1303-1308. 2006.

LIU, Y.; CHEN, H.; ZHUANG, D.; JIANG, D.; LIU, J.; WU, G.; YANG, M.; SHEN, S. Characterization of a DRE-binding transcription factor from asparagus (*Asparagus officinalis* L.) and its overexpression in *Arabidopsis* resulting in salt- and drought-resistant transgenic plants. **Journal of Plant Sciences.**, v.171, p. 12-23. 2010.

LIU, L, WHITE, M. J.; MACRAE, T. H. Transcription factors and their genes in higher plants functional domains, evolution and regulation. **Eur. J. Biochem.**, v.262, p. 247-257. 1999.

MAPA. **Ministério da Agricultura, Pecuária e Abastecimento**, Disponível em: <http://www.agricultura.gov.br/arq_editor/file/3894_relatorio_do_mercado_cafeeiro_-_marco_2011.pdf>. Acesso em: 5 abr. 2011.

MCGRATH, K. C.; DOMBRECHT, B.; MANNERS, J. M.; SCHENK, P. M.; EDGAR, C. I.; MACLEAN, D. J, SCHEIBLE, W. R.; UDVARDI, M. K.; KAZAN, K. Repressor- and activator-type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genome-wide screen of *Arabidopsis* transcription factor gene expression. **Plant Physiol.**, v.139, p. 949–959. 2005.

MONDEGO, J. M. C.; VIDAL, R. O.; CARAZZOLLE, M. F.; TOKUDA, E. K.; PARIZZI, L. P.; COSTA, G. G. L.; PEREIRA, L. F. P.; ANDRADE, A. C.; COLOMBO, C. A.; VIEIRA, L. G. E.; PEREIRA, G. A. G. Brazilian Coffee Genome Project Consortium. An EST-based analysis identifies new genes and reveals distinctive gene expression features of *Coffea arabica* and *Coffea canephora*. *BMC Plant Biology*, v.11, p. 30. 2011.

NAKANO, T.; SUZUKI, K.; FUJIMURA, T.; SHINSHI, H. Genomewide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiology*, v.140, p. 411-432. 2006.

OHTA, M.; MATSUI, K.; HIRATSU, K.; SHINSHI, H.; OHME-TAKAGI, M. Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. *Plant Cell*, v.13, p. 1959–1968. 2001.

PARK, J. M.; PARK, C. J.; LEE, S. B.; HAM, B. K.; SHIN, R.; PAEK, K. H. Overexpression of the tobacco Tsi1 gene encoding an EREBP/ AP2-type transcription factor enhances resistance against pathogen attack and osmotic stress in tobacco. *Plant Cell*, v.13, p. 1035–1046. 2001.

PAY, E: **The market for organic and fair-trade coffee**. FAO Rome 2009.

PEREIRA, L. F. P.; GALVAO, R. M.; KOBAYASHI, A. K.; CAÇÃOII, S. M. B.; ESTEVES, V. L. G. Ethylene production and acc oxidase gene expression during fruit ripening of *Coffea arabica* L. *Braz. J. Plant Physiol.*, v.17, p. 283-289. 2005.

PETITOT, A. S.; LECOULS, A. C.; Fernandez, D. Sub-genomic origin and regulation patterns of a duplicated WRKY gene in the allotetraploid species *Coffea Arabica*. *Tree Genetics & Genomes*, v.4, n.3, p. 379-390. 2008.

PIRRELLO, J.; JAIMES-MIRANDA, F.; SANCHEZ-BALLESTA, M. T.; TOURNIER, B.; KHALIL-AHMAD, Q.; REGAD, F.; LATCHE, A.; PECH, J. C.; BOUZAYEN, M. Sl-ERF2, a tomato ethylene response factor involved in ethylene response and seed germination. *Plant Cell Physiol.*, v.47, n.9, p. 1195-1205. 2006.

QIN, Q. L.; LIU, J. G.; ZHANG, Z.; PENG, R. H.; XIONG, A. S.; YAO, Q. H.; CHEN, J. M. Isolation, optimization, and functional analysis of the cDNA encoding transcription factor RdreB1 in *Oryza Sativa* L. *Mol. Breed.*, v.19, p. 329-340. 2007.

REYES, J. C.; MURO-PASTOR M. I.; FLORENCIO, F. J. The GATA family of transcription factors in *Arabidopsis* and rice. *Plant Physiol.*, v.134, p. 1718-1732. 2004.

RIECHMANN J. L.; MEYEROWITZ, E. M. The AP2/EREBP family of plant transcription factors. *Biol. Chem.*, v.379, p. 633–646. 1998.

RIECHMANN, J. L.; HEARD, J.; MARTIN, G.; REUBER, L.; JIANG, C.; KEDDIE, J.; ADAM, L.; PINEDA, O.; RATCLIFFE, O. J.; SAMAHA, R. R, et al. *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. *Science* v.290, p. 2105-2110. 2000.

SAKUMA Y.; LIU Q.; DUBOUZET J. G.; ABE H.; SHINOZAKI, K. DNA binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold inducible gene expression. **Biochem Biophys Res Commun.**, v.290, p. 998-1009. 2002.

SÉMON M.; WOLFE, K. H. Consequences of genome duplication. **Curr. Opin. Gen. De.**, v. 17, p. 505-512. 2007.

SHARONI, A. M.; NURUZZAMAN, M.; SATOH, K.; SHIMIZU, T.; KONDOH, H, SASAYA, T.; CHOI, IR.; OMURA, T.; KIKUCHI, S. Gene Structures, Classification and Expression Models of the AP2/EREBP Transcription Factor Family in Rice **Plant Cell Physiol.**, v.52, n.2, p. 344–360. 2011.

SOLTIS, D. E.; BELL, C. D.; KIM, S.; SOLTIS, P. S. Origin and early evolution of angiosperms. **Ann. New York Acad. Sci.**, v.1133, p. 3-25. 2008.

SHARMA, M. K.; KUMAR, R.; SOLANKE, A. U.; SHARMA, R.; TYAGI, A. K.; SHARMA A. K. Identification, phylogeny, and transcript profiling of ERF family genes during development and abiotic stress treatments in tomato. **Mol. Genet. Genomics**, 2010.

SHINOZAKI, K. Y.; SHINOZAKI, K. A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress. **The Plant Cell.**, v.6, p. 251-264. 1994.

STOCKINGER, E. J.; GILMOUR, S. J.; THOMASHOW, M. F. 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. **Proc. Natl. Acad. Sci. USA.**, v.94, p. 1035-1040. 1997.

TAMURA, K.; DUDLEY, J, NEI M, KUMAR S. MEGA4. Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. **Mol. Biol. Evol.**, v.24, p.1596-1599. 2007.

THOMASHOW, M. F. PLANT COLD ACCLIMATION: freezing tolerance genes and regulatory mechanism. **Annual Review of Plant Physiology and Plant Molecular Biology**, v.50, p. 571–599. 1999.

TIWARI, S. B.; HAGEN, G.; GUILFOYLE, T. J. Aux/IAA proteins contain a potent transcriptional repression domain. **Plant Cell**, v.16, p. 533-543. 2004.

TOURNIER, B.; SANCHEZ-BALLESTA, M. T.; JONES, B.; PESQUET, E.; REGAD, F.; LATCHE, A.; PECH, J. C.; BOUZAYEN, M. New members of the tomato ERF family show specific expression pattern and diverse DNA-binding capacity to the GCC box element. **FEBS Lett**, v.550, p.149-154. 2003.

- TSUTSUI, T.; KATO, W.; ASADA, Y.; SAKO, K.; SATO, T.; SONODA, Y.; KIDOKORO, S.; YAMAGUCHI-SHINOZAKI, K.; TAMAOKI, M.; ARAKAWA, K.; ICHIKAWA, T.; NAKAZAWA, M.; SEKI, M.; SHINOZAKI, K.; MATSUI, M.; IKEDA, A.; YAMAGUCHI, J. DEAR1, a transcriptional repressor of DREB protein that mediates plant defense and freezing stress responses in Arabidopsis. **J. Plant. Res.**, v.122, p. 633-643. 2009.
- VIDAL, R. O.; MONDEGO, J. M.; POT, D.; AMBROSIO, A. B.; ANDRADE, A. C.; PEREIRA, L. F.; COLOMBO, C. A.; VIEIRA, L. G.; CARAZZOLLE, M. F.; PEREIRA, G. A. A high-throughput data mining of SNPs in Coffea spp ESTs suggests differential homeologous gene expression in the allotetraploid Coffea arabica. **Plant Physiology**, v.154, n.3, p.1053-1066. 2010.
- VAN RAEMDONCK, D.; PESQUET, E.; CLOQUET, S.; BEECKMAN, H.; BOERJAN, W.; GOFFNER, D.; EL, JAZIRI, M.; BAUCHER, M. Molecular changes associated with the setting up of secondary growth in aspen. **J. Exp. Bot.**, v.56, p. 2211–2227. 2005.
- VIEIRA, L. G. E.; ANDRADE, A. C.; COLOMBO, C. A.; MORAES, A. H. A.; METHA, A.; OLIVEIRA A. C.; LABATE, C. A.; MARINO, C. L.; MONTEIRO-VITORELLO, C. B.,; MONTE, D. C. et al: Brazilian coffee genome project: an EST-based genomic resource. **Brazil J Plant Physiol.** v.18, p. 95-108. 2006.
- WANG, W.; VINOCUR, B.; ALTMAN, A. Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. **Planta**. v.218, p. 1-14. 2003.
- WANG, A.; TAN, D.; TAKAHASHI, A.; ZHONG Li, T.; HARADA, T.; 2007. MDERFS, two ethylenesresponse factors involved in apple fruit ripening. **J. Exp. Bot.** v.58, n.13, p. 3743–3748. 2007.
- WANG, A.; TAN D. M.; TAKAHASHI, A.; LI, T. Z.; HARADA, T. MdERFs, two ethylene-response factors involved in apple fruit ripening. **J. Exp. Bot.**, v.58, p 3743-3748. 2007.
- YAMAGUCHI, S. K.; SHINOZAKI, K. Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. **Annu. Rev. Plant Biol.** v.57, p. 781-803. 2006.
- YANG, T. W.; ZHANG, L. J.; ZHANG, T. G.; ZHANG, H.; XU, S. J.; AN, L. Z. Transcriptional regulation network of cold-responsive genes in higher plants **Plant Science**, v.169, p. 987-995. 2005.
- YIN, X-R.; ALLAN, A. C.; CHEN, K-S.; FERGUSON, I. B. Kiwifruit EIL and ERF Genes Involved in Regulating Fruit Ripening. **Plant Physiology**, v.153, n.3, p. 1280-1292. 2010.
- ZHANG, G.; CHEN, M.; CHEN, X.; XU, Z.; GUAN, S, Li LC, Li A, GUO J, MAO L, MA Y. Phylogeny, gene structures, and expression patterns of the ERF gene family in soybean (*Glycine max L*). **J. Exp. Bot.** v.59, p. 4095-4107. 2008.

ZHANG, G.; CHEN, M.; Li, L.; XU, Z.; CHEN, X.; GUO, J.; MA, Y. Overexpression of the soybean GmERF3 gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. **J. Exp. Bot.**, v.60, n.13, p. 3781-3796. 2009.

ZHUANG, J.; CAI, B.; PENG, R-H.; ZHU, B.; JIN, X. F.; XUE Y.; GAO, F.; FU, X. Y.; TIAN, Y. S, ZHAO, W.; QIAO, Y. S.; ZHANG, Z.; XIONG, A. S.; YAO, Q. H. Genome-wide analysis of the AP2/ERF gene family in *Populus trichocarpa*. **Biochem. Biophys. Res. Commun.**, v. 371, n. 3, p. 468-474. 2008.

ZHUANG, J.; CHEN, J. M.; YAO, Q-H.; XIONG, F.; SUN, C. C.; ZHOU, X. R.; ZHANG, J. A.; XIONG, A. S. Discovery and expression profile analysis of AP2/ERF family genes from *Triticum aestivum*. **Mol. Biol. Rep.**, v.38, n.2, p. 745–753. 2011a.

ZHUANG, J.; PENG, R-H.; CHENG, Z-M.; ZHANG, J.; CAI, B.; ZHANG, Z.; Gao, F.; ZHU, B.; FU, X-Y.; JIN, X-F.; CHEN, J-M, QIAO, Y-S.; XIONG, A-S.; YAO, Q-H. Genome-wide analysis of the putative AP2/ERF family genes in *Vitis vinifera*. **Sci Hortic.**, v.123, n.1, p. 73-81. 2009.

ZHUANG, J.; DENG, De-X.; YAO.; Q-H.; ZHANG, J.; XIONG, F.; CHEN, J-M.; XIONG, Ai- S. Discovery, phylogeny and expression patterns of AP2-like genes in maize. **Plant Growth Regul.**, v.62, n.1, p.51-58. 2010.

ZHUANG, J.; YAO, Q-H.; XIONG, A. S.; ZHANG, J. Isolation, Phylogeny and Expression Patterns of AP2-Like Genes in Apple (*Malus × domestica* Borkh). **Plant Mol. Biol. Rep.**, v.29, n.1, p. 209–216. 2011b.

5 ARTIGO B**ANÁLISE DA EXPRESSÃO GÊNICA DE TREZE ERFS DE *Coffea arabica*
SUBMETIDOS A DIFERENTES ESTRESSES ABIÓTICOS**

ANÁLISE DA EXPRESSÃO GÊNICA DE TREZE ERFs DE *Coffea arabica* SUBMETIDOS A DIFERENTES ESTRESSES ABIÓTICOS

5.1 RESUMO

Neste trabalho, 13 ERFs de *C. arabica* (CaERF) foram identificados e caracterizados *in silico* e *in vivo*. Com base na caracterização das sequências, foram classificados em oito dos 10 grupos principais da família ERF. A análise da expressão dos 13 ERFs exibiu padrões e níveis de expressão diferencial em resposta aos estresses hídrico, salino e térmico. ERFs dos grupos VIII e IX apresentaram padrões distintos de expressão no estresse hídrico. As ERFs do grupo VIII apresentaram um aumento pronunciado de transcritos quando a planta foi submetida a um estresse hídrico severo, decrescendo no recuperado. Já nas ERFs do grupo IX foi observado um aumento de transcritos no déficit hídrico severo e este aumento se manteve na recuperação após o estresse. Isso sugere que membros desses grupos podem estar envolvidos nos mecanismos de regulação da tolerância a seca. Com exceção dos ERFs *CaERF04*, *CaERF06*, *CaERF13* e *CaERF23* todos os demais foram induzidos por mais de um tipo de estresse abiótico, o que sugere a existência de respostas cruzadas (crosstalk) entre as diferentes vias de sinalização dos estresses. Os dados gerados neste trabalho ajudarão na seleção de genes candidatos para futuras análises funcionais da família ERF em *C. arabica*, o que ajudará a compreender os determinantes genéticos da tolerância a estresses abióticos o que constitui um passo fundamental nos programas de melhoramento genético do cafeeiro.

Palavras-chave: AP2/ERF. Calor. Expressão gênica. Salinidade. Seca.

5.2 ABSTRACT

In this work, 13 *Coffea arabica* ERFs (CaERF) were identified and characterized *in silico* and *in vivo*. Based on the characterization of the sequencing, they were classified in eight of the 10 main groups of the ERF family. The analysis of the expression of the 13 ERFs showed patterns and differential expression levels in response to water deficit, salt and heat stresses. The ERFs of the groups VIII and IX showed distinct patterns of expression in water deficit stress. The ERFs from group VIII presented a high transcript increase when the plant was subjected to the severe water deficit stress, decreasing in the recovery. On the other hand, on the ERFs from group IX it was observed an increase of transcripts in severe water deficit stress, and this rise so remained in recovery after the stress. This suggests that members of these groups may be involved in the regulation mechanisms of tolerance to drought. Except for the ERFs *CaERF04*, *CaERF06*, *CaERF13*, *CaERF23*, all the others were induced by more than one kind of abiotic stress, which suggests the existence of crosstalk between the different signaling pathways of the stresses. The data generated in this work will help in the selection of candidate genes for future functional analyses of the ERF family in *C. arabica* in, which will help to understand the genetic determinants of tolerance to abiotic stresses that consists of a fundamental step in genetic improvement programs on the coffee plant.

Key words: AP2/ERF. Heat. Gene expression. Salinity. Drought.

5.3 INTRODUÇÃO

Com uma produção mundial de 134 milhões de sacas de grãos em 2010 (ICO, 2011), o café é um importante produto agrícola cultivado em mais de 80 países. O Brasil é o maior produtor mundial e junto com a Colômbia e o Vietnã são responsáveis por mais de 50% da produção mundial (MAPA, 2011). O café é uma das maiores commodities agrícolas do mundo, representa uma importante fonte de renda para os países tropicais em desenvolvimento (PAY, 2009). No gênero *Coffea*, duas espécies são responsáveis pela quase totalidade da produção de grãos de café: *Coffea canephora* e *Coffea arabica*, que contribuem aproximadamente com 30 e 70% da produção mundial, respectivamente. A produção de café está sujeita a oscilações regulares é que explicado, principalmente, pelo ciclo bienal natural, mas também pelos efeitos adversos das condições climáticas. Entre eles, a seca e as altas temperaturas são fatores que afetam o desenvolvimento e produção do cafeeiro (ASSAD et al, 2004; DAMATTA, 2004; DAMATTA, 2006).

Condições ambientais desfavoráveis podem limitar o crescimento e o desenvolvimento de plantas ocasionando reduções na produção agrícola (LU et al, 2011). Para sobreviver aos diferentes estresses bióticos e abióticos, as plantas desenvolveram uma complexa rede de sinalização que conduz à expressão orquestrada de genes que são responsáveis pela tolerância às condições ambientais adversas (WANG et al, 2002; XIONG; ZHU 2003). A regulação gênica em nível de transcrição é um dos principais pontos no controle dos processos biológicos onde os Fatores de Transcrição (TFs) são importantes para o controle da expressão funcional desses genes (SHARMA et al, 2010). Os fatores de transcrição como bZIP, NAC, MYB, MYC, *zinc-finger*, WRKY e *ethylene response factor* (ERF) desempenham um papel crucial no controle da expressão de genes específicos relacionados ao estresse através da ligação aos elementos *cis-acting* nos promotores de genes responsivos aos estresses, ou através da interação com outros fatores de transcrição (CHEN et al, 2004; UMEZAWAET et al, 2006).

A família ERF é definida pela presença de um domínio AP2/ERF, que consiste de 58-59 aminoácidos, no qual estão envolvidas em uma variedade de mecanismos de regulação durante todo o ciclo de vida da planta (SHARONI et al, 2011). A família ERF é de particular interesse devido ao seu papel no desenvolvimento e nas respostas das plantas aos diferentes estresses (NAKANO et

al, 2006). Os processos de desenvolvimento regulados pelos fatores de transcrição ERFs incluem a formação de raízes (BANNO et al, 2001), germinação de sementes (PIRRELLO et al, 2006) e os desenvolvimentos: do embrião (BOUTILIER et al, 2002), da flor (ELLIOTT et al, 1996; LIU et al, 2011) e dos frutos e maturação (WANG et al, 2007; EL-SHARKAWY et al, 2009; YANG et al, 2010). Os fatores de transcrição ERF pode ser divididos nas subfamílias DREB e ERF.

A subfamília DREB desempenha um papel importante na tolerância de plantas a estresses abióticos, reconhecendo o elemento responsivo a desidratação (Dehydration Responsive Element - DRE), com um motivo *core* de A/GCCGAC (LIU et al, 1998). A superexpressão de genes DREB em *Arabidopsis* ativam a expressão de muitos genes relacionados ao estresse melhorando a tolerância à seca, sal e frio (LIU et al, 1998; KASUGA et al, 1999; SAKUMA et al, 2006). Por exemplo, a superexpressão do *AoDREB* de *Asparagus officinalis* L. em plantas transgênicas de *Arabidopsis* induziu a expressão dos genes *rd29A* e *COR15A*, o que resultou em maior tolerância das plantas transgênicas à seca e à salinidade (LIU et al, 2010). O mesmo foi observado por Krishnaswamy et al, (2011) com a superexpressão do *DREB19* em *Arabidopsis*. A subfamília ERF tem demonstrado participar na resposta a estresses bióticos, como o ataque de diferentes patógenos. As proteínas ERFs reconhecem o elemento *cis-acting* AGCCGCC, conhecido como GCC-box e modulam a expressão dos genes relacionados à patogenicidade (PR) (SOLANO et al, 1998; GU et al, 2002; ONATE-SANCHEZ et al, 2007). A superexpressão constitutiva de *AtERF2* mostrou induzir a expressão do gene PDF1.2 (BROWN et al, 2003). Da mesma forma, Liu et al, (2011) demonstrou que a superexpressão de ERFs de *Bupleurum kaoi* em plantas transgênicas de *Arabidopsis* conferiu resistência a *Botrytis cinerea*. No entanto, além de sua participação nas respostas bióticas, alguns membros da família ERF podem integrar diferentes vias de sinalização. Por exemplo, Tsi1 de tabaco (PARK et al, 2001), TERF1 de tomate (ZHANG et al, 2005.) e CaFP1 de pimenta (YI et al, 2004) são integradores de respostas aos estresses biótico e abiótico através da interação com GCC-box e DRE. Assim, as proteínas ERF desempenham um papel importante não apenas em respostas de defesa ao ataque a patógenos, mas também em tolerância a diferentes estresses abióticos.

Entretanto não existem informações dos ERFs em *Coffea arabica*, e seus prováveis papéis na tolerância aos diferentes estresses ainda são

desconhidos. Desta forma, o objetivo do presente trabalho foi caracterizar o padrão transcricional de 13 putativos ERFs em resposta aos estresses hídrico, térmico e salino.

5.4 MATERIAL E MÉTODOS

Análise *in silico* da família ERF em *Coffea arabica*

Para isolar os genes da família ERF em *Coffea arabica*, o domínio AP2 do ERF4 (Genebank: AAO34706) de *Solanum lycopersicum* foi utilizado para pesquisas no Banco de Dados Genoma Café Brasileiro (www.lge.ibi.unicamp.br./café), usando o programa BlastP. Buscas também foram realizadas utilizando as palavras chaves: ERF, Fator de Resposta ao Etileno e EREBP, visando aumentar as chances de identificação de novas ESTs. As sequências foram confrontadas com outras sequências depositadas no banco de dados do GenBank, utilizando os programas BlastP e BlastX (National Center for Biotechnology Information - NCBI - <http://www.ncbi.nlm.nih.gov>). O alinhamento e a árvore filogenética foram realizados a partir da sequência completa de *C. arabica*, *Arabidopsis thaliana*, *Vitis vinifera*, *Solanum lycopersicum* e *Populus trichocarpa* usando o algoritmo ClustalW versão 2.0 (LARKIN et al, 2007). A árvore filogenética foi construída utilizando-se o método neighbor-joining (NJ) usando a opção *pair-wise deletion* com a ajuda do programa MEGA versão 4.0 (TAMURA et al, 2007). Para testar a confiabilidade das análises foram utilizadas 1.000 réplicas de *bootstrap*. A identificação dos putativos motivos de localização nuclear e ativador transcricional (domínios ácidos) foi realizada utilizando o programa PSORT (<http://psort.hgc.jp/>).

Material vegetal

O genótipo utilizado para os experimentos de estresses foi o *Coffea arabica* cv. IAPAR-59. Foram utilizadas plantas com 18 meses de idade. Os experimentos foram realizados em casa de vegetação a 25° C com fotoperíodo controlado (12h:12h) (déficit hídrico e salino) ou em câmara de crescimento Fitotron (estresse térmico). As folhas foram coletadas no mesmo estágio de desenvolvimento (segundo par de folhas de cada ramo plagiotrópico). Para cada amostragem foram

feitas coletas de folhas (três *pools* de folhas de diferentes plantas, sob as mesmas condições de estresse onde cada *pool* constituiu uma repetição biológica). Os materiais coletadas foram imersos imediatamente em N líquido e armazenados em freezer – 80°C.

Experimento de estresse hídrico

Foram utilizadas 24 plantas submetidas a um período sem adição de água. As amostragens foram realizadas conforme as leituras do potencial hídrico, sendo submetidos aos tratamentos: sem estresse (-1.34 MPa), estresse moderado (-2.4 MPa), estresse severo (-4.5 MPa) e recuperado (material coletado 72 h após irrigação das plantas atingirem o potencial hídrico estabelecido como estresse severo). A coleta das plantas para a realização dos *pools* foi baseada no potencial hídrico em função dos níveis de estresses estabelecidos, realizados através do monitoramento do potencial hídrico de cada planta por psicrometria da seguinte forma: o potencial de água na folha foi monitorado através de psicrômetros de termopar (modelo C-30, Wescor, Inc., Logan, Ut, U.S.A) acoplados a um datalogger (Campbell Scientific, Inc., Logan, Ut, U.S.A., modelo CR-7). A cada dia de leitura, entre 10h e 11h, uma amostra de aproximadamente 2 cm² da folha, coletada da parte mediana do ramo plagiotrópico, posicionado no meio da cada planta, foi coletada e colocada nos psicrômetros. O datalogger foi programado para que as leituras fossem efetuadas a cada 10 min até o equilíbrio da pressão de vapor na câmara. A microvoltagem fornecida pelo sistema foi convertida em potencial da água (MPa) em função de prévia calibração dos sensores com soluções de cloreto de sódio. Após a obtenção do potencial total da água (Ψ_t), os sensores foram imersos durante 4 min em nitrogênio líquido e as leituras retomadas para obtenção do potencial osmótico (Ψ_s). O cálculo do potencial de pressão foi realizado através de $\Psi_t - \Psi_s$.

Experimento de estresse salino

Foram utilizadas 9 plantas. Para evitar choque osmótico, as plantas foram irrigadas no primeiro dia com 50 mM de NaCl, e no segundo dia com 100 mM. A partir do terceiro dia até o final do experimento as plantas foram irrigadas

diariamente com 150 mM de NaCl. Foram realizadas as seguintes amostragens: dia 0 (controle sem adição de NaCl), dia 1 (24 horas após a adição de 50 mM de NaCl), dia 2 (24 horas após a adição de 100 mM de NaCl) e dias 4, 6 e 12 (150 mM NaCl).

Experimento de estresse térmico

Foram utilizadas 9 plantas. As plantas foram aclimatadas durante sete dias com temperatura de 25°C com fotoperíodo controlado (12h:12h). Após esse período, deu-se o início aos tratamentos de estresse térmico: a temperatura de 37°C por cinco dias e manutenção diária de irrigação. Foram realizados quatro pontos de amostragem para o experimento de estresse térmico: Dia 0 - folhas amostradas um dia antes do início do tratamento em plantas mantidas a 25°C, Dia 1, Dia 3 e Dia 5 - folhas amostradas um, três e cinco dias após o início do tratamento térmico, respectivamente. Foram realizadas medições das trocas gasosas nas plantas utilizando o sistema portátil de fotossíntese (Li6200-Licor, Licon, NE, EUA), com câmara de 1 litro, para a obtenção dos valores da taxa fotossintética após o início do estresse térmico.

Extração de RNA total e síntese de cDNA

Foi realizada a extração do RNA total das folhas de *Coffea arabica* cv. IAPAR-59 para os experimentos de estresses abióticos descritos anteriormente. O material vegetal foi coletado e armazenado a -80°C até a sua utilização. As extrações foram conduzidas utilizando o tampão de extração CTAB de acordo com o protocolo descrito por Chang et al, (1993). O RNA foi tratado com DNase I (Invitrogen®) de acordo com as recomendações do fabricante. A primeira-fita de cDNA foi sintetizada utilizando 5 µg do RNA total purificado, utilizando o Kit da transcriptase reversa Superscript III (Invitrogen®), seguindo recomendações do fabricante.

Análise da expressão semi-quantitativa por RT-PCR

As reações de amplificação para análise da expressão semi-quantitativa por RT-PCR foram realizadas em um volume final de 20 µL, contendo: 1

X tampão (75 mM Tris-HCl pH 9; 50 mM KCl, 2 mM de $(\text{NH}_4)_2\text{SO}_4$); 0,25 mM de cada dNTP, 0,5 mM de oligonucleotídeos; 1,5 unidade de Taq DNA polimerase; 1 μl cDNA diluído 10X e água ultra-pura para completar o volume. As reações foram realizadas em termociclador modelo PT-100 (MJ Research, Massachusetts, EUA), com as seguintes condições: 1 ciclo inicial de 5 min a 94°C e 28-32 ciclos de: 30 seg a 94°C para desnaturação, 45 seg a 60-63°C para anelamento, 1 min a 72 °C para extensão e 1 ciclo final de 10 min a 72 °C. Os oligonucleotídeos utilizados para a amplificação, temperatura de anelamento (T_m) e o tamanho dos fragmentos encontram-se na Tabela 1. O gene *GAPDH* foi utilizado como controle interno para normalizar o nível de expressão dos genes alvo entre os diferentes tratamentos. Em alguns ERFs, a temperatura de anelamento foi fixada em 63°C para evitar a amplificação dos produtos de PCR inespecíficos. A reação foi repetida por 28-32 ciclos para obter uma adequada quantidade de DNA. As condições e número de ciclos foram determinados para evitar a saturação do DNA amplificado. O produto de PCR obtido foi submetido à eletroforese em gel de agarose 1,2% corados com brometo de etídio. Os géis foram fotografados utilizando câmera Kodak Digital Science™ - Eletrophoresis Documetation and Analysis System 120 e analisados densitometricamente (FRESCHI et al, 2009) utilizando o software IMAGEJ 1.43u (<http://imagej.nih.gov/ij/index.html>). Para cada reação foram realizadas três repetições biológicas.

Tabela 1 – Sequencias dos oligonucleotídeos de *CaERF* utilizados para a semi-quantitativa RT-PCR.

ERF	Primer Foward (5')	Primer Reverse (3')	Tm (°C)	Fragmento (bp)
<i>CaERF01</i>	TGGATTTCTGGAATAGTACACCTTTAGA	GGAGGAAGCACCAGATTTTCATAA	60	103
<i>CaERF04</i>	AACACCTTTGATGATTCATTGAACAG	TCAGATCTTGAAGGTGATAATGATGAT	63	100
<i>CaERF06</i>	TCTTAGTGGGCGATACTACTG	CTCTCGAACCCACCTCTATT	60	88
<i>CaERF09</i>	CAACGCCCCCGCTTATAA	AAAGGTTGCTGTCGTGATTAGCTA	62	71
<i>CaERF13</i>	AATAGGCATCAGCCCTCAATCTC	GTGATCTCCGGTTCGTCTATGG	62	100
<i>CaERF19</i>	GTGGCGGAGCTATAATTTCC	CTTCTTAATGTCCTTGCTATCAGG	62	121
<i>CaERF22</i>	CGATGGATTATGCTAACGTGGTT	TTGTCGCCGTCGCTTATATTATC	60	106
<i>CaERF23</i>	GATGGATTATGCTAACGTGGTTG	TGTCGCCGTCGCTTATATTA	60	99
<i>CaERF27</i>	CGGAAGACATGGTAATTTACTG	GGCTCAGCTTTAACTCTAAC	60	114
<i>CaERF29</i>	ATCCGATATCCGCTGAACTC	TCGACCTTAAGAGGCAAGTC	60	129
<i>CaERF30</i>	GGCAGAATTACCTCTTAACGA	CTTGAAGAAGGGAAATTTGAGG	60	120
<i>CaERF31</i>	CTGAAACTTCTGCCTTGGAG	AATCGGAGGAATTGTTAGACTG	60	137
<i>CaERF36</i>	GCTTCACATCCTCTGCTGCTT	CGACGGTGCCGAAAA	60	101
<i>GAPDH</i>	AGGCTGTTGGGAAAGTTCTTC	ACTGTTGGAAGTTCGGAATGC	60	100

Análise da expressão quantitativa Real Time-PCR

Os oligonucleotídeos dos genes específicos *CaERF19* (5'GTGGCGATCTGCCTGATAGC e 3'CTCGTCGTCTAGGTCAATGATCTC) e *RPL39* (5'GAACAGGCCCATCCCTTATTG e 3'CGGCGCTTGGCATTGTA) foram desenhados com o auxílio do software Primer Express (Applied Biosystems, Foster City, CA, USA) com fragmento de amplificação de 100bp. *RPL39* foi utilizado como controle interno para normalizar o nível de expressão do gene alvo entre os diferentes tratamentos. A PCR quantitativa em tempo real foi realizada em um volume final de 25 µL utilizando o Kit Platinum[®] SYBR[®] Green qPCR Super Mix-UDG conforme recomendação do fabricante. As reações foram realizadas em termociclador ABI 7500 Real Time PCR System (Applied Biosystems, Foster, CA, USA), com as seguintes condições: 1 ciclo de 2 min a 50°C; 1 ciclo de 2 min 95 °C; 40 ciclos de: 15 seg a 95°C e 30 seg a 60°C. Para cada reação foram realizadas três repetições biológicas em triplicata. Os dados foram analisados usando o método $\Delta\Delta C_t$ (LIVAK; SCHMITTGEN, 2001).

Análise Estatística

Os dados de expressão foram submetidos à análise de variância e as médias comparadas pelo Teste de Tukey a 5% de probabilidade, utilizando-se o programa estatístico SISVAR (FERREIRA, 1999).

5.5 RESULTADOS

Características moleculares dos ERFs de *Coffea arabica*

Os ERF utilizados neste trabalho foram identificados a partir de análise *in silico* do Banco de Dados Genoma Café Brasileiro e selecionados de modo a representar o maior número possível dos grupos filogenéticos e de acordo com estudos preliminares de expressão. Treze ERFs de *Coffea arabica* (*CaERF*) selecionado e nomeados *CaERF01*, *CaERF04*, *CaERF06*, *CaERF09*, *CaERF13*, *CaERF19*, *CaERF22*, *CaERF23*, *CaERF27*, *CaERF29*, *CaERF30*, *CaERF31* e *CaERF36* codificam proteínas de 270, 360, 164, 311, 246, 303, 222, 228, 244, 264,

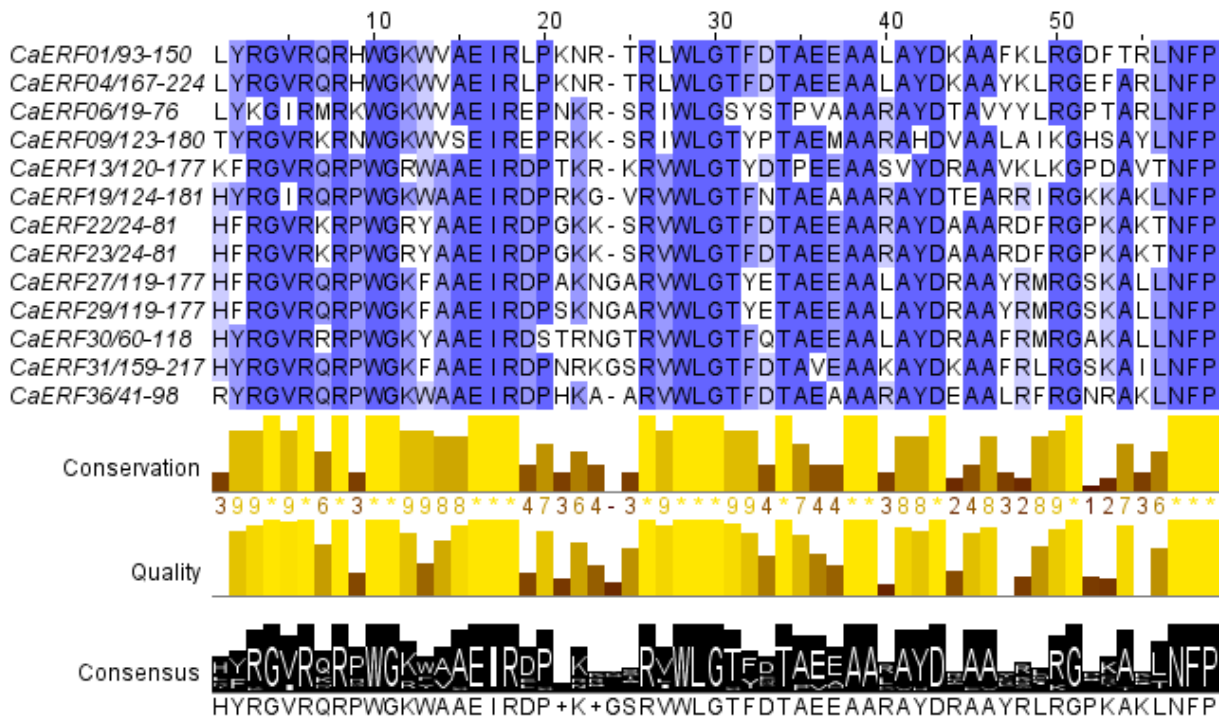
193, 317 e 256 resíduos de aminoácidos, com pesos moleculares de 30.1, 39.9, 11.7, 33.5, 28.2, 34.0, 23.5, 24.6, 26.4, 28.4, 21.3, 35.0 e 27.7 kDa, respectivamente. As relações entre a predição das seqüências de aminoácidos de *Coffea arabica* indicado pelas porcentagens da identidade e similaridade em relação à seqüência inteira são apresentadas na Tabela 2. A identidade entre elas variou de 42% (*CaERF09* e *CaERF22*) a 98% (*CaERF22* e *CaERF23*), enquanto que a similaridade entre os *CaERFs* variou de 59% (*CaERF09* e *CaERF19*) a 98% (*CaERF22* e *CaERF23*). As seqüências deduzidas de aminoácidos dos ERFs de café constituem um domínio conservado de ligação ao DNA, o domínio ERF/AP2 variando de 58-59 aminoácidos, que é característico da família de genes ERF em plantas (Figura 1). O domínio ERF das proteínas de café em relação ao de outras espécies já caracterizados incluindo *Arabidopsis*, uva, tomate e *populus* apresentaram alta homologia da seqüência (75-98% de identidade e de 92-100% de similaridade). Por outro lado, as seqüências completas dos ERFs apresentaram divergências (43-77% similaridade e 55-90% de similaridade), em relação aos ERFs dessas espécies. A baixa similaridade da seqüência completa é devido ao baixo nível de homologia existente fora do domínio ERF.

A subfamília de genes ERF inclui dois resíduos de aminoácidos principais no domínio AP2, a alanina (A) na posição 14 e o aspartato (D) na posição 19, que acredita-se pode contribuir para uma atividade funcional de ligação à GCC box em muitos ERFs (SAKUMA et al, 2002). Por outro lado, a subfamília de genes CBF/DREB possui uma valina (V) e um ácido glutâmico (E) nas posições 14 e 19, respectivamente. O domínio AP2 de todos os *CaERFs* (*CaERF13*, *CaERF19*, *CaERF22*, *CaERF23*, *CaERF27*, *CaERF29*, *CaERF30*, *CaERF31* e *CaERF36*) apresentaram o resíduo conservado para a subfamília ERF. Na subfamília DREB, todos os genes apresentaram o resíduo conservado V14, já na posição 19 os genes *CaERF01* e *CaERF04* contem uma Leucina (L), enquanto que os genes *CaERFs06* e *CaERF09* apresentam um ácido glutâmico (E). Além disso, todos os treze cDNAs possuem uma região básica em suas proteínas que poderiam funcionar como um sinal de localização nuclear (NLS) e mantém um domínio ácido que pode atuar como um domínio de ativação (AD) para transcrição (Figura 2).

Tabela 2 – Comparação entre as seqüências completas preditas de aminoácidos dos fatores de resposta ao etileno (*CaERF*) cDNAs. Acima da diagonal estão os valores da identidade enquanto que abaixo da diagonal estão os valores da similaridade entre os ERFs de *Coffea arabica*.

		Similaridade e Identidade dos aminoácidos (%)												
No aa		<i>CaERF01</i>	<i>CaERF04</i>	<i>CaERF06</i>	<i>CaERF09</i>	<i>CaERF13</i>	<i>CaERF19</i>	<i>CaERF22</i>	<i>CaERF23</i>	<i>CaERF27</i>	<i>CaERF29</i>	<i>CaERF30</i>	<i>CaERF31</i>	<i>CaERF36</i>
<i>CaERF01</i>	270	-	50	59	61	56	63	61	61	68	68	68	66	64
<i>CaERF04</i>	360	63	-	62	60	64	64	63	63	69	66	68	61	65
<i>CaERF06</i>	164	72	74	-	55	54	57	53	53	54	54	47	47	62
<i>CaERF09</i>	311	79	81	70	-	56	46	42	42	59	59	53	56	58
<i>CaERF13</i>	246	66	75	69	74	-	58	70	70	67	67	53	49	64
<i>CaERF19</i>	303	75	79	74	59	73	-	67	65	68	68	62	72	75
<i>CaERF22</i>	222	75	77	72	60	78	81	-	98	71	71	54	60	57
<i>CaERF23</i>	228	75	77	72	60	78	80	98	-	71	71	54	60	63
<i>CaERF27</i>	244	83	85	69	76	80	77	81	81	-	92	45	51	68
<i>CaERF29</i>	264	83	82	70	76	82	77	81	81	94	-	46	54	52
<i>CaERF30</i>	193	80	82	64	68	71	76	66	66	60	63	-	77	65
<i>CaERF31</i>	317	75	72	67	73	66	81	75	75	65	71	89	-	58
<i>CaERF36</i>	256	74	74	76	72	74	84	68	75	76	65	72	73	-

Figura 1 – Alinhamento dos domínios AP2/ERF a partir das proteínas ERF de *C. arabica*. As colunas de cores, azul escuro e azul claro, indicam resíduos de aminoácidos idênticos e conservados, respectivamente.



espécies de plantas mostrou que os ERFs de café pertencem a oito dos 10 grupos ERFs caracterizados de acordo com a classificação proposta por Nakano et al, (2006).

As proteínas *CaERF01* e *CaERF04* são classificadas como membros do grupo I, subfamília DREB (NAKANO et al, 2006) (Figura 3). Ambas as seqüências de aminoácidos compartilham 46-62% de identidade e 54-72% de similaridade com seus homólogos em Arabidopsis, uva, tomate e populus (Tabela 3). *CaERF01* e *CaERF04* possuem o motivo putativo NLS na região C-terminal, o domínio ERF constituído por 58 aminoácidos localizado perto do centro da proteína, além de compreender um domínio ácido, possível ativador da transcrição, localizados na região C-terminal (Figura 2). *CaERF04* além de um domínio ácido na região C-terminal, possui outro domínio ácido na região N-terminal. A proteína predita *CaERF06*, pertencente a subfamília DREB, é classificado como membro do grupo II (NAKANO et al, 2006) (Figura 3). Essa seqüência compartilha 51-59% de identidade e 66-73% de similaridade com seus homólogos em Arabidopsis, uva, tomate e populus. Esse ERF apresenta um domínio conservada de 58 aa próximo da extremidade N-terminal, um putativo NLS localizado logo antes do domínio ERF e um domínio ácido na região C-terminal. O *CaERF09* pertence ao grupo ERF III, da subfamília DREB. Esse ERF possui um putativo NLS localizado na região N-terminal, o domínio ácido na região C-terminal e um domínio ERF composto por 58 aa. Essa sequencia compartilha de 55-64% de identidade e 64-72% de similaridade com seus homólogos Arabidopsis, uva, tomate e populus.

O *CaERF13* é membro do grupo VI, pertencente à subfamília ERF. Essa proteína compartilha 48-62% de identidade e 61-73% de similaridade com seus homólogos putativos das espécies arabidopsis, uva, tomate e populus. *CaERF19* codifica para as proteínas que pertencem ao grupo VII (NAKANO et al, 2006). *CaERF19*, membro da subfamília ERF compartilha 46-56% de identidade e 59-72% de similaridade com seus homólogos putativos das espécies Arabidopsis, uva, tomate e populus. Em concordância com outros ERFs pertencentes a esse grupo, *CaERF19* tem um domínio ERF conservado de 58 aa, perto da extremidade N-terminal, um putativo NLS (KRKRK/118-122aa) localizado logo antes do domínio ERF e um domínio ácido (22-24aa), perto da extremidade N-terminal (Figura 2) (YI et al, 2004, ZHANG et al, 2004). Além disso, possui uma seqüência altamente conservada (MCGGAIL/L) na região N-terminal identificada por Tournier et al,

(2003). No grupo VIII estão os ERFs *CaERF22* e *CaERF23*, onde todos os membros desse grupo possuem o motivo conservado repressor EAR (ERF-associado repressão anfifílico) na extremidade C-terminal (OHTA et al, 2001). *CaERF22* e *CaERF23* e seus homólogos em *Arabidopsis*, uva, tomate e *populus* compartilham 50-77% de identidade e 58-84% de similaridade. O domínio ERF dos membros desse grupo são localizados próximo à região N-terminal e consiste de 58 aa. O motivo putativo NLS foi identificado na região do domínio ERF, enquanto que o domínio ácido de 26 aa é localizado pouco antes do motivo EAR (Figura 2) (TOURNIER et al, 2003). As proteínas preditas *CaERF27*, *CaERF29*, *CaERF30* e *CaERF31* são classificados como membros do grupo ERF IX (NAKANO et al, 2006) (Fig. 2). Essas seqüências compartilham 43-78% de identidade e 55-90% de similaridade com seus homólogos em *Arabidopsis*, uva, tomate e *populus*. O domínio ERF conservado desse grupo é de 59 aa e todos os ERFs desse grupo possuem o domínio ácido na região N-terminal (FUJIMOTO et al, 2000). O domínio ácido de *CaERF27*, *CaERF29* e *CaERF30* foi previamente denominado de 24-aminoácido conhecido como motivo DMLV, cuja sequência é (WGDLPLKENDSEDMLVYGILNDAF de AtERF1) (GUTTERSON; REUBER, 2004). O putativo NLS dos ERFs do grupo IX encontram-se na região C-terminal, exceto de *CaERF30* cujo sinal encontra-se na região N-terminal, logo antes do domínio ERF. A proteína predita *CaERF36* pertence ao grupo ERF X (NAKANO et al, 2006) (Fig. 2). Essa seqüência compartilha 60-75% de identidade e 73-89% de identidade com seus homólogos em *Arabidopsis*, uva, tomate e *populus*. O putativo NLS desse ERFs encontra-se na região N-terminal, logo antes do domínio ERF, enquanto o domínio ácido encontra-se na região C-terminal.

Figura 3 – Análise filogenética entre os ERFs de *Coffea arabica* e dos organismos que apresentaram maior similaridade em *Arabidopsis thaliana* AtERF059 (AT1G78080), AtERF010 (AT5G67190), AtERF034 (AT2G44940), AtERF066 (AT3G14230), AtERF075 (AT3G14230), AtERF078 (AT3G15210), AtERF100 (AT4G17500), AtERF112 (AT2G33710), *Vitis vinifera* VvERF004 (GSVIVP00015253001), VvERF007 (GSVIVP00019541001), VvERF029 (GIDVvP00033798001), VvERF054 (GSVIVP00006483001), VvERF059 (GSVIVP00034010001), VvERF064 (GSVIVP00003062001), VvERF077 (GSVIVP00014229001), VvERF113 (GSVIVP00019350001), *Solanum lycopersicum* SlERF6, SlERF13, SlERF21, SlERF31, SlERF76, SlERF35, SlERF70, SlERF83 e *Populus trichocarpa* PtDREB19 (XP_002301042.1), PtDREB34 (XP_002307277.1), PtDREB82 (XP_002334238.1), PtERF82 (XP_002319978.1), PtERF38 (XP_002331217.1), PtERF48 (XP_002305121.1), PtERF13 (XP_002304592.1), PtERF76 (XP_002306712.1).

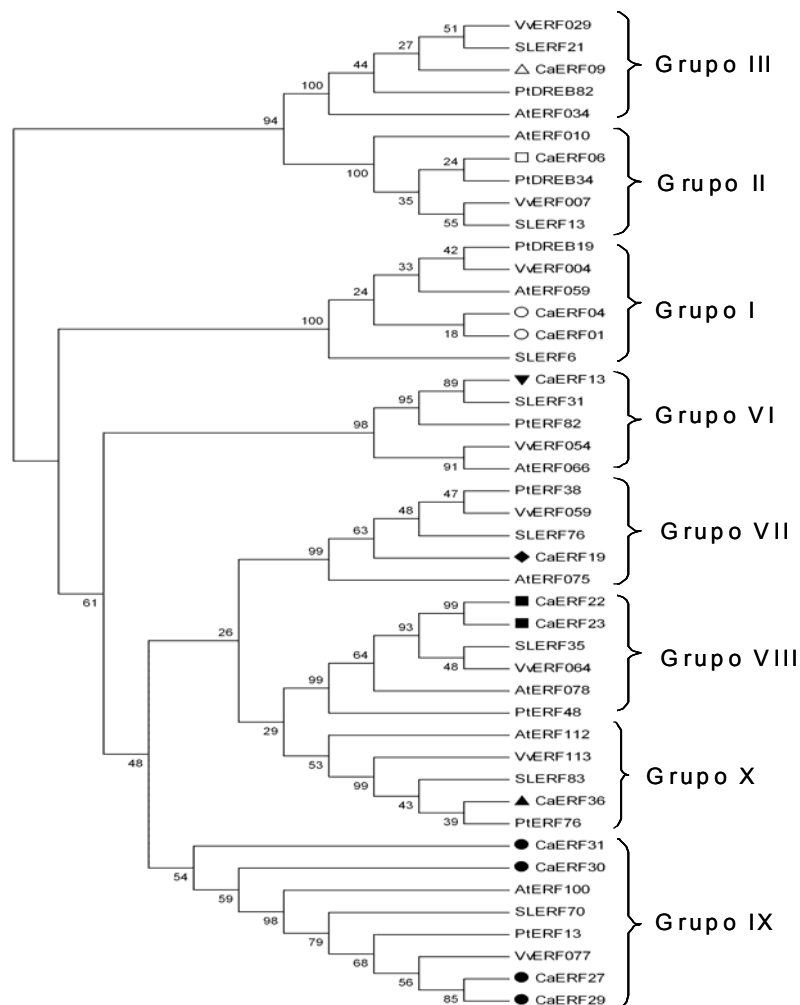


Tabela 3 - Organismos que apresentaram maior similaridade em *Arabidopsis thaliana*, *Vitis vinifera*, *Solanum lycopersicum* e *Populus trichocarpa* genes ERFs de *Coffea arabica*.

<i>Coffea arabica</i>			<i>Arabidopsis thaliana</i> ^a			<i>Vitis vinifera</i> ^b			<i>Solanum lycopersicum</i> ^c			<i>Populus trichocarpa</i> ^d							
ERF	ERF	Genebank	iden.	simil.	e-value	ERF	Genebank	iden.	simil.	e-value	ERF	iden.	simil.	e-value	ERF	Genebank	ident.	simil.	e-value
CaERF01	AtERF059	AT1G78080	54	64	1.00E-68	VvERF004	GSVIVP00015253001	57	65	2.00E-76	SLERF6	62	72	1.00E-83	PtDREB19	XP_002301042.1	53	65	5.00E-63
CaERF04	AtERF059	AT1G78080	49	60	5.00E-62	VvERF004	GSVIVP00015253001	46	54	2.00E-69	SLERF6	49	59	3.00E-70	PtDREB19	XP_002301042.1	49	62	2.00E-80
CaERF06	AtERF010	AT5G67190	51	66	1.00E-41	VvERF007	GSVIVP00019541001	58	68	4.00E-50	SLERF13	59	73	9.00E-47	PtDREB34	XP_002307277.1	59	71	4.00E-42
CaERF09	AtERF034	AT2G44940	55	64	6.00E-48	VvERF029	GIDVvP00033798001	58	68	3.00E-55	SLERF21	64	72	2.00E-63	PtDREB82	XP_002334238.1	59	68	2.00E-45
CaERF13	AtERF066	AT4G27950	48	64	3.00E-30	VvERF054	GSVIVP00006483001	51	61	3.00E-38	SLERF31	62	73	3.00E-45	PtERF82	XP_002319978.1	55	70	5.00E-36
CaERF19	AtERF075	AT3G14230	46	64	8.00E-33	VvERF059	GSVIVP00034010001	46	59	8.00E-62	SLERF76	54	67	2.00E-85	PtERF38	XP_002331217.1	56	72	5.00E-49
CaERF22	AtERF078	AT3G15210	70	79	3.00E-30	VvERF064	GSVIVP00003062001	50	58	2.00E-38	SLERF35	58	65	1.00E-51	PtERF48	XP_002305121.1	76	80	1.00E-19
CaERF23	AtERF078	AT3G15210	70	79	5.00E-30	VvERF064	GSVIVP00003062001	77	84	1.00E-36	SLERF35	58	67	1.00E-48	PtERF48	XP_002305121.1	50	55	1.00E-38
CaERF27	AtERF100	AT4G17500	60	73	9.00E-53	VvERF077	GSVIVP00014229001	65	74	3.00E-68	SLERF70	71	79	3.00E-77	PtERF13	XP_002304592.1	67	73	4.00E-47
CaERF29	AtERF100	AT4G17500	55	68	5.00E-52	VvERF077	GSVIVP00014229001	58	67	1.00E-71	SLERF70	59	68	8.00E-76	PtERF13	XP_002304592.1	57	64	5.00E-67
CaERF30	AtERF100	AT4G17500	49	63	1.00E-30	VvERF077	GSVIVP00014229001	48	67	3.00E-35	SLERF70	48	67	9.00E-34	PtERF13	XP_002304592.1	43	55	3.00E-31
CaERF31	AtERF100	AT4G17500	78	90	3.00E-25	VvERF077	GSVIVP00014229001	50	69	1.00E-32	SLERF70	67	80	7.00E-32	PtERF13	XP_002304592.1	52	69	8.00E-33
CaERF36	AtERF112	AT2G33710	75	89	4.00E-28	VvERF113	GSVIVP00019350001	60	74	1.00E-32	SLERF83	60	73	5.00E-53	PtERF76	XP_002306712.1	62	73	2.00E-25

^aNAKANO et al, (2006), ^bLICAUSI et al, (2010), ^cSHARMA et al, (2010) e ^dZHUANG et al, (2008).

Análise da expressão por semi-quantitativa RT-PCR dos ERFs em *C. arabica* em resposta a estresses abióticos

A análise do padrão de expressão de novos genes pode ajudara revelar suas possíveis funções biológicas (JIN et al, 2010a). Neste trabalho treze genes foram selecionados, de modo a representar o maior número possível dos grupos filogenéticos da família de FT ERF para análise da expressão em diferentes estresses abióticos, sendo quatro DREBs e nove ERFs. Os padrões de expressão dos *CaERFs* foram avaliados e os níveis de mRNAs correspondentes foram quantificados em resposta aos estresses hídrico, térmico e salino por semi-quantitativa RT-PCR.

Estresse hídrico

Através da análise do perfil transcricional dos ERFs de *Coffea arabica*, observou-se que dos treze ERFs, oito apresentaram um aumento pronunciado de transcritos nas folhas quando a planta foi submetida a um estresse hídrico severo (Figura 4A). Dentro da subfamília DREB, foi possível observar que para o gene *CaERF01* a máxima atividade transcricional foi encontrada no estresse hídrico severo, enquanto que *CaERF09* teve sua expressão reprimida no estresse hídrico severo. Os genes *CaERF04* e *CaERF06* apresentaram uma expressão constante durante todo o período do estresse. Dentre da subfamília ERF, os genes *CaERF19*, *CaERF22*, *CaERF23* e *CaERF30* apresentaram um padrão da transcrição muito semelhante (Figura 5) onde se observou um aumento pronunciado de transcritos quando a planta foi submetida a um estresse hídrico severo, havendo um decréscimo no recuperado. Para os genes *CaERF27*, *CaERF29* e *CaERF30*), houve um aumento transcricional acentuado em plantas submetidas ao déficit hídrico severo e este aumento se manteve na recuperação após o estresse. A análise dos 13 ERF identificou cinco padrões de expressão demonstrados na Figura 5.

Estresse salino

Através da análise do perfil transcricional dos ERFs de *Coffea arabica*, observou-se que dos treze ERFs, sete apresentaram um aumento de transcritos nas folhas quando a planta foi submetida ao estresse salino (Figura 4B). Dentre os DREBs, *CaERF01* e *CaERF09* apresentaram um padrão da expressão semelhante, onde se observou um aumento de transcritos 6 dias após o início do tratamento. Os genes *CaERF04* e *CaERF06* apresentaram uma expressão constante durante todo o período do estresse. Dentre os ERFs, o gene *CaERF30* apresentou uma indução da expressão a partir do primeiro dia de estresse, atingindo o máximo acúmulo de transcritos 2 dias após o início do tratamento. Porém uma rápida queda no número de transcritos foi observado no dia 6 de estresse. Por outro lado, o gene *CaERF31* mostrou um padrão de expressão constante durante o período inicial do estresse sendo observado um aumento no quarto dia e no 12 dia após o início do tratamento. Para o *CaERF22* verificou-se um acúmulo da atividade transcricional 4 dias após o início do tratamento de estresse salino. *CaERF19* e *CaERF29* apresentaram o padrão de expressão muito similar, onde foi observado um aumento da atividade transcricional no sexto dia de estresse. Entretanto, para *CaERF27* o padrão de transcrição manteve-se alto em plantas com 12 dias de estresse.

Estresse térmico

Através da análise do perfil transcricional dos ERFs de *Coffea arabica*, observou-se que dos treze ERFs, cinco apresentaram um aumento pronunciado de transcritos nas folhas quando a planta foi submetida ao estresse térmico (Figura 4C). Dentre os DREBs, foi observado um aumento da expressão em *CaERF01* a partir do primeiro dia de estresse térmico, que se manteve até o quinto dia do estresse. Um aumento da atividade transcricional também foi observado para *CaERF06*, mas no quinto dia de estresse. *CaERF09* teve sua expressão reprimida no terceiro dia do estresse térmico, de forma similar a resposta ao estresse hídrico. O gene *CaERF04* apresentou uma expressão constante durante todo o período do estresse, também similarmente ao observado durante o estresse hídrico. Dentre os ERFs, o gene *CaERF19* apresentou alta expressão nos dias 3 e 5 após o início do

tratamento, sendo observado um grande aumento da atividade transcricional no terceiro dia de estresse, diminuindo no quinto dia de estresse térmico, porém com uma transcrição maior que no período inicial de estresse (D1). *CaERF30* e *CaERF31* apresentaram um padrão da expressão semelhante onde se observou um aumento gradativo no acúmulo de transcritos atingindo o máximo 5 dias após o início do tratamento. Os demais ERFs apresentaram uma expressão constante durante todo o período do estresse.

Figura 4A – Expressão relativa dos genes ERFs de *Coffea arabica* submetidos ao estresse hídrico (C - controle, EM - estresse moderado, ES - estresse severo e R - recuperado). Os níveis de expressão foram determinados por RT-PCR semi-quantitativa usando oligonucleotídeos específicos e normalizadas com o gene GAPDH. As imagens fotográficas da eletroforese em gel são mostradas na parte inferior de cada gráfico. Entre parênteses está o número de ciclos. Os dados representam as médias \pm desvio padrão de três repetições biológicas. Médias seguidas de mesma letra não diferem entre si por Tukey 5%.

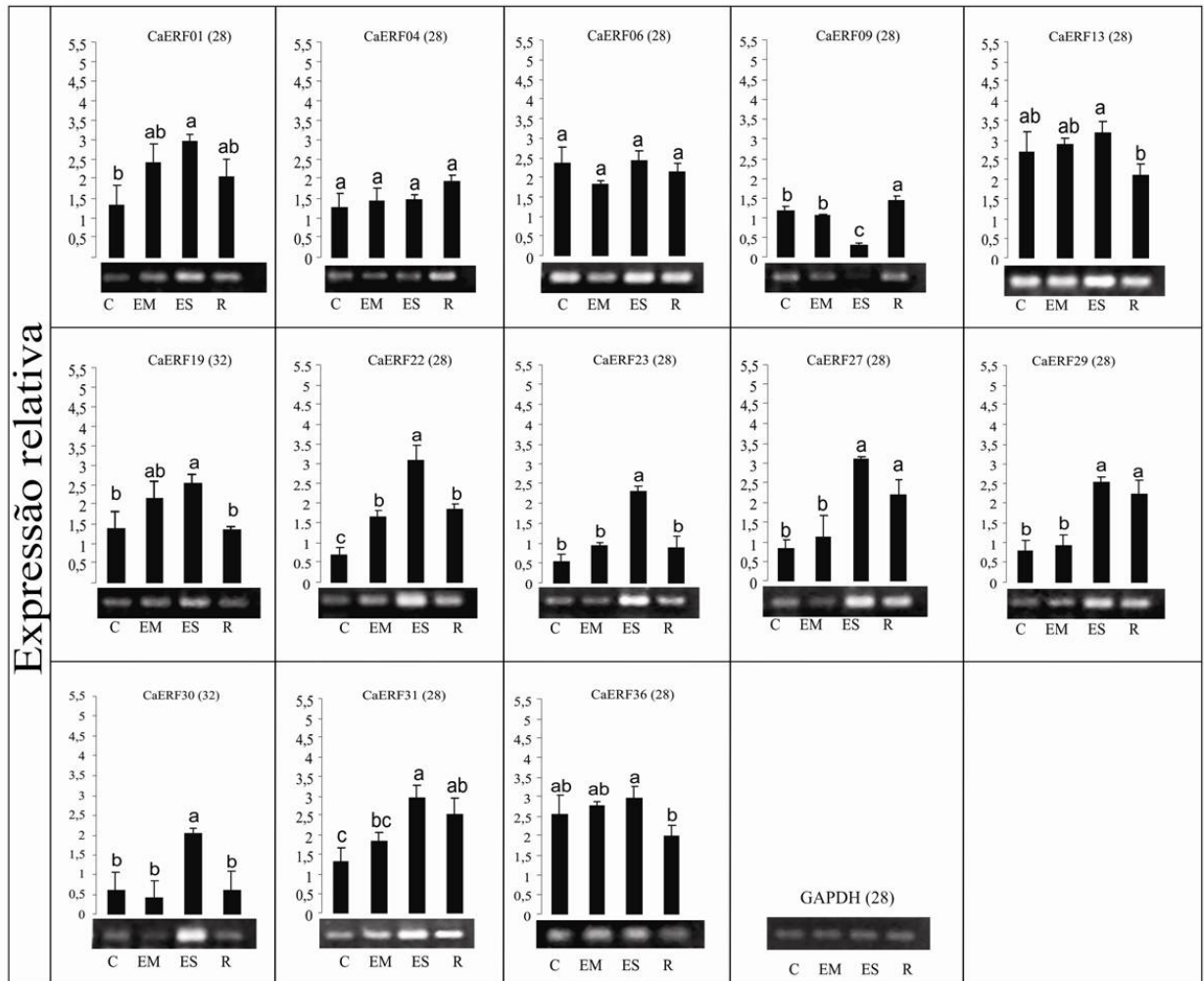


Figura 4B – Expressão relativa dos genes ERFs de *Coffea arabica* submetidos ao estresse salino (C-controle, D1- dia 1, D2- dia 2, D4-dia 4, D6 - dia 6 e D12 - dia 12). Os níveis de expressão foram determinados por RT-PCR semi-quantitativa usando oligonucleotídeos específicos e normalizadas com o gene GAPDH. As imagens fotográficas da eletroforese em gel são mostradas na parte inferior de cada gráfico. Entre parênteses está o número de ciclos. Os dados representam as médias \pm desvio padrão de três repetições biológicas. Médias seguidas de mesma letra não diferem entre si por Tukey 5%.

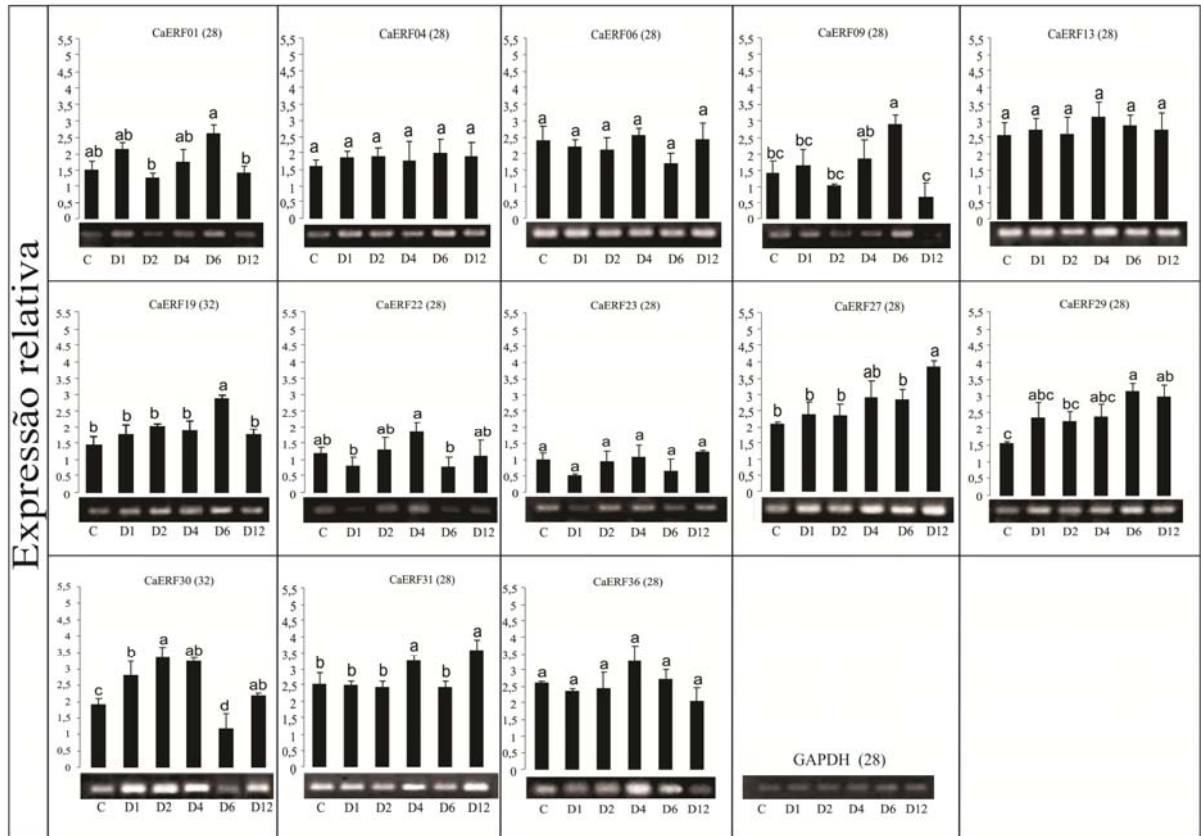


Figura 4C – Expressão relativa dos genes ERFs de *Coffea arabica* submetidos ao estresse térmico (C-controle, D1- dia 1, D3- dia 3 e D5-dia 5). Os níveis de expressão foram determinados por RT-PCR semi-quantitativa usando oligonucleotídeos específicos e normalizadas com o gene GAPDH. As imagens fotográficas da eletroforese em gel são mostradas na parte inferior de cada gráfico. Entre parênteses está o número de ciclos. Os dados representam as médias \pm desvio padrão de três repetições biológicas. Médias seguidas de mesma letra não diferem entre si por Tukey 5%.

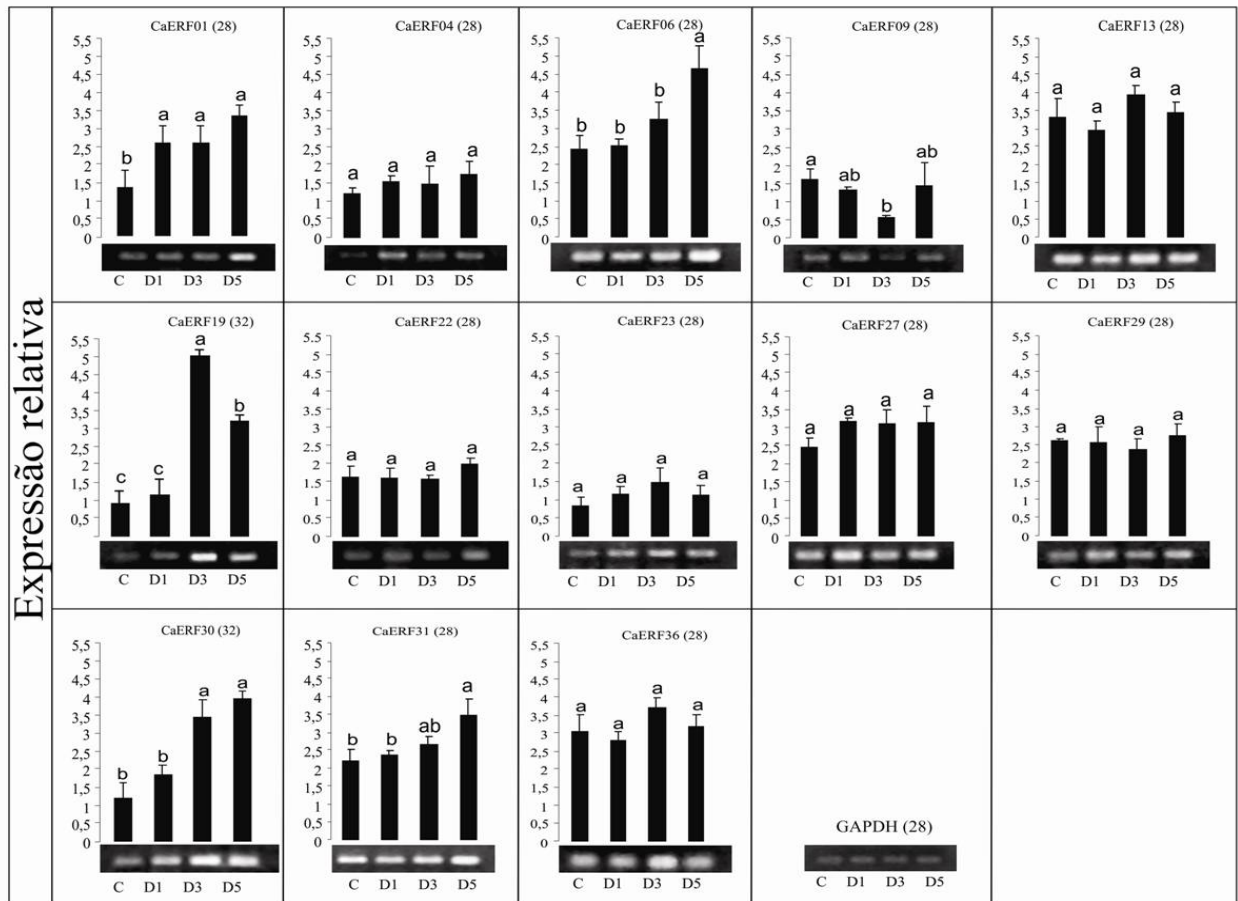
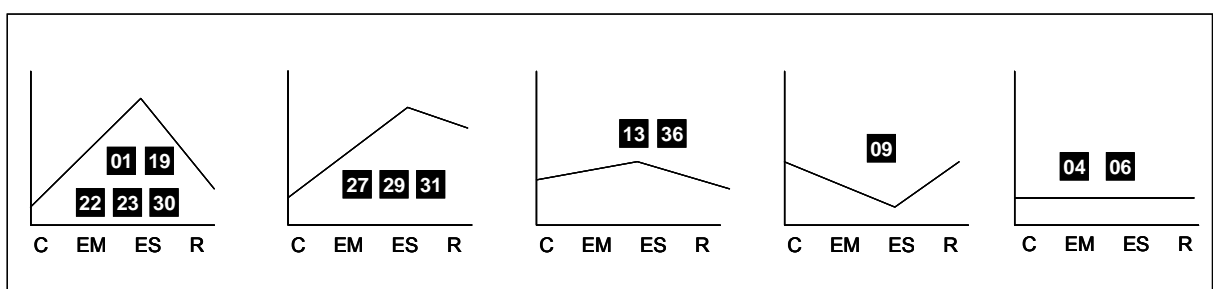


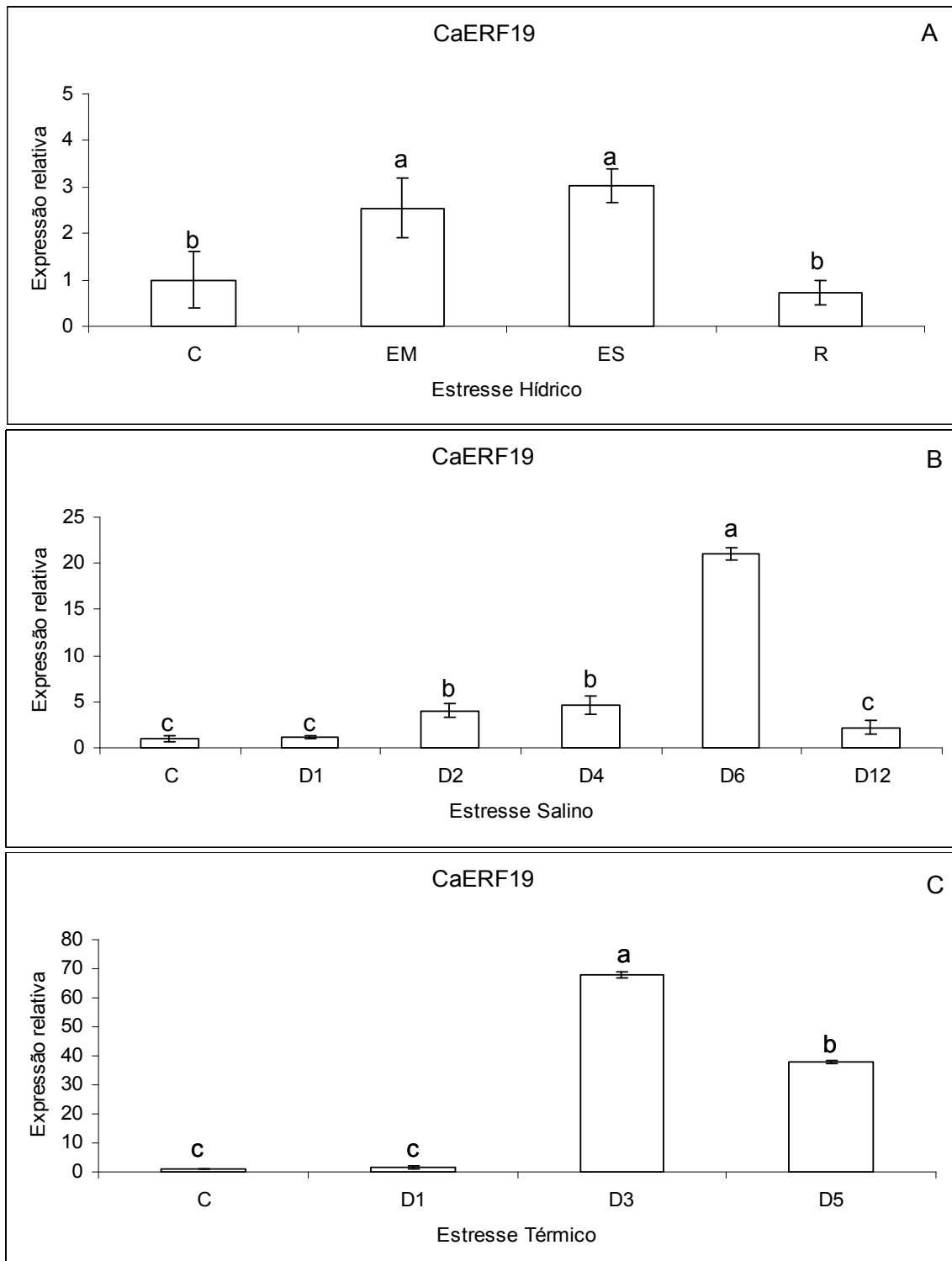
Figura 5 – Padrão de expressão exibido pela família ERF em *Coffea arabica* durante o estresse hídrico. Foram observados 5 padrões de expressão. O eixo x representa: C-controle, EM-estresse moderado, ES-estresse severo e R-recuperado e nas caixas em preto está o número de cada CaERF.



Análise da expressão por RTq-PCR do gene *CaERF19* em resposta a estresses abióticos

A expressão de *CaERF19* também foi analisada por RTq-PCR durante os estresses térmico, salino e hídrico. Os resultados observados foram similares aos obtidos por PCR semi-quantitativo, porém a análise dos estresses térmico e salino apresentaram uma maior diferença nos níveis de expressão relativa. Quando submetido a estresse hídrico (Figura 6A) foi observado um aumento transcricional acentuado em plantas submetidas ao déficit hídrico moderado, atingindo o máximo acúmulo de transcritos no déficit hídrico severo e decrescendo na recuperação após o estresse. Para o estresse salino (Figura 6B), diferentemente da análise utilizando RT-PCR semiquantitativo, a análise por RTq-PCR detectou um aumento de atividade a partir do segundo e quarto dia de estresse. Entretanto, o padrão de expressão das duas análises foi similar, com um maior aumento de transcritos no sexto dia após o estresse e diminuição nas análises no 12º dia. Durante o estresse térmico (Figura 6C) foi observado um grande aumento da atividade transcricional no terceiro dia, com uma diminuição no quinto dia, mas ainda em níveis superiores ao primeiro dia de estresse.

Figura 6 - A análise da expressão de *CaERF19* em resposta aos tratamentos de estresses abióticos por quantitativa RT-PCR. Figura 6A: Estresse hídrico (C-controle, EM-estresse moderado, ES-estresse severo e R-recuperado). Figura 6B: Estresse térmico (C-controle, D1 - dia 1, D3 - dia 3 e D5 - dia 5) e Figura 6C: Estresse salino (C-controle, D1 - dia 1, D2 - dia 2, D4 - dia 4, D6 - dia 6 e D12 - dia 12). Os dados representam as médias \pm desvio padrão de três repetições biológicas. Médias seguidas de mesma letra não diferem entre si por Tukey 5%.



5.6 DISCUSSÃO

Os FT ERFs exercem um papel importante na resposta a estresses abióticos. No entanto a maioria dos estudos tem sido conduzido em plantas modelos como *Arabidopsis* e arroz (NAKANO et al, 2006; SHARONI et al, 2011), e pouco são os estudos em espécies de importância Agrônômica, como o cafeeiro. Neste trabalho treze membros da família FT ERFs foram identificados e caracterizados em *Coffea arabica*. A identidade entre as seqüências foi alta, variando de 42% a 98% entre os diferentes ERFs de *C. arabica*. Todos os ERFs possuem elementos altamente conservados associados aos fatores de transcrição, incluindo a presença de um domínio AP2/ERF que é responsável pela atividade de ligação ao DNA, o domínio ativador ácido (AD), e uma região de aminoácidos básicos que supostamente funciona como um sinal localização nuclear (NLS) (FUJIMOTO et al, 2000; OHTA et al, 2001; SAKUMA et al, 2002; TOURNIER et al, 2003; CAO et al, 2006). No entanto, as variações dentro dos ERFs são suficientes para dividi-los em subfamílias ou mesmo em grupos. Baseado nos resíduos conservados do domínio AP2/ERF foi possível identificar os ERFs de *Coffea arabica* pertencentes às subfamílias DREB e ERF. Para a subfamília DREB os quatro ERFs de *Coffea arabica* apresentaram a valina na posição 14. Já na posição 19, *CaERF01* e *CaERF04* apresentam uma leucina, enquanto que *CaERF06* e *CaERF09* apresentaram um ácido glutâmico. Trabalhos anteriores sugerem que V14 e E19 no domínio AP2/ERF são essenciais para ligação específica ao elemento responsivo DRE (CAO et al, 2001; SAKUMA et al 2002). No entanto, *CaERF01* e *CaERF04* possuem uma L19. Essa mesma diferença também foi observado em outras plantas, incluindo arroz, soja e milho. Em monocotiledôneas como arroz, trigo e cevada, o aminoácido da posição 19 do domínio AP2/ERF de DREB1 é a valina (DUBOUZET et al, 2003). Estas observações sugerem que a função da V14 é provavelmente mais importante do que o aminoácido da posição 19 para a atividade de ligação ao DNA (WANG et al, 2011). Todos os membros de *Coffea arabica* pertencentes à subfamília ERF apresentaram os resíduos conservados alanina e aspartato nas posições 14 e 19, respectivamente, concordando com os resultados encontrados por Sakuma et al (2002).

Baseado em suas organizações estruturais e através da construção da árvore filogenética contendo os ERFs de *Arabidopsis thaliana*, *Vitis vinifera*,

Solanum lycopersicum e *Populus trichocarpa* foi observado que dos 10 principais grupos identificados nessas espécies, 8 grupos ocorreram em *C. arabica*. *CaERF01* e *CaERF04* pertencem ao grupo I da subfamília DREB que funcionam como ativadores da transcrição (LIN et al, 2008; XU et al, 2010). Da subfamília DREB, o *CaERF06* pertence ao grupo II também de ativadores da transcrição. Entretanto, alguns membros desse grupo podem conter o motivo EAR, com função de repressão das atividades transcricionais. Este domínio também é encontrada em outras proteínas DREBs, incluindo RAP2.1 de *Arabidopsis*, GhDBP1 de algodão, GmDREB1 de soja e OsRAP2.1 de arroz (NAKANO et al, 2006; CIFTCI-YILMAZ et al, 2007; DONG; LIU, 2010). Esses trabalhos sugerem que pode haver um mecanismo de adaptação molecular nas respostas de estresse da planta de forma harmoniosa mediado por proteínas DREB que funcionam como ativadores ou repressores. O *CaERF09*, pertencente ao grupo III da subfamília DREB, também caracterizada como um grupo ativador da transcrição (KIZIS; PAGES, 2002; NAKANO et al, 2006).

Quanto aos membros da subfamília ERF, *CaERF13* pertence ao grupo VI caracterizado como ativador da transcrição. *CaERF19* é caracterizado como membro do grupo VII devido a localização do suposto NLS e a presença do motivo (MCGGAI/L) conservado na região N-terminal (TOURNIER et al, 2003). Tournier et al, (2003) demonstrou que esse motivo provavelmente seja necessário para a localização nuclear ou para a ligação com o GCC-box. Mais tarde, Xu et al, (2007) verificaram que a supressão desse motivo na região N-terminal não teve efeito sobre a caracterização molecular da proteína *TaERF1*, incluindo a localização subcelular e a ativação transcricional, entretanto foi verificado que houve um aumento da interação com a proteína quinase. Há divergências sobre a função do motivo MCGGAI/L e trabalhos devem ser realizados a fim de descobrir a potencial função desse motivo. Os ERFs *CaERF22* e *CaERF23* pertencem ao grupo VIII e alguns membros desse grupo têm demonstrado ser repressores ativos da transcrição (FUJIMOTO et al, 2000; OHTA et al, 2001; MATSUI; OHME-TAKAGI, 2010). O domínio de repressão está situado na região C-terminal denominado como ERF-associado a repressão anfifílicas (EAR) e contém uma seqüência de consenso conservada (L/F)DLN(L/F)xP (OHTA et al, 2001). Dois membros do grupo VIII, *AtERF4* e *AtERF7*, foram bem caracterizados. Ambas as proteínas regulam negativamente o gene PDF1.2, no qual contém GCC-box na região promotora.

AtERF4 demonstrou ser um regulador negativo da expressão de etileno, JA e genes ABA-responsivos (OHTA et al, 2001; McGRATH et al, 2005). Além disso, foi mostrado que AtERF7 atua como um regulador negativo nas respostas ao ABA (YANG et al, 2005). Os FT ERF do grupo IX e dos quais fazem parte *CaERF27*, *CaERF29*, *CaERF30*, *CaERF31* e do grupo X do qual faz parte *CaERF36* também foram caracterizados como ativadores da transcrição (CHAMPION et al, 2010; WANG 2007).

A atividade de grande parte dos genes da família ERF está relacionada com aumento da tolerância das plantas a estresses bióticos e abióticos (PARK et al, 2001; GUO et al, 2004; ZHANG et al, 2009). Com base na análise filogenética e sua expressão de resposta ao estresse abiótico, a maioria dos *CaERFs* possuem um possível papel como ativadores de transcrição que podem fazer parte dos mecanismos de aumento da tolerância aos estresses abióticos em *C. arabica*. Dentre os treze ERFs analisados, onze foram diferencialmente regulados em níveis de mRNA nas folhas das plantas de café submetidas ao estresse hídrico. Dentre os DREBs estudados, *CaERF01* foi o único gene *up-regulated* durante o estresse hídrico, onde o nível máximo de transcritos foram acumulados durante o estresse severo, decrescendo na recuperação. Níveis elevados dos DREBs *ZmDBF1* e *GhDBP2* foram encontrados em milho e algodão, respectivamente, quando as plantas foram submetidas a seca (KIZIS; PAGES, 2002; HUANG et al, 2008). Os autores mostraram que tanto *ZmDBF1* como *GhDBP2* podem estar envolvidos na regulação de alguns genes do tipo LEA (Late Embryogenesis Abundant). Essas proteínas acumulam-se durante os estágios avançados da embriogênese e também em tecidos expostos a estresses como desidratação, estresse osmótico e baixa temperatura desempenhando um papel importante na estabilização das proteínas e estrutura da membrana durante a desidratação celular (HONG-BO et al, 2005).

Dentre da subfamília ERF, todos os genes analisados foram *up-regulated* quando submetidos ao estresse hídrico, com um aumento predominante de transcritos durante o estresse severo. Níveis de expressão elevados do ERF *CitERF* em condições de seca também foram observados em plantas de citrus, mostrando que ele pode estar envolvido nos mecanismos de tolerância a seca (YANG et al, 2011). Baseado no padrão de expressão dos *CaERFs* pode-se observar que houve um padrão distinto em alguns grupos, quando submetidos ao

estresse hídrico. *CaERF22* e *CaERF23* ambos pertencentes ao grupo VIII, exibiram um aumento pronunciado no acúmulo de transcritos durante o estresse severo, decrescendo quando a planta foi reidratada (Figura 5). Zhang et al, (2010) mostrou que a expressão constitutiva de *GmERF4* (membro do grupo VIII) em plantas transgênicas de tabaco, promoveu um aumento da tolerância aos estresses salino e a seca, porém suscetíveis a infecção bacteriana. Sugere-se que membros deste grupo podem estar envolvidos na regulação nos mecanismos de resposta ao estresse hídrico levando a um possível aumento da tolerância à seca. Outro tipo de padrão foi observado no grupo IX. Em *CaERF27*, *CaERF29* e *CaERF31* observou-se um aumento transcricional acentuado em plantas submetidas ao déficit hídrico severo e este aumento se manteve na recuperação após o estresse (Figura 5). Fukao et al, (2011) observou que genótipos que tiveram a introgressão de *SUB1*, um fator de transcrição ERF encontrado em arroz, melhorou a recuperação de plantas na fase vegetativa submetidas a períodos de seca prolongados, através da redução da perda de água pela folha, da peroxidação lipídica e aumento da expressão de genes associados com aclimatação a desidratação. Oh et al, (2009) demonstrou que a expressão constitutiva de *AP59* pertencente à família AP2/ERF do grupo IXa, aumentou a recuperação de plantas submetidas a períodos de déficits hídricos em arroz. Entretanto, o papel da família ERF na recuperação de plantas foi estudo somente em plantas monocotiledoneas. Neste trabalho, mostrou-se que alguns membros do grupo IX também podem estar envolvidos nos mecanismos de regulação na recuperação de plantas dicotiledoneas como o *C. arabica* após períodos de déficit hídrico.

O perfil de expressão de mRNA dos 13 ERFs de *C. arabica* sob o estresse salino demonstrou que estes foram diferencialmente induzidos. Dentro da subfamília DREB, somente *CaERF09* foi *up-regulated*, onde se observou um aumento de transcritos 6 dias após o início do tratamento. Huang;Liu (2006) mostrou que a expressão do gene *GhDBP1* foi induzido sob condições de estresse osmótico, como seca e alta salinidade. A expressão estesse induzido de *GmDREB1*, um ortólogo *DREB* da soja, conferiu tolerância a salinidade em plantas transgênicas de alfafa através do aumento do nível de transcrição de P5CS, o que consequentemente, levou a um maior acúmulo de prolina livre (JIN et al, 2010a). Singh et al, (2000) mostrou que sob condições de estresse salino, a prolina livre tem as funções de ajuste osmótico, proteção de macromoléculas celulares e eliminação

de radicais hidroxila. Esses resultados sugerem que esses DREB/ERFs podem estar envolvidos na regulação da transdução de sinal de diferentes vias responsivos ao estresse.

Apesar de *CaERF22* e *CaERF23* pertencerem ao mesmo grupo (grupo VIII) e ambos terem sido induzidos pelo estresse hídrico, estes apresentaram padrões de expressão distintos quando submetidos ao estresse salino. *CaERF22* não foi induzido durante o estresse salino, enquanto que em *CaERF23* observou-se um aumento no acúmulo de transcritos quatro dias após o estresse. Esses resultados sugerem que expressão desses CaERFs são diferentemente induzidos pelos estresses hídricos e salino. Todos os membros do grupo IX foram induzidos pelo estresse salino, de forma similar ao observado no estresse hídrico. A máxima atividade transcricional ocorreu dois e seis dias após o início do tratamento para os genes *CaERF30* e *CaERF29*, respectivamente. Enquanto que em *CaERF27* e *CaERF31* o máximo de acúmulos de transcritos ocorreu 12 dias após o início do estresse. Sharoni et al, (2011) demonstrou que dos 19 genes identificados em arroz pertencentes ao grupo IX, sete foram *up-regulated* em resposta a estresses abióticos. Os genes desse grupo foram responsivos a mais de um tipo de estresse. Similarmente, as quatro CaERFs aqui estudadas também foram induzidos pelos estresses hídrico e salino.

Somente seis ERFs foram diferentemente regulados durante o estresse térmico, um número bem abaixo do que encontrado durante os estresses hídrico e salino. Dentre os DREBs, *CaERF01* e *CaERF06* foram *up-regulated*. Em *CaERF01* foi observado um acúmulo de transcritos a partir do primeiro dia de tratamento e este se manteve até o 5º dia de estresse térmico, enquanto que em *CaERF06* o acúmulo de transcritos se deu de forma gradual no início do estresse, atingindo o máximo 5 dias após o início do estresse. *CaERF06* foi induzido somente pelo estresse térmico (calor) , enquanto que *CaERF01* foi induzido pelos estresses hídrico e térmico. *CaERF09* foi *down-regulated* de forma similar ao observado no estresse hídrico. Embora análises *in silico* tenham mostrado que os genes contendo DRE/CRT podem ser induzidos por choque térmico, poucos trabalhos foram realizados nesse sentido (GEISLER et al, 2006). Sakuma et al, (2006) mostrou que a indução da expressão ao estresse térmico de HsfA3 em *Arabidopsis* depende de *DREB2A*. Mais tarde, Chen et al, (2010) demonstraram que HsfA3 também é transcricionalmente ativado durante o choque térmico por *DREB2C*. Em plantas

transgênicas de *Arabidopsis*, a superexpressão constitutivamente de *DREB2A* de *Arabidopsis* e *ZmDREB2A* de milho ativou a expressão de muitos genes induzidos por estresse ao calor, tais como proteínas de choque térmico (Hsps) e HsfA3 cuja indução resultou em uma maior termotolerância (QIN et al, 2007; SCHRAMM et al, 2008). Desta forma, *CaERF01* e *CaERF06* podem desempenhar um papel essencial como fatores de transcrição da subfamília DREB na sinalização em resposta ao estresse térmico em *C. arabica*.

Na subfamília ERF, somente *CaERF19*, *CaERF30* e *CaERF31* foram *up-regulated* sob o estresse térmico. Membros do grupo IX, *CaERF30* e *CaERF31* foram induzidos pelos três estresses abióticos estudados. Ambos se expressaram de forma similar quando submetidos ao estresse térmico, onde se observou um aumento de transcritos gradual no início do estresse, com acentuado acúmulo no 5º dia de estresse. Assim como para subfamília DREB, poucos trabalhos foram realizados a fim de estudar a sinalização da subfamília ERF em resposta ao calor. Estes resultados são similares ao encontrado por Sharma et al, (2010). Este autor mostrou que dos 16 membros da subfamília ERF em tomate, pertencentes aos grupos V, VI, VII, VIII, IX e X foram *down-regulated* quando submetidos ao estresse ao calor, enquanto que, somente *SlERF42* pertencente ao grupo XI foi *up-regulated* duas vezes mais que o controle sem tratamento térmico. Sendo assim, mostrou-se nesse trabalho que alguns membros da subfamília ERF podem desempenhar um importante papel em resposta ao calor em *C. arabica*.

Pertencente ao grupo VII, *CaERF19* foi diferencialmente regulado em níveis de mRNA em *C. arabica* quando submetidas aos estresses hídrico, salino e térmico. O padrão de expressão foi analisado por semiquantitativa-PCR e por RTq-PCR durante os três estresses, com resultados semelhantes. A expressão do gene *CaERF19* foi *up-regulated* 67, 21 e 3 vezes por RTq-PCR; enquanto que, por semiquantitativa PCR foi *up-regulated* 5, 3 e 2 vezes para os estresses térmico, salino e hídrico, respectivamente. Os resultados observados foram similares nas duas técnicas, entretanto a análise dos estresses térmico e salino apresentaram uma maior diferença nos níveis de expressão relativa. Embora esse grupo esteja associado à maturação de frutos (TOURNIER et al, 2003; EL-SHARKAWY et al, 2009) trabalhos tem mostrado que ERFs do grupo VII podem estar envolvidos nas respostas aos diferentes estresses. Níveis elevados de expressão foi observado nos ERFs *GhERF2*, *GhERF3* e *GhERF6* em algodão quando as plantas foram

submetidos aos estresses abióticos como o salino, o frio e a seca (JIN et al, 2010b). A expressão ectópica do gene *GmERF3* em plantas transgênicas de tabaco aumentou a tolerância à salinidade e seca. Além disso, a superexpressão de *GmERF3* em plantas de tabaco transgênico levou a níveis elevados de prolina livre e carboidratos solúveis em plantas sob condições de seca (ZHANG et al, 2009). Neste trabalho, vários ERFs foram induzidos por mais de um tipo de estresse. A indução de genes da família ERF em resposta a mais de um estresse é mais uma indicação da existência de respostas cruzadas (crosstalk) entre as diferentes vias de sinalização dos estresses. O acúmulo de transcritos do gene *CaERF04* não foi alterado nos estresses hídricos, salino e térmico. O padrão de expressão constitutivo de *CaERF04* durante todo o período dos três estresses estudados, indica que este pode ser um gene candidato para ser utilizado como normalizador em estudos de expressão em plantas submetidas a estresses abióticos. Entretanto futuros trabalhos devem ser realizados para confirmar o padrão transcricional desse gene.

Através da caracterização *in silico* e *in vivo* de 13 ERFs neste trabalho foi possível identificar genes que devem desempenhar um papel essencial na regulação de genes de resposta aos estresses abióticos em *C. arabica*. O controle da regulação pela família ERF e os seus genes alvos é fundamental para o entendimento de como as plantas de cafeeiro respondem aos diferentes estresses ambientais. Desta forma, os dados gerados neste trabalho fornecem subsídios para selecionar os melhores candidatos para futuras análises funcionais da família ERF em *C. arabica*, no qual ajudará a entender os determinantes genéticos da tolerância a estresses abióticos. Isso constituirá um passo fundamental para os programas de melhoramento genético do cafeeiro que têm como objetivo aumentar a qualidade e a produção de frutos de café e a obtenção de um sistema agrícola mais sustentável.

5.7 AGRADECIMENTOS

O primeiro autor agradece às agências financiadoras, CAPES e CNPq pela concessão das bolsas de doutorado no exterior e no país, respectivamente.

5.8 REFERENCIA

ASSAD, E. D.; PINTO, H. S.; ZULLO, JUNIOR, J.; AVILA, A. M. H. Impacto das mudanças climáticas no zoneamento agroclimático no café no Brasil. **Pesq. Agropec. Bras.**, v.39, n.11, p. 1057-1064. 2004.

BANNO, H.; IKEDA, Y.; NIU Q. W.; CHUA, N. H. Overexpression of Arabidopsis ESR1 induces initiation of shoot regeneration. **The Plant Cell**, v.13, p. 2609–2618. 2001.

BOUTILIER, K, OFFRINGA, R.; SHARMA, V. K, et al. Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. **The Plant Cell**, v.14, p. 1737–1749. 2002.

BROWN, R. L, KAZAN, K.; MCGRATH, K. C.; MACLEAN, D. J.; MANNERS, J. M. A role for the GCC-box in jasmonate-mediated activation of the PDF1.2 gene of Arabidopsis. **Plant Physiol.**, v.132, p. 1020–1032. 2003.

CAO, Y, SONG F.; GOODMAN, R. M.; ZHENG, Z. Molecular characterization of four rice genes encoding ethylene-responsive transcriptional factors and their expressions in response to biotic and abiotic stress. **J. Plant. Physiol.**, v.163, p. 1167–1178. 2006.

CAO, Z. F.; LI, J.; CHEN, F.; LI, Y. Q.; ZHOU, H. M.; LIU, Q. Effect of two conserved amino acid residues on DREB1A function. **Biochemistry**, v.66, p. 623–627. 2001.

CHAMPION, A.; HEBRARD, E.; PARRA, B.; BOURNAUD, C.; MARMEY, P.; TRANCHANT, C.; NICOLE, M. Molecular diversity and gene expression of cotton ERF transcription factors reveal that group IXa members are responsive to jasmonate, ethylene and Xanthomonas. **Molecular Plant Pathology**, v.10, n.4, p. 471-485. 2009.

CHANG, S.; PURYEAR, J.; CAIRNEY, J. A simple and efficient method for isolating RNA from pine trees. **Plant Molecular Biology**, v.11, p.113-116, 1993.

CHEN, H.; HWANG, J. E.; LIM, C. J.; KIM, D. Y.; LEE, S. Y.; LIM, C. O. Arabidopsis DREB2C functions as a transcriptional activator of HsfA3 during the heat stress response. **Biochemical and Biophysical Research Communications**, v.401 p. 238–244. 2010.

CHEN, W. J.; ZHU, T. Networks of Transcription Factors with Roles in Environmental Stress Response, **Trends Plant Sci.**, v.9, p. 591–596. 2004.

CIFTCI-YILMAZ, S.; MORSY, M. R.; SONG, L.; COUTU, A.; KRIZEK, B. A.; LEWIS, M. W.; WARREN, D.; CUSHMAN, J.; CONNOLLY, E. L.; MITTLER, R. The EAR-motif of the Cys2/His2-type zinc finger protein Zat7 plays a key role in the defense response of Arabidopsis to salinity stress. **J. Biol. Chem.**, v.282, p. 9260-9268. 2007.

- DAMATTA, F. M. Exploring drought tolerance in coffee: a physiological approach with some insights for plant breeding. **Brazilian Journal of Plant Physiology**, v.16, n.1, p. 1-6. 2004.
- DAMATTA, F. M.; RAMALHO, J. D. C. Impacts of drought and temperature stress on coffee physiology and production: a review, **Brazilian Journal of Plant Physiology**, v.18, n.1, p. 55-81, 2006.
- DONG, C. J.; LIU, J. Y. The Arabidopsis EAR-motif-containing protein RAP2.1 functions as an active transcriptional repressor to keep stress responses under tight control. **BMC Plant Biology**, v.10, p. 47. 2010.
- DUBOUZET, J. G.; SAKUMA, Y.; ITO, Y.; KASUGA, M.; DUBOUZET, E. G.; MIURA, S.; SEKI, M.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Os-DREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. **Plant J.** v.33, p. 751–763. 2003.
- ELLIOTT, R. C.; BETZNER, A. S.; HUTTNER, E.; OAKES, M. P.; TUCKER, W. Q.; GERENTES, D. et al. AINTEGUMENTA, an APETALA2-like gene of Arabidopsis with pleiotropic roles in ovule development and floral organ growth. **Plant Cell**, v.8, p. 155–168. 1996.
- EL-SHARKAWY, I.; SHERIF, S.; MILA, I.; BOUZAYEN, M.; JAYASANKAR, S. Molecular characterization of seven genes encoding ethylene-responsive transcriptional factors during plum fruit development and ripening. **J. Exp. Bot.**, v.60, p. 907–922. 2009.
- FERREIRA, D. F. **Sistema de análise de variância (Sisvar)**. Versão 4.6. Lavras: Universidade Federal de Lavras, 1999. (CD-ROM).
- Freschi, L., Nievola, C.C., Rodrigues, M.A., Domingues, D.S., Sluys, M.A.V., Mercier, H., Thermoperiod affects the diurnal cycle of nitrate reductase expression and activity in pineapple plants by modulating the endogenous levels of cytokinins. **Physiologia Plantarum**, v.137, p. 201–212. 2009.
- FUJIMOTO, S. Y.; OHTA, M.; USUI, A.; SHINSHI, H.; OHME-TAKAGI, M. Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. **Plant Cell**, v.12, p. 393–404. 2000.
- FUKAO, T.; YEUNG, E.; BAILEY-SERRES, J. The Submergence Tolerance Regulator SUB1A Mediates Crosstalk between Submergence and Drought Tolerance in Rice. **The Plant Cell**, v.23, p. 412–427. 2011.
- GEISLER, M.; KLECZKOWSKI, L. A.; KARPINSKI, S. A universal algorithm for genome-wide in silico identification of biologically significant gene promoter putative cis-regulatory-elements; identification of new elements for reactive oxygen species and sucrose signaling in Arabidopsis, **Plant J.**, v. 45, p. 384–398. 2006.
- GUO, Z. J.; CHEN, X. J.; WU, X. L.; LING, J. Q.; XU, P. Overexpression of the AP2/EREBP transcription factor OPBP1 enhances disease resistance and salt tolerance in tobacco. **Plant Mol Biol.**, v.55, p. 607–618. 2004.

GUTTERSON, N.; REUBER, T. L. Regulation of disease resistance pathways by AP2/ERF transcription factors. **Curr. Opin. Plant Biol.**, v.7, p. 465–471. 2004.

GU, Y. Q.; WILDERMUTH, M. C.; CHAKRAVARTHY, S.; LOH, Y. T.; YANG, C.; HE, X.; HAN, Y.; MARTIN, G. B. Tomato transcription factors pti4, pti5, and Pti6 activate defense responses in Arabidopsis. **Plant Cell**, v.14, p. 817-831. 2002.

HONG-BO, S.; ZONG-SUO, L.; MING-AN, S. LEA proteins in higher plants: structure, function, gene expression and regulation. *Colloids Surf B: Biointerfaces*, v.45, p. 131–5. 2005.

HUANG, B.; LIU, J. Y. Cloning and functional analysis of the novel gene GhDBP3 encoding a DRE-binding transcription factor from *Gossypium hirsutum*. *Biochimica et Biophysica Acta*, v.1759, p. 263–269. 2006.

HUANG, B.; JIN, L.; LIU, J. Y. Identification and characterization of the novel gene GhDBP2 encoding a DRE-binding protein from cotton (*Gossypium hirsutum*). **Journal of Plant Physiology**, v.165, p. 214-223. 2008.

JIN, T.; CHANG, Q.; LI, W.; YIN, D.; LI, Z.; WANG, D.; LIU, B.; LIU, L. Stress-inducible expression of GmDREB1 conferred salt tolerance in transgenic alfalfa. **Plant Cell Tiss Organ Cult.**, v.100, p. 219–227. 2010a.

JIN, L. G.; LI, H.; LIU, J. Y. Molecular Characterization of Three Ethylene Responsive Element Binding Factor Genes from Cotton. **Journal of Integrative Plant Biology**, v.52, n.5, p. 485–495. 2010b.

KRISHNASWAMY, S.; VERMA, S.; RAHMAN, M. H.; KAV, N. N. V. Functional characterization of four APETALA2-family genes (RAP2.6, RAP2.6L, DREB19 and DREB26) in Arabidopsis. **Plant Mol Biol.**, v.75, p. 107–127. 2011.

ICO. **International Coffee Organization**. Disponível em:
<<http://www.ico.org/prices/po.htm>> Acesso em: 20 mai. 2011.

LARKIN, M. A.; BLACKSHIELDS, G.; BROWN, N. P.; CHENNA, R.; MCGETTIGAN, P. A.; MCWILLIAM, H.; VALENTIN, F.; WALLACE, I. M.; WILM, A.; LOPEZ, R.; THOMPSON, J. D.; GIBSON, T. J.; HIGGINS, D. G. Clustal W and Clustal X version 2.0. **Bioinformatics**, v.23, p. 2947-2948. 2007.

LICAUSI, F.; GIORGI, F. M.; ZENON, S.; OSTI, F.; PEZZOTTI, M.; PERATA, P. Genomic and transcriptomic analysis of the AP2/ERF superfamily in *Vitis vinifera*. **BMC Genomics**, v.11, p. 719. 2010.

LIN, R. C.; PARK, H. J.; WANG, H. Y. Role of Arabidopsis RAP2.4 in regulating light- and ethylene-mediated developmental processes and drought stress tolerance. **Molecular Plant**, v.1, n.1, p. 42-57. 2008.

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

LIU, W. Y.; CHIOU, S. J.; KO, C. Y.; LIN, T. Y. Functional characterization of three ethylene response factor genes from *Bupleurum kaoi* indicates that BkERFs mediate resistance to *Botrytis cinerea*. **Journal of Plant Physiology**, v.168, p. 375-381. 2011.

LIU, Y.; CHEN, H.; ZHUANG, D.; JIANG, D.; LIU, J.; WU, G.; YANG, M.; SHEN, S. Characterization of a DRE-binding transcription factor from asparagus (*Asparagus officinalis* L.) and its overexpression in arabidopsis resulting in salt- and drought-resistant transgenic plants. **International Journal of Plant Sciences**. v.171, p. 12-23. 2010.

LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. **Methods**, v.25, p. 402-408. 2001.

LU, C. W.; SHAO, Y.; LI, L.; CHEN, A. J.; XU, W. Q.; WU, K. J.; LUO, Y. B.; ZHU, B. Z. Overexpression of SIERF1 Tomato Gene Encoding an ERF-Type Transcription Activator Enhances Salt Tolerance. **Russian Journal of Plant Physiology**, v.58, n. 1, p. 118–125. 2011.

KASUGA, M.; LIU, Q.; MIURA, S.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. **Nature Biotechnology**, v.17, n.3, p. 287–291. 1999.

KIZIS, D.; PAGES, M. Maize DRE-binding proteins DBF1 and DBF2 are involved in rab17 regulation through the drought-responsive element in an ABA-dependent pathway. **Plant J.** v.30, p. 679–689. 2002.

NAKANO, T.; SUZUKI, K.; FUJIMURA, T.; SHINSHI, H. Genome-wide analysis of the ERF gene family in Arabidopsis and rice. **Plant Physiol.**, v.140, p. 411–432. 2006.

OH, S.-J.; KIM, Y. S.; KWON, C.-W.; PARK, H. K.; JEONG, J.S.; KIM, J.-K. Overexpression of the transcription factor AP37 in rice improves grain yield under drought conditions. **Plant Physiol.**, v.150, p. 1368–1379. 2009.

OHTA, M.; MATSUI, K.; HIRATSU, K.; SHINSHI, H.; OHME-TAKAGI, M. Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. **The Plant Cell**, v.13, p. 1959–1968. 2001.

ONATE-SANCHEZ, L.; ANDERSON, J. P.; YOUNG, J.; SINGH, K. B. AtERF14, a member of the ERF family of transcription factors, plays a nonredundant role in plant defense. **Plant Physiol.**, v.143, p. 400–409. 2007.

PAY, E. **The market for organic and fair-trade coffee**. FAO Rome. 2009.

PARK, J. M.; PARK, C. J.; LEE, S. B.; HAM, B. K.; SHIN, R.; PAEK, K. H. Overexpression of the tobacco Tsi1 gene encoding an EREBP/ AP2-type transcription factor enhances resistance against pathogen attack and osmotic stress in tobacco. **Plant Cell**, v.13, p. 1035–1046. 2001.

PIRRELLO, J.; JAIMES-MIRANDA, F.; SANCHEZ-BALLESTA, M. T.; TOURNIER, B.; KHALIL-AHMAD Q.; REGAD, F.; LATCHE, A.; PECH, J. C.; BOUZAYEN, M. SIERF2, a tomato ethylene response factor involved in ethylene response and seed germination. **Plant Cell Physiol**, v.47, p. 1195–1205. 2006.

QIN, F.; KAKIMOTO, M.; SAKUMA, Y.; MARUYAMA, K.; OSAKABE, Y.; TRAN, L. S.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Regulation and functional analysis of ZmDREB2A in response to drought and heat stresses in *Zea mays* L, **Plant J.**, v.50, p. 54–69. 2007.

MAPA. **Ministério da Agricultura, Pecuária e Abastecimento**. Disponível em: <http://www.agricultura.gov.br/arq_editor/file/3894_relatorio_do_mercado_cafeeiro_-_marco_2011.pdf>. Acesso em: 05 abr. 2011.

MATSUI, K.; OHME-TAKAGI, M. Detection of protein–protein interactions in plants using the transrepressive activity of the EAR motif repression domain. **The Plant Journal**, v.61, p. 570–578. 2010.

McGRATH, K. C.; DOMBRECHT, B.; MANNERS, J. M.; SCHENK, P. M.; EDGAR, C. I, MACLEAN D. J.; SCHEIBLE, W. R.; UDVARDI, M. K.; KAZAN, K. Repressor- and activator-type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genomewide screen of Arabidopsis transcription factor gene expression. **Plant Physiology**, v.139, p. 949–959. 2005.

SAKUMA, Y.; LIU, Q.; DUBOUZET, J. G.; ABE, H.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. **Biochem. Biophys. Res. Commun.**, v.290, p. 998–1009. 2002.

SAKUMA, Y.; MARUYAMA, K.; QIN, F.; OSAKABE, Y.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Dual function of an Arabidopsis transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression, **Proc. Natl. Acad. Sci. USA.**, v.103, p. 18822–18827. 2006.

SHARMA, M. K.; KUMAR, R.; SOLANKE, A. U.; SHARMA, R.; TYAGI, A. K.; SHARMA A. K. Identification, phylogeny, and transcript profiling of ERF family genes during development and abiotic stress treatments in tomato. **Mol. Genet. Genomics.**, v.284, n.6, p. 455-475. 2010.

SHARONI, A. M, NURUZZAMAN, M.; SATOH, K.; SHIMIZU, T.; KONDOH, H.; SASAYA, T.; CHOI, I. R.; OMURA, T.; KIKUCHI, S. Gene Structures, Classification and Expression Models of the AP2/EREBP Transcription Factor Family in Rice. **Plant Cell Physiol.**, v.52, n.2, p. 344–360. 2011.

SCHRAMM, F.; LARKINDALE, J.; KIEHLMANN, E.; GANGULI, A.; ENGLICH, G, VIERLING, E.; VON KOSKULL-DÖRING, P. A cascade of transcription factor DREB2A and heat stress transcription factor HsfA3 regulates the heat stress response of Arabidopsis, **Plant J.**, v.53, p. 264–274. 2008.

SINGH, S. K.; SHARMA, H. C.; GOSWAMI, A. M.; DATTA, S. P.; SINGH, S. P. In vitro growth and leaf composition of grapevine cultivars as affected by sodium chloride. **Biol. Plant.**, v.43, p. 283–286. 2000.

SOLANO, R.; STEPANOVA, A.; CHAO, Q.; ECKER J. R. Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. *Genes and Development* v.12, p. 3703–3714. 1998.

TAMURA, K.; DUDLEY, J.; NEI, M.; KUMAR, S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.*, v.24, p. 1596–1599. 2007.

TOURNIER, B.; SANCHEZ-BALLESTA, M. T.; JONES, B.; PESQUET, E.; REGAD, F.; LATCHE, A.; PECH, J. C.; BOUZAYEN, M. New members of the tomato ERF family show specific expression pattern and diverse DNA-binding capacity to the GCC box element. *FEBS Lett.*, v.550, p. 149–154. 2003.

UMEZAWA, T.; FUJITA, M.; FUJITA, Y.; YAMAGUCHI_SHINOZAKI, K.; SHINOZAKI, K. Engineering Drought Tolerance in Plants: Discovering and Tailoring Genes to Unlock the Future, *Curr. Opin. Biotechnol.*, v.17, p. 113–122. 2006.

WANG, K. L.; LI, H.; ECKER, J. R. Ethylene biosynthesis and signaling networks. *Plant Cell*, v.14, p. 131–151. 2002.

WANG, A.; TAN, D.M.; TAKAHASHI, A.; LI, T.Z.; HARADA, T. MdERFs, two ethylene-response factors involved in apple fruit ripening. *J. Exp. Bot.* v.58, p. 3743–3748. 2007.

WANG, X.; CHEN, X.; LIU, Y.; GAO, H.; WANG, Z.; SUN, G. CkDREB gene in *Caragana korshinskii* is involved in the regulation of stress response to multiple abiotic stresses as an AP2/EREBP transcription factor. *Mol. Biol. Rep.*, v.38, p. 2801–2811. 2011.

XIONG, L.; ZHU, J. K. Regulation of abscisic acid biosynthesis. *Plant Physiol.*, v.33, p. 29–36. 2003.

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. Isolation and molecular characterization of the *Triticum aestivum* L. ethylene-responsive factor 1 (TaERF1) that increases multiple stress tolerance. *Plant Mol. Biol.*, v.65, p. 719–732. 2007.

XU, H.; WANG, X. C.; CHEN, J. Overexpression of the Rap2.4f transcriptional factor in *Arabidopsis* promotes leaf senescence. *Science China-Life Sciences*. v.53, n. 10, p. 1221–1226. 2010.

YANG, X. Y.; XIJX, LUA, X. P.; LIUYZ, PENG, S. A. Isolation of a citrus ethylene responsive element binding factor gene and its expression in response to abiotic stress, girdling and shading. *Scientia Horticulturae*, v.127, p. 275–281. 2011.

YANG, Y.; WU, Y.; PIRRELLO, J.; REGAD, F.; BOUZAYEN, M.; DENG, W.; LI, Z. Silencing SI-EBF1 and SI-EBF2 expression causes constitutive ethylene response phenotype, accelerated plant senescence, and fruit ripening in tomato. *Journal of Experimental Botany*, v.61, n.3, p. 697–708, 2010.

YANG, Z.; TIAN, L.; LATOSZEK-GREEN, M.; BROWN, D.; WU, K. Arabidopsis ERF4 is a transcriptional repressor capable of modulating ethylene and abscisic acid responses. **Plant Molecular Biology**, v.58, p. 585–596. 2005.

YI, S. Y.; KIM, J. H.; JOUNG, Y. H.; LEE, S.; KIM, W. T.; YU, S. H.; CHOI, D. The pepper transcription factor CaPF1 confers pathogen and freezing tolerance in Arabidopsis. **Plant Physiol.**, v.136, p. 2862–2874. 2004.

ZHANG, H.; HUANG, Z.; XIE, B.; CHEN, Q.; TIAN, X.; ZHANG, X.; LU, X.; HUANG, D.; HUANG, R. The ethylene-, jasmonate-, abscisic acid- and NaCl-responsive tomato transcription factor JERF1 modulates expression of GCC box-containing genes and salt tolerance in tobacco. **Planta**, v.220, p. 262–270. 2004.

ZHANG, J. Y.; BROECKLING, C. D.; BLANCAFLOR, E. B.; SLEDGE, M. K.; SUMNER, L. W.; WANG, Z. Y. Overexpression of WXP1, a putative *Medicago truncatula* AP2 domain-containing transcription factor gene, increases cuticular wax accumulation and enhances drought tolerance in transgenic alfalfa (*Medicago sativa*). **Plant J.**, v.42, p. 689–707. 2005.

ZHANG, G.; CHEN, M.; LI, L.; XU, Z.; CHEN, X.; GUO, J.; MA, Y. Overexpression of the soybean GmERF3 gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. **Journal of Experimental Botany**, v.60, n.13, p. 3781–3796, 2009.

ZHANG, G.; CHEN, M.; CHEN, X.; XU, Z.; LI, L.; GUO, J.; MA, Y. Isolation and characterization of a novel EAR-motif-containing gene GmERF4 from soybean (*Glycine max* L.) **Mol. Biol. Rep.**, v.37, p. 809–818. 2010.

ZHUANG, J.; CAI, B.; PENG, R. H.; ZHU, B.; JIN, X. F.; XUE, Y.; GAO, F.; FU, X. Y.; TIAN, Y. S.; ZHAO, W.; QIAO, Y. S.; ZHANG, Z.; XIONG, A. S.; YAO, Q. H. Genome-wide analysis of the AP2/ERF gene family in *Populus trichocarpa*. **Biochem Biophys. Res. Commun.**, v.371, p. 468–474. 2008.

6 CONCLUSÕES GERAIS

Foram identificados 36 ERFs em *C. arábica*, classificados em nove grupos, dos quais 12 pertencem a subfamília DREBe 24 à subfamília ERF.

Níveis elevados de expressão foram observados em bibliotecas derivadas de tecidos de frutos, folhas e flores, e em bibliotecas submetidas ao estresse hídrico.

A análise da atividade transcricional dos 13 ERFs exibiu padrões e níveis de expressão diferencial em resposta aos estresses hídrico, salino e térmico.

Os grupos VIII e IX apresentaram padrões distintos de expressão no estresse hídrico. As ERFs do grupo VIII (*CaERF22* e *CaERF23*) apresentaram um aumento pronunciado de transcritos quando a planta foi submetida ao estresse hídrico severo, decrescendo no recuperado. As ERFs do grupo IX (*CaERF27* e *CaERF29* e *CaERF31*) apresentaram um aumento pronunciado de transcritos quando a planta foi submetida ao estresse hídrico severo e este se manteve na recuperação após o estresse.

Os ERFs foram induzidos por mais de um tipo de estresse. A resposta a mais de um estresse é uma indicação da existência de respostas cruzadas (crosstalk) entre as diferentes vias de sinalização dos estresses.

O padrão de expressão estável e constitutiva de *CaERF04* nos estresses hídrico, salino e térmico, indica que este pode ser um candidato para ser utilizado como normalizador em estudos de expressão em plantas submetidas a estresses abióticos.

APÊNDICES

APÊNDICE A – Alinhamento dos domínios AP2/ERF de *Coffea arabica* e *Arabidopsis thaliana*

CLUSTAL 2.1 multiple sequence alignment

```

At4g11140 -----
At4g23750 -----
CaERF13 -----
At5g53290 -----
At4g27950 -----
At2g46310 -----
AT2G20880 -----MATAKNKGKSIKRVLTSEAEKKDEMELEEEEFQFSSGKYKDSGPGSDMWLGDAS 53
AT4G28140 -----MDFDEELNLCITKGKNVDHS---FGGEAS 26
AT1g22190 -----MTTSMDFYSN-KTFQQSDPFGGELM 24
AT1g78080 -----MAAAMNLYTCSRFSQDS---GGELM 22
AT1G36060 -----MADLFGG-----GHGGELM 14
CaERF01 -----
CaERF02 -----
CaERF05 -----MEALEPFMKSGASSPSYSSSYNPQSFSTASPSFFPSM 38
AT2g22200 -----
AT4g39780 -----MAAIDMFNSNTDPFQEELMKALQPY 25
CaERF04 -----MKALEPFMKGASSLPSSPTTTSSPF 25
CaERF03 -----
AT5g65130 -----MALNMNAYVDEFMEALEPFMKVTSSSSTSN 30
At4g31060 -----
At2g23340 -----
At4g36900 -----
At5g67190 -----
At3g50260 -----
CaERF06 -----
At4g06746 -----
At1g46768 -----
At1g21910 -----
At1g77640 -----
At1g44830 -----
At1g22810 -----
At1g71520 -----
At1g19210 -----

```

At1g74930 -----
At5g21960 -----
At2g40220 -----
At1g71450 -----
At1g33760 -----
At1g63040 -----
At1g12630 -----
At5g52020 -----
CaERF08 -----
CaERF07 -----
At2g36450 -----
At4g25470 -----
At4g25480 -----
At4g25490 -----
At5g51990 -----
At1g63030 -----
At1g12610 -----
At3g60490 -----
CaERF09 -----
At2g35700 -----
At4g16750 -----
At2g44940 -----
At5g25810 -----
At5g11590 -----
At2g25820 -----
At4g32800 -----
At1g77200 -----
At3g16280 -----
At3g11020 -----
At5g05410 -----
CaERF10 -----
At2g40350 -----
At2g40340 -----
CaERF11 -----
At2g38340 -----
At1g75490 -----
At5g18450 -----
At3g57600 -----
CaERF12 -----
At5g11190 -----

At5g25390	-----	
At1g15360	-----	
At5g25190	-----	
At5g19790	-----	
At3g15210	-----	
At1g53170	-----	
At1g28370	-----	
CaERF22	-----	
CaERF23	-----	
CaERF25	-----	
At1g03800	-----	
At5g44210	-----	
At1g50640	-----	
At3g20310	-----	
CaERF24	-----	
CaERF26	-----	
At1g28360	-----	
At1g12980	-----	
At5g13910	-----	
At5g18560	-----	
At1g28160	-----	
At1g12890	-----	
At1g24590	-----	
At2g47520	-----	
At1g72360	-----	
CaERF14	-----	
CaERF15	-----	
CaERF16	-----	
At3g16770	-----	
CaERF17	-----	
CaERF18	-----	
At1g53910	-----	
At3g14230	-----	
CaERF19	-----	
CaERF20	-----	
CaERF21	-----	
At5g61890	-----MYGKRPFGGDESEER---	EEDEN 20
At5g07310	-----MANSNGYKRPFRGDESDEKKEADDEN	28
CaERF36	-----	
At5g13330	-----	

At1g43160	-----	
At5g50080	-----	
CaERF37	-----	
At5g64750	MCVLKVANQEDNVGKKAESIRDDDHRTLSEIDQWLYLFAAEDDHRHSFPTQQPPSSSS	60
At2g33710	-----	
At2g31230	-----	
At1g06160	-----	
At3g23240	-----	
At5g43410	-----	
At1g04370	-----	
At3g23220	-----	
At3g23230	-----	
At4g18450	-----	
At4g17500	-----	
At5g47220	-----	
CaERF27	-----	
CaERF28	-----	
CaERF29	-----	
At2g44840	-----	
CaERF30	-----	
CaERF31	-----	
At5g47230	-----	
At4g17490	-----	
At5g61600	-----	
At5g51190	-----	
CaERF32	-----	
CaERF34	-----	
CaERF33	-----	
At5g07580	---MTPSLKPLRERQNHSEFFVSLPHPWLLKSCDYLLVSFLFSFPLLYKNQNIHSFCSII	57
At5g61590	-----	
At1g80580	-----	
CaERF35	-----	
At2g20350	-----	
At5g67000	-----	
At5g67010	-----	
At1g25470	-----	
At1g68550	-----	
At3g25890	-----	

At4g11140	-----METEKKVSLP-----	10
At4g23750	-----MEAEKKMVLPRIKFTE----HK	18
CaERF13	-----MEIGISPQSPMYSRQRVKFSEHVVTNKPIDD	32
At5g53290	-----MDEYIDFRPLKYTEHKTSMTKYTKKSS-----	27
At4g27950	-----MMMDEFMDLRPVKYTEHKTVIRKYTKKSSM-----	30
At2g46310	-----MKSRRVRKSKYTVHRKITS-----	18
AT2G20880	STSPRSLRKTTRTFDRHNPYLVSYYATPQPPTTTTCSVSFPFYLPPIQN-----	102
AT4G28140	STSPRSMKMKSPSRPKPYFQSS-----SSPYSLEAFPFSLDPTLQN-----	68
AT1g22190	EALLPFIK----SPSNDSSAFA-----FSLP-----	46
AT1g78080	DALVPFIKSVSDSPSSSSAASASAFSLHPSAFSLPPLPGYY-----	62
AT1G36060	EALQPFYKSASTSASNPAFASNDAFASAPNDLFSSSSYNPHASLFP-----	62
CaERF01	-----MYPESECYSSQSTYGLV-----	16
CaERF02	-----	
CaERF05	DDQFFSFSSSSPSTSYS-APSSFPAQYQAQFNMYPESECYSSQSTYGLV-----	85
AT2g22200	-METASLSF---PVPNTSFGVN-----	18
AT4g39780	TTNTDSSSP---TYSNTVFGFN-----	44
CaERF04	SSSSSSLSP---PSTSSSSSCSSYYPFDSSSFSEPNMYSSFTSYVPTTAHMFSQGLSSF	82
CaERF03	-----	
AT5g65130	SSNPKPLTPNFI PNNDQVL PVS N-----	53
At4g31060	-----	
At2g23340	-----	
At4g36900	-----	
At5g67190	-----	
At3g50260	-----	
CaERF06	-----	
At4g06746	-----	
At1g46768	-----	
At1g21910	-----	
At1g77640	-----	
At1g44830	-----	
At1g22810	-----	
At1g71520	-----	
At1g19210	-----	
At1g74930	-----	
At5g21960	-----	
At2g40220	-----	
At1g71450	-----	
At1g33760	-----	
At1g63040	-----MADPNNPITEPKAIIQSSTSS-SVTIVPVPTCGDS	34
At1g12630	-----	

At5g52020	-----MSNNNN---SPTTVNQETTTTS-REVSITLPTDQSP	31
CaERF08	-----MLESDQPNNSLRPDDDPDDPPQPKQPKPTGIIA	35
CaERF07	-----MLESDQPNNSLRPDDDPDDPPQPKQPKPTGIIA	35
At2g36450	-----	
At4g25470	-----MNSF-SAFSEMFG--SDYESP-----	18
At4g25480	-----MNSF-SAFSEMFG--SDYESS-----	18
At4g25490	-----MNSF-SAFSEMFG--SDYEP-----	17
At5g51990	-----MNPFYSTFPDSFLSISDHRSP-----	21
At1g63030	-----MEND-----	4
At1g12610	-----MNND-----	4
At3g60490	-----MGKQINIESSATHHQDN---IVSV	21
CaERF09	-----MLSLAASTTPPLIIMGKQINIGSLANHDSNLSSSSTTT	38
At2g35700	-----MERDDCRRFQDSPAQTTER-----	19
At4g16750	-----MQDSSSHESQR-----	11
At2g44940	-----MARQINIESSVSQVTFISSAIPAV	24
At5g25810	-----	
At5g11590	-----MAEYYSLRSERVTVQ-----	15
At2g25820	-----	
At4g32800	-----	
At1g77200	-----MTESSII-----	7
At3g16280	-----MTSLNSSASPTSSSSDQSDA-----	20
At3g11020	-----	
At5g05410	-----	
CaERF10	-----	
At2g40350	-----	
At2g40340	-----	
CaERF11	-----	
At2g38340	-----	
At1g75490	-----	
At5g18450	-----	
At3g57600	-----	
CaERF12	-----	
At5g11190	-----	
At5g25390	-----	
At1g15360	-----	
At5g25190	-----	
At5g19790	-----	
At3g15210	-----	
At1g53170	-----	
At1g28370	-----	

CaERF22	-----	
CaERF23	-----	
CaERF25	-----	
At1g03800	-----	
At5g44210	-----	
At1g50640	-----	
At3g20310	-----	
CaERF24	-----	
CaERF26	-----	
At1g28360	-----	
At1g12980	-----	
At5g13910	-----	
At5g18560	-----	
At1g28160	-----	
At1g12890	-----	
At1g24590	-----	
At2g47520	-----MCGGAIISDFIWSKSE-----SEPSQLGS-----	24
At1g72360	-----MCGGAVISDYIAPEKI-----ARSSGKSS-----	24
CaERF14	-----MCGGAIISGG-SPINP---GRKLTPKELWA-----	26
CaERF15	-----MCGGAIISGG-SPINP---GRKLTPKELWA-----	26
CaERF16	-----MCGGAIISGG-SPINP---GRKLTPKELWA-----	26
At3g16770	-----MCGGAIISDY-APLVTKAKGRKLTAEELWS-----	29
CaERF17	-----MCGGAIILEGLIPPRNNNDGTRRLSSADLWP-----	30
CaERF18	-----MC-----	2
At1g53910	-----MCGGAIISDFIPPPRS----RRVTSEFIWP-----DLKKNLKGSK	36
At3g14230	-----MCGGAIISDFIPPPRS----LRVTNEFIWP-----DLKNKVKASK	36
CaERF19	-----MCGGAIISDYIQPSRSS---HRLSADLLWG----IGGSGDLPDSKDIKKK	43
CaERF20	-----	
CaERF21	-----MCGGAIISDFIPPPRSS---RRLTADLLWGTGADFPGTTKDINKKKKNNK	47
At5g61890	LFPVFSARSQHDMRVMVSALTQVIGNQQSKSH-----	52
At5g07310	IFPFFSARSQYDMRAMVSALTQVIGNQSS-SH-----	59
CaERF36	-----MTAMVSALSQVIGSTDN-----	17
At5g13330	-----MVSALSRVIENPTD-----	14
At1g43160	-----MVSMLTNVVSGETEPSA-----	17
At5g50080	-----MSAMVSALTQVVSARSQTE-----AEGAHSSSSSAGHKRWLGID	40
CaERF37	-----MLDEAASVITP-----TTTA-----	15
At5g64750	SSLISGFSREMEMSAIVSALTHVVAGNVPQHQQGGEGSGEGTSNSSSSSGQKRRREVEE	120
At2g33710	MHSGKRPLSPESMAGNREEKELCCCSTLSES-----	32
At2g31230	-----MEYSQ-----SSMYSSPSSWSSS-----	18
At1g06160	-----MEYQTNFLSGEFSPENSSSSSWSS-----	24

At3g23240	-----MDPFLIQSPFSGFSPEYSIGSSPDSF-----	26
At5g43410	-----	
At1g04370	-----	
At3g23220	-----	
At3g23230	-----	
At4g18450	-----MAFGNIQELDGEILKNVWANYIGTPQTDTRSIQVP-----	35
At4g17500	---MSMTADSQSDYAFLESIRRHLG---ESEPILSESTASSVTQSCVTGQSIKPVYGRN	54
At5g47220	---MYGQCNIESDYALLESITRHLGGGGENELRLNESTPSS-----	39
CaERF27	---MYHPISAELPPANSS-----GEFPAYRRSSSFSS-----	29
CaERF28	---MYHPISAELPPANSS-----GEFPAYRRSSSFSS-----	29
CaERF29	---MYHPISAELPPANSS-----GEFPAYRGSCSFSS-----	29
At2g44840	---MSSSDSVNNG-----VNSRMYFRNPSFSN-----	24
CaERF30	-----	
CaERF31	-----	
At5g47230	-----MAT-----PNEVSALWFIKHLLEASPVATDPWMKHESSSATESSSDS	44
At4g17490	-----MAT-----PNEVSALFLIKKYLLDELSPLPT-----TATTNRWMN	35
At5g61600	-----MAT-----KQEALAI DFISQHLITDFVSMETD-----HPSLFTN	34
At5g51190	-----MASSHQQQQEQDQSALDLITQHLLTDFPSLDT-----FAS	35
CaERF32	-----MAS---QTQITETSALELIKQHLLDDFAFMEN-----YCVSDDRLSQ	39
CaERF34	-----MAS---QTQITETSALELIKQHLLDDFAFMEN-----YCASDDRLSQ	39
CaERF33	-----MAS---QTQITETSALELIKQHLLDDFAFMEN-----YCVSDDRLSQ	39
At5g07580	TCRFFLSFVKMAS-----FEESSDLEAIQSHLLEDLLVCDG-----FMG	96
At5g61590	-----MET-----FEESDLDVIQKHLFEDLMIPDG-----FIE	29
At1g80580	-----MENSYTVDGHRLQYSVPLSSMHETSQNSETYG-----	32
CaERF35	-----MTIEVSSVQIFLIKRVLTTLNRKRKPTNG-----	30
At2g20350	-----MENSENVPSYDQINIFT-----	17
At5g67000	-MLTPFCSSHHLQEKMNSCQSNPTKMDNSENVLFNQENENFTL-----	42
At5g67010	-----MDYSENV---QNKNT-----	13
At1g25470	-----MKSFK-----	6
At1g68550	-----MVAIRKEQSLSGVSSEIKKRAKR-----	23
At3g25890	-----MAERKKRSSIQTNKPNKKPMKKP-----	24
At4g11140	-----RILRISVTDPYATDSSSDEEEVDFDALSTKRRRVKKYVKEVVL	54
At4g23750	TNTTTIVSELTNTHQTRILRISVTDPDATDSSSDDEEE-EHQRFVSKRRRVKKFVNEVYL	77
CaERF13	RRSLTPSGQSLTVVRQKVVRIIYTDADATDSSDDEEET-----RNVRVKRHVEEINF	85
At5g53290	-----EKLSGGKSLKRVISICYTDPDATDSSSDEE-----DFLFPRRRVKRFVNEITV	76
At4g27950	-----ERKTSVRDSARLVRSMTDRDATDSSSDEE-----FLFPRRVKRLINEIRV	78
At2g46310	-----TPFDGFPKIVKIIIVTDPDATDSSSDEEND-----NKSVA PRVKRYVDEIRF	64
AT2G20880	-----QQRFL-----HPNDPSGQRQQQMI SFDPPQQVQPYVAQQQQQQHLLQYWRD	149

AT4G28140	-----QQQQLGSYVPVLEQRQDPTMQGQKQMISFSPQQ-----	-----QQQQQQYMAQYWSD	115
AT1g22190	-----APISYGSDLH-----		56
AT1g78080	-----PDSTFLTQPFYSYGSDDLQQTGSLIGLNNLSSSQIHQIQSQIHHP		105
AT1G36060	-----SHSTTSYPDIYSGSMTYPSSFGSDLQQPEN-----	-----YQSQFHYQNT	102
CaERF01	-----DQQGSLGLNQLTPSQIHDIQSRFSLQENR-----	-----HHMISGPFHQQW	58
CaERF02	-----MISGPFHQQW		10
CaERF05	-----DQQGSLGLNQLTPSQIHDIQSRFSLQENR-----	-----HHMISGPFHQQW	127
AT2g22200	-----KSAMPLGLNQLTPYQIHQIQN-----		38
AT4g39780	-----QTTSGLNQLTPYQIHQIQN-----		64
CaERF04	DEMGFDEQSGGSIIGNHLTPYQILQVEAQIKLQNNQQ-----	-----QYLAAQSLQNYNYS	132
CaERF03	-----QTGPIGLNQLTPTQILQIQTELHLR-----		78
AT5g65130			
At4g31060			
At2g23340		-----MET	3
At4g36900		-----MET	3
At5g67190		-----MEG	3
At3g50260		-----MDA	3
CaERF06		-----MEG	3
At4g06746		-----MVI	3
At1g46768		-----MER	3
At1g21910		-----MVKQERKIQT	10
At1g77640		-----MVKQELKIQV	10
At1g44830		-----MVKTLQKTPK	10
At1g22810			
At1g71520			
At1g19210			
At1g74930			
At5g21960			
At2g40220		-----MDPLASQHQHNHLED	15
At1g71450		-----MAGLR	5
At1g33760		-----MENT	4
At1g63040	LSDSATCENPCPLDTITTTTTTVCFAAPSSTASGNDIN-----		72
At1g12630	-----MN-----		2
At5g52020	QTS-----PGSSSSPSRPSG-----		47
CaERF08	TASWSN-----TCECPPPPSQPSPA-----		56
CaERF07	TASGSN-----TCECPPPPSQPSPA-----		56
At2g36450			
At4g25470	-----VSSGGDYSPKLATSC-----		33
At4g25480	-----VSSGGDYIPTLASSC-----		33
At4g25490	-----QGGDYCPTLATSC-----		30

At5g51990	-----VSDSSECSPKLASSC-----	36
At1g63030	-----DITVAEMK-----	12
At1g12610	-----DIILAEMR-----	12
At3g60490	ITATISSSS-----VVTSSSDSWSTSKRS-----LVQ	48
CaERF09	TTTTTSSSSSFSTTSNALNSPSTTSTSSSSSSSNQSTKATTSTRTAKAGKAQESSQELQQ	98
At2g35700	-----RVKY-KPKK--KRA-----	30
At4g16750	-----NLRSPVPEKTGKSS-----	25
At2g44940	SSSSSITAS-----ASLSSSPTTSSSSSSSTNSNFIEEDNSKRKASRRSLSSLVS	74
At5g25810	-----MIASESTKSWEASAVRQE-----	18
At5g11590	-----LLVPNSESDSVSDKSKAE-----	33
At2g25820	-----MVDSHGSDTECSSKKKKE-----	18
At4g32800	-----MADSS-----SDKEKKE-----	12
At1g77200	-----SVKQSSPVPEEEDHHHHQ-----	25
At3g16280	-----TTTTSTHLSEEEAPPRNN-----	38
At3g11020	----MAVYEQTG-----TEQPKKRKSRARAGGLTVADRLKKWKEYNEIVEASAVKEG--	48
At5g05410	----MAVYDQSGDRNRTQIDTSRKRKRSRSGDGTVAERLKRWKEYNETVEEVSTK----	52
CaERF10	----MALLDQVSNVASVPMDYSRKRKSRRRDGTNNVADTLAKWKEYNDKLDLDEEG---	53
At2g40350	----MP-----RKRKSR---GTRDVAEILRWREYNEQTEADSCIDGGG	37
At2g40340	----MPS-----EIVDRKRKSR---GTRDVAEILRWREYNEQIEAESCIDGGG	42
CaERF11	---MSPLKPASDETSLTLPFGRKNSKRRSDASKHVAETLNKWKKEYFDKIESINGK----	53
At2g38340	-----MEKEDNGSKQSSSASVVSSRRRRRVVEPVEATLQRWEEEG	40
At1g75490	-----MVGANKKQ-----	8
At5g18450	-----MEEEQPPAK-----	9
At3g57600	-----MEKSSS-----	6
CaERF12	-----MDADRSNGNGRPG-----	14
At5g11190	-----	
At5g25390	-----	
At1g15360	-----	
At5g25190	-----	
At5g19790	-----	
At3g15210	-----	
At1g53170	-----	
At1g28370	-----	
CaERF22	-----	
CaERF23	-----	
CaERF25	-----	
At1g03800	-----MTT	3
At5g44210	-----MAP	3
At1g50640	-----	
At3g20310	-----	

CaERF24	-----	
CaERF26	-----MDV	3
At1g28360	-----	
At1g12980	-----MEKALRNFTTESTHSPDPNPLTKFFTEP	27
At5g13910	-----	
At5g18560	-----MSTSKTLDHNKPFETSQTQ	19
At1g28160	-----MEF	3
At1g12890	-----	
At1g24590	-----MEEAIMRL	8
At2g47520	-----VSSR	28
At1g72360	-----WRSNGVFDCSIYDFDGNFDELESDE	49
CaERF14	-----EFDTISP--FWGFDPSGQAAD	45
CaERF15	-----EFDTISP--FWGFDPSGQAAD	45
CaERF16	-----EFDTISP--FWGFDPSGQAAD	45
At3g16770	-----ELDASAADDFWGFYSTSKLHP	50
CaERF17	-----SATDFWPNSHLAKPSRSNPTDAPH	54
CaERF18	-----	
At1g53910	KSSKN-----RSNFFDFDAEFEADFQGFKDDSSIDCDDDFDV	73
At3g14230	KRSNK-----RSDFFDLDDDFEADFQGFKDDSAFDCEDD	70
CaERF19	KSTSNFYSKQLRSEIIDLDDEFEADFREF---SDQAVEEL	80
CaERF20	-----MEEL	4
CaERF21	KNTSDLYSKPLRSEVVDLDDEFETDFQEFNDHSDDEAEEL	87
At5g61890	-----DNISSIDDNYPSVYNP	68
At5g07310	-----DN---NQHQPVVYNQ	71
CaERF36	-----NNSVAVVGNYPALADP	33
At5g13330	-----PPVKQE	20
At1g43160	-----SATWTMGHKREREFEFS	33
At5g50080	-----SAPIPSSFARVDSS	54
CaERF37	-----STIATTVQS	24
At5g64750	GGAKAVKAANTLTVDQYFSGGSSTSKVREASSNMSGPGPT	160
At2g33710	-----DVSDFVSELTG	43
At2g31230	-----QESLLWNESCFLDQSSEPQAFFC-PNYDYSD	D 49
At1g06160	-----QESFLWEES-FLHQSFQSFLLSSPTDNYCD	D 55
At3g23240	-----SSSSSNYSLFPFNENDSEEMFLYGLIEQSTQ	Q 58
At5g43410	-----MDQGGR	6
At1g04370	-----MDQGGRS	7
At3g23220	-----MERIES	6
At3g23230	-----MESSNRS	7
At4g18450	-----EVSRTWEALPTLDDIPEGSREMLQSLDMSTE	D 67
At4g17500	PSFSKLYPCFTESWGDLPKENDSEDMLVYGILNDAFHG	G 94

At5g47220 -----CFTESWGGLPLKENDSEDMLVYGLLKDAFH----- 69
 CaERF27 -----LVPCLTETWGDLPKVVDDSEDMVIYWLLRDAVNV-----G 64
 CaERF28 -----LVPCLTETWGDLPKVVDDSEDMVIYWLLRDAVNV-----G 64
 CaERF29 -----LVPCLTETWGDLPKVVDDPEDMVIYWILRDAVNV-----G 64
 At2g44840 -----VILNDNWSDLPLSVDDSQDMAIYNTLRDAVSS-----G 57
 CaERF30 -----MVPYPQAEPLNENDSQEMVLYEVLTEANA----- 30
 CaERF31 -----MVPYPQAEPLNENDSQEMVLYEVLTEANA----- 30
 At5g47230 SSIIFGSSSSSFAPIDFSESVCCKPEIIDLDTPRSMEFLSIPFEFDSEVSVSDFDFKPSNQ 104
 At4g17490 DFTSFDQTG-----FEFSEFETKPEIIDLVTPKPEIFD---FDVKSEIPSESNDSTTFQS 87
 At5g61600 QLHNFHSETG-----PRTITNQSPK-----PNS 57
 At5g51190 TIH--HCTT-----S 43
 CaERF32 FSHLSTSQSNSSSDSDTSSFATSNASFATSSSSSSIN-----YMELETKPNVVALTTP 94
 CaERF34 FSHLSTSQSNSSSDSDTSSFVTSNASFSFATSSSSSSIN-----YMELETKPNVVALTTP 94
 CaERF33 FSHLSTSQSNSSSDSDTSSFATSNASFATSSSSSSIN-----YMELETKPN-VALTTP 93
 At5g07580 DFDFD-ASFVSGLWCIEPH--VPKQEPDSPVLDPDSFVN-----EFLQ 136
 At5g61590 DFVFDDTAFVSGLWSLEPFNVPKLEPSSPVLDPDSYVQ-----EILQ 72
 At1g80580 -----LSKESPLVCMPLFETNTTTSFDISS-LFSFNPKPEPENTHRVMD 74
 CaERF35 -----TFLDPVFFHKTVFSSYKQIIAEGENLDFCIAGDEDETRYLWD 73
 At2g20350 -----PN-----LTRDQEHVIMVSALQQVISNVGGD--TNSNAWEA--DLPPL 56
 At5g67000 -----VAPHPSSSYLTRDQEHVIMVSALRQVISNSGADDASSNLIIT--SVPPP 90
 At5g67010 -----PISQPPNLTRDQEHAIMVSTLRQVISNTGGDT-SSSHYIAASEALPPS 60
 At1g25470 -----PERDSLRLTRIVFTDPDATDDSSSSSDE-WLPKPRKVKRFVHEITFLP-Q 55
 At1g68550 ---NTLSSLPQETQPLRKVRIIVNDPYATDDSSSDEEELKVPKPRKMKRIVREINFPSME 80
 At3g25890 FQLNHLPLGLEDLKTMRKLRFFVNDPYATDYSSEEEEE---RSQRRKRYVCEIDL PFAQ 80

At4g11140 DS--VVSDKEKPMK-KRRKKRVVT-----VPVVVTTAT 84
 At4g23750 DSGAVVTGSCGQMESKKRQKRAVKSE-----STVSPVVSATTT 115
 CaERF13 GPPPKTTKIEQPRNGKRSHQSLTP-----PAEPDVSS 117
 At5g53290 EPSCNNV--VTGVSMKDRKRLSSSSDETQSPAS-----SRQRPNNKVSVS 119
 At4g27950 EPSSSSTGDVSASPTKDRKRINVDS-----TVQKP---SVS 111
 At2g46310 CDEDDEPKPARKAKKSPAAAAENG-----GDL 92
 AT2G20880 ILKLSPSGRMMMMNMLRQESDLPLTRPPVQP----- 180
 AT4G28140 TLNLSPRGRMMMM-----MSQEAVQP----- 136
 AT1g22190 -----SFSHHLSP-KPVSMKQTGTS-----A 76
 AT1g78080 LPPTHNNNNNSFSNLLSP-KPLLMKQSGVAGSCFAYG-----SGVP 145
 AT1G36060 IITYHQDNNTCMLNFIEPSQPGFMTQPGPS-----SGSV 136
 CaERF01 PATRQSQSQQQALPFLAP-KPVPMKQSASP----- 87
 CaERF02 PATRQSQSQQQALPFLAP-KPVPMKQSASP----- 39
 CaERF05 PAIRQSQSQQQALPFLAP-KPVPMKQSASP----- 156

AT2g22200	---QLNHRRSTIS-NLSP-NRIRMKNLTPST-----	64
AT4g39780	---QLNQRRNIISP NLAP-KPVPMKNMT-----	88
CaERF04	STRQLQHQNHSPLSFLGP-KPVQMKQVGT-----	161
CaERF03	-----QVATP-----	5
AT5g65130	---QNQSRRRAGSHLLTA-KPTSMKKIDVA-----	104
At4g31060	-----MPPSPPKSPFISSSLKGAHED-----	21
At2g23340	EAAVTATVT--AATMGIGTRKRD-----	25
At4g36900	ADEVATVVSTPAVTVAAVATRKR-----	27
At5g67190	GGVADVAVP-----GTRKR-----	18
At3g50260	GVAVKADVA-----VKMKRE-----	18
CaERF06	GGGIGAGAG-----KKKSE-----	17
At4g06746	QYKRKQEFPMVKEGMVMTEKPKRNLISSN-----	32
At1g46768	EQEESTMRKRRQPPQEEVPNHVATR-----	28
At1g21910	SSTKKEMPLSSSPSSSSSSSSSSSSSSCKNKKNK-----	44
At1g77640	TT-----SSSSLSHSSSSSSSSSTSALRHQSCK-----	37
At1g44830	RM-----SSPSSSSSSSSSTSSSSIR-----	31
At1g22810	-----MDYRESTGESQS-----	12
At1g71520	-----MDSRDTGETDQS-----	12
At1g19210	-----MEGS----SSSMQ-----	9
At1g74930	-----MVKQAMKEEEKRNTAMQ-----	18
At5g21960	-----MDAS-----	4
At2g40220	NNQTLTHNNPQSDSTTDSSTSSAQKRKKGKGGPD-----	49
At1g71450	-----NSGNSDKAQNDGKGV-----	20
At1g33760	-----YVQQRDYRFNVN-----	16
At1g63040	-----TLMATDTDISRRKK-----	86
At1g12630	-----SSMAS-AGLGSRRK-----	15
At5g52020	-----GSPARRTATGLSGK-----	61
CaERF08	-----TIPSPTGSSGPRS-----	70
CaERF07	-----TIPSPTGSSGPRS-----	70
At2g36450	-----MQGTSKDNNGR-----	11
At4g25470	-----PKKPAGRKKFRETR-----	47
At4g25480	-----PKKPAGRKKFRETR-----	47
At4g25490	-----PKKPAGRKKFRETR-----	44
At5g51990	-----PKKRAGRKKFRETR-----	50
At1g63030	-----PKKRAGRRIFKETR-----	26
At1g12610	-----PKKRAGR RVFKETR-----	26
At3g60490	DNDSSGGK---RRKSNVSDDNKNPTS-----	70
CaERF09	VAKINGNENQPRKKHKNGDDEKHPT-----	123
At2g35700	-----KDDDEKVVSKHPN-----	44
At4g16750	-----KTKNEQKGVSKQPN-----	39

At2g44940	VEDDDDQNGGGGKRRKTNGGDKHPT-----	99
At5g25810	-----NEEEKKKPKVSDSGK-----	32
At5g11590	-----QSEKTKRGRDSGK-----	47
At2g25820	-----KTKEKG-----	24
At4g32800	-----NNKQP-----	17
At1g77200	-----QDSHRTNTKKRVRS-----	39
At3g16280	-----NTRKRRRDSSSSASSSSSMQ-----	57
At3g11020	EKPKRKVPAKGSKKGCMKGKGGPDNS-----	74
At5g05410	---KRKVPKGSKKGCMKGKGGPENS-----	75
CaERF10	-KVIRKVPKGSKKGCMKGKGGPENA-----	78
At2g40350	SKPIRKAPPKRSRKGCMKGKGGPENG-----	63
At2g40340	PKSIRKPPPKGSRKGCMKGKGGPENG-----	68
CaERF11	--LVRKAPGRGSKKGCMKGKGGPENF-----	77
At2g38340	LARARRVQAKGSKKGCMRGKGGPENP-----	66
At1g75490	---RTVQAS-SRKGCMRGKGGPDNA-----	29
At5g18450	----KRNMGRSRKGCMKGKGGPENA-----	30
At3g57600	-----MKQWKKGPARGKGGPQNA-----	24
CaERF12	----KKPAQASSRKGCMRGKGGPENA-----	36
At5g11190	-----MVHS-----	4
At5g25390	-----MVHS-----	4
At1g15360	-----MVQT-----	4
At5g25190	-----MARPQ-----	5
At5g19790	-----MEHQTTPKQKTKEKSKGNK-----	19
At3g15210	--MAKMGLKPDP-ATTNQTHNN-----	19
At1g53170	-MPNITMGLKPDPVAPTNPTHHESN-----	24
At1g28370	-----MAPTVKTA AVKTNEG-----	15
CaERF22	-----MAMKEKSKDVKPSTGKAN-----	18
CaERF23	-----MAMEEKSKDVKPSTGKAN-----	18
CaERF25	-----MAMKEKSKDVKPSTGKAN-----	18
At1g03800	EKENVTTAVAVKDGGEKSKEVSDKGVKKRKNVTKALAVN-----DGGE	46
At5g44210	RQANGRSIAVSEGGGGKTMTMTT-----	26
At1g50640	----MRRGRGSSAVAGPTVVAAIN-----G	21
At3g20310	----MRKGRGS-SVVG PALPVTAG-----G	20
CaERF24	----MRRGRAAAAQAAAAAAGET-----NGSG	26
CaERF26	DHLHQEHNNANA VRRGVQVTPSS-----LR	29
At1g28360	-----MASTTC-----	6
At1g12980	TASPVSRNRKLSKDTT VTIAGAG-----	51
At5g13910	-----MNTTSSKSKKKQ-----DD	14
At5g18560	MGFALIHQNTSANTTTTTTTGERRRR-----SK	48
At1g28160	NGNLNAGSCSRSKKSHRQKQQQPQP-----QQ	32

At1g12890	-----MLKSSNKRKSKEEKKLQEG-----	19
At1g24590	EGAEHRETNIHSLKRKPSRTSSTAPGSPGGVTTAKAASG-----AGAS	51
At2g47520	-----KRRK-----PVSVSEERDGK	43
At1g72360	-----PFVFSSTHK-----HHASGSASDGK	69
CaERF14	-----TKATDKNKH-----PKKGAGNQKSEKISQ	69
CaERF15	-----TKATDKNKH-----PKKGAGNQKAEKISQ	69
CaERF16	-----TKATDKNKH-----PKKGAGNQKSEKISQ	69
At3g16770	-----TNQVNVKEE-----AVK--KEQATEPGKR	72
CaERF17	-----KATHETCHDPH-----DEAGGKQQVLPKGAK	80
CaERF18	-----KQTGGKQQVLPKGAK	17
At1g53910	-----GDVFADV KPFVFTSTPKPAVSAAA-----EGSVFGKKVTGLDGD AEKSAN	118
At3g14230	-----DDV FVNVPFVFTATT KPVASAFVSTGIYLVGSAYAKKTVESAEQA EKSSK	121
CaERF19	-----EDVKPF GFSASKH---SAILS----RGT KCGN-SSDSSDES GKYLK	118
CaERF20	-----EDVKPF GFSASKH---SAILS----RGT KCGN-SSDSSDES GKYLK	42
CaERF21	-----EDVKPF GFSASKH---SAIPP----RGT KYGK-SSDSSDESEKYSK	125
At5g61890	-----QDPNQOVA-----PTHQDQG	83
At5g07310	-----QDPNP-PA-----PPTQDQG	85
CaERF36	-----QPAPAIEN-----RAQQEQG	48
At5g13330	-----LDKSD-----QHQPQDQD	32
At1g43160	-----LPPQPLITGSA-----VTKECESSMS	54
At5g50080	-----HNPIEESMS-----KAFPEEA	70
CaERF37	-----QNPPAEAA-----EET	35
At5g64750	-----YEYTTTATASS-----ETSSFSG	178
At2g33710	-----QPIPSSIDDQ-----SSSLTLQEK S	63
At2g31230	FFSFESPEMMIKEEIQNGDVSNS EEEEEKVGIDEE-----	83
At1g06160	FFAFESS--IIKEEGKEATVA AEEEE-----	79
At3g23240	TYIDS DS-----QDLPIKSVSSRKSE-----	79
At5g43410	-----GVGAEHGK-----	14
At1g04370	-----SGSGGGGAEQ GK-----	19
At3g23220	-----YNTNEMK-----	13
At3g23230	-----SNNQSQDDKQAR-----	19
At4g18450	QEWTEILD AIASFPNKTNHDPLTNPTIDSCSLSSRVSC-----	105
At4g17500	WEPSSSSSDEDRSSFPSVKIETPE SFAAVDSVPVKKEKT-----SPVSAAVTAA	143
At5g47220	FDTSSSDL S-CLFDFPAVKVEPTENFTAMEEKPKK-----AIPVTETAV	112
CaERF27	WSPFNLTASSDVSVDVRVKAEP SDAPEPATTSPAADHCS-----APPPATAGAA	113
CaERF28	WSPFNLTASSDVSVDVRVKAEP SDAPEPATTSPAADHCS-----APPPATAGAA	113
CaERF29	WSPFNLTASSDVSVDVRVKAEP SDAPEPATTSPAADHCS-----APPPAAAGAA	113
At2g44840	WTPSVPPVT-----SPA EENKPPATKASGSHAPRQ-----	87
CaERF30	I SP SNFPSSRNQASV NKMLHRP SHATG-----	57
CaERF31	I SP SNFPSSRNQASV NKMLHRP SHATG-----	57

At5g47230 NQNQFEPELKSQIRKPPLKISLPAKTEWIQFA-----AENTKPEVTKP-VS 149
 At4g17490 NPPRVTVQSN---RKPPLKIAPPNRTKWIQFA-----TGPNKPELPVPPVVA 130
 At5g61600 TLNQ-----RKPPLPNLSVSRT-----VSTKTEK 81
 At5g51190 TLSQ-----RKPPLATIAVPTT-----APVVQEN 67
 CaERF32 NHPRQTATSFSE-RKPVLNIAIPPPVKLLDFNKMNTSNTTTTTYEQPKAKAVEQKKVQEG 153
 CaERF34 NHPRQTATSFSE-RKPVLNIAIPPPVKLLDFNKMNTSN-TTTYEEPKAKAVEQKKVQEG 152
 CaERF33 NHPRQTATSFSE-RKPVLNIAIPPPVKLLDFNKMNTSNTTTTTYEQPKAKAVEQKKVQEG 152
 At5g07580 VEGESSSSSPELNSSSSTYETDQSVK-----KAERFEE 170
 At5g61590 MEAESSSSS---TTTSPEVETVSNRK-----KTKRFEE 103
 At1g80580 DSIAAVVGENVLFQDKNKVSDDLTKEGGVKRGK-----M 109
 CaERF35 TGPSSSTST-----DENEIIFSLQKE----- 94
 At2g20350 NAGPCPLCSVTGCGYCVFPRHEAIIK----- 82
 At5g67000 DAGPCPLCGVAGCYGCTLQRPHEVVK----- 116
 At5g67010 DAGPCPLCGVTSCYGCAFP-QHEEIK----- 85
 At1g25470 VSE-----SSQD--RSNAVKTPRRK-----S 74
 At1g68550 VSEQPSESSSQDSTKTDGKIAVSAS-----P 106
 At3g25890 AATQAESESSYCQESNNGVSKTKISACSKKVLRSKASP-----VVGRS 124

At4g11140 ----RKFGRVQR-PWGKWAAEIRDPS-----RRVRVWLGTTFDTAEEAAIVYDAAI 131
 At4g23750 TTGEKKFRGVRQR-PWGKWAAEIRDPL-----KRVRLWLGTYNATAEEAAMVYDAAI 166
 CaERF13 ---RKKFRGVRQR-PWGRWAAEIRDPT-----KRKRVLWLTGTYDTPPEEAASVYDRAAV 165
 At5g53290 GQ-IKKFRGVRQR-PWGKWAAEIRDPE-----QRRRIWLGTFFETAEEAAVVYDAAI 169
 At4g27950 GQNQKKYRGVRQR-PWGKWAAEIRDPE-----QRRRIWLGTFFATAEEAAIVYDAAI 162
 At2g46310 VKSVVKYRGVRQR-PWGKFAAEIRDPS-----SRTRLWLGTFFATAEEAAIGYDRAAI 143
 AT2G20880 FSATKLYRGVRQR-HWGKWVAEIRKPR-----NRTRLWLGTTFDTAEEAAMAYDREAF 231
 AT4G28140 YIATKLYRGVRQR-QWGKWVAEIRKPR-----SRARLWLGTTFDTAEEAAMAYDRQAF 187
 AT1g22190 AKPTKLYRGVRQR-HWGKWVAEIRLPR-----NRTRLWLGTTFDTAEEAALAYDKAAY 127
 AT1g78080 SKPTKLYRGVRQR-HWGKWVAEIRLPR-----NRTRLWLGTTFDTAEEAALAYDKAAY 196
 AT1G36060 SKPAKLYRGVRQR-HWGKWVAEIRLPR-----NRTRLWLGTTFDTAEEAALAYDRAAF 187
 CaERF01 PKPTKLYRGVRQR-HWGKWVAEIRLPK-----NRTRLWLGTTFDTAEEAALAYDKAAF 138
 CaERF02 PKPTKLYRGVRQR-HWGKWVAEIRLPK-----NRTRLWLGTTFDTAEEAALAYDKAAF 90
 CaERF05 PKPTKLYRGVRQR-HWGKWVAEIRLPK-----NRTRLWLGTTFDTAEEAALAYDKAAF 207
 AT2g22200 SKTKNLYRGVRQR-HWGKWVAEIRLPK-----NRTRLWLGTFFETAEKAALAYDQAAF 115
 AT4g39780 --AQKLYRGVRQR-HWGKWVAEIRLPK-----NRTRLWLGTTFDTAEEAAMAYDLAAY 137
 CaERF04 PKPAKLYRGVRQR-HWGKWVAEIRLPK-----NRTRLWLGTTFDTAEEAALAYDKAAY 212
 CaERF03 PKPAKLYRGVRQR-HWGKWVAEIRLPK-----NRTRLWLGTTFDTAEEAALAYDKAAY 56
 AT5g65130 TKPVKLYRGVRQR-QWGKWVAEIRLPK-----NRTRLWLGTFFETAQEAALAYDQAAH 155
 At4g31060 -RKFKCYRGVRKR-SWGKWVSEIRVPK-----TGRIIWLGSYDAPEKAARAYDAALF 71
 At2g23340 ----KPYKGIRMR-KWGKWVAEIREPN-----KRSRIWLGSYATPEAAARAYDTAVF 72

At4g36900 ----KPYKGIRMR-KWGKWVAEIREPN-----KRSRIWLGSYSTPEAAARAYDTAVF 74
 At5g67190 ----RPYKGIRMR-KWGKWVAEIREPN-----KRSRLWLGSYSTPEAAARAYDTAVF 65
 At3g50260 ----RPFKGIRMR-KWGKWVAEIREPN-----KRSRLWLGSYSTPEAAARAYDTAVF 65
 CaERF06 ----RLYKGIRMR-KWGKWVAEIREPN-----KRSRIWLGSYSTPVAAARAYDTAVY 64
 At4g06746 ---EKRYKGIRMR-KWGKWVAEIREPN-----KRSRIWLGSYKTAVAAARAYDTAVF 80
 At1g46768 ----KPYRGIRRR-KWGKWVAEIREPN-----KRSRLWLGSYTTDIAAARAYDVAVF 75
 At1g21910 -SKIKKYKGVRRM-SWGSWVSEIRAPN-----QKTRIWLGSYSTAEEAARAYDVALL 94
 At1g77640 -NKIKKYKGVRRM-SWGSWVTEIRAPN-----QKTRIWLGSYSTAEEAARAYDAALL 87
 At1g44830 ---MKKYKGVRRM-SWGSWVSEIRAPN-----QKTRIWLGSYSTAEEAARAYDAALL 79
 At1g22810 -----KYKGIRRR-KWGKWVSEIRVPG-----TRDRLWLGSFSTAEGA AVAHDVAFF 58
 At1g71520 -----KYKGIRRR-KWGKWVSEIRVPG-----TRQRLWLGSFSTAEGA AVAHDVAFY 58
 At1g19210 ----SKYKGVRRK-KWGKWVSEIRLPN-----SRERIWLGSYDTPEKAARAFDAALY 56
 At1g74930 ----SKYKGVRRK-KWGKWVSEIRLPH-----SRERIWLGSYDTPEKAARAFDAAQF 65
 At5g21960 ----PKYTGVRKR-KWGKWVAEIRLPN-----SRDRIWLGSFSDSAEKAARAFDAALY 51
 At2g40220 -NSKFRYRGVRRQ-SWGWVVAEIREPR-----KRTRKWLGTFFATAEDAARAYDRAAV 99
 At1g71450 ---PSAYRGVRRK-KWGKWVSEIREPG-----TKNRIWLGSFETPEMAAATAYDVAAF 68
 At1g33760 ---QLSYRGIRRR-KWGKWVSEIREPG-----KKTRIWLGSYETAEMAAAAYDAAAL 64
 At1g63040 ---NPVYRGIRCR--SGKWVSEIREPK-----KTTRVWLGTYPTEMAAAAAYDVAAL 133
 At1g12630 ---DPVYRGIRCR--SGKWVSEIREPR-----KTTRIWLGTYPMAEMAAAAYDVAAM 62
 At5g52020 ---HSIFRGIRLR--NGKWVSEIREPR-----KTTRIWLGTYPPPEMAAAAAYDVAAL 108
 CaERF08 ---HPHFRGVRSR--SGKWVSEIREPR-----KTTRIWLGTYPTPEMAAAAAYDAATL 117
 CaERF07 ---HPHFRGVRSR--SGKWVSEIREPR-----KTTRIWLGTYPTPEMAAAAAYDAATL 117
 At2g36450 ---HPLYRGVRRQKNSNKWVSEIREPR-----KPNRIWLGTFSTPEMAAIAYDVAAL 60
 At4g25470 ---HPIYRGVRRQ-NSGKWVCEVREPN-----KKTRIWLGTFQTAEMAARAHDAVAI 95
 At4g25480 ---HPIYRGVRRR-NSGKWVCEVREPN-----KKTRIWLGTFQTAEMAARAHDAVAAL 95
 At4g25490 ---HPIYRGVRRQ-NSGKWVCEVREPN-----KKTRIWLGTFQTAEMAARAHDAVAAL 92
 At5g51990 ---HPIYRGVRRQ-NSGKWVCEVREPN-----KKSRIWLGTFPTVEMAARAHDAVAAL 98
 At1g63030 ---HPIYRGVRRR-DGDKWVCEVREPI-----HQRRVWLGTYPPTADMAARAHDAVAVL 74
 At1g12610 ---HPVYRGIRRR-NGDKWVCEVREPT-----HQRRIWLGTYPTADMAARAHDAVAVL 74
 At3g60490 -----YRGVRRM-SWGWVSEIREPR-----KKSRIWLGTYPTAEMAARAHDAVAAL 115
 CaERF09 -----YRGVRRK-NWGKWVSEIREPR-----KKSRIWLGTYPTAEMAARAHDAVAAL 168
 At2g35700 -----FRGVRRM-QWGWVSEIREPK-----KKSRIWLGTFSTAEMAARAHDAVAAL 89
 At4g16750 -----FRGVRRM-QWGWVSEIREPR-----KKSRIWLGTFSTPEMAARAHDAVAAL 84
 At2g44940 -----YRGVRRM-SWGWVSEIREPR-----KKSRIWLGTYPTAEMAARAHDAVAAL 144
 At5g25810 ---HPVYRGVRRK-NWGKWVSEIREPR-----KKSRIWLGTFPSPPEMAARAHDAVAAL 80
 At5g11590 ---HPVYRGVRRM-NWGKWVSEIREPR-----KKSRIWLGTFPTPEMAARAHDAVAAL 95
 At2g25820 -----VYRGARRM-SWGWVSEIREPR-----KKSRIWLGTFPTAEMAARAHDAVAAL 70
 At4g32800 -----VYRGVRRM-SWGWVSEIREPR-----KKSRIWLGTFPTAEMAMRAHDVAAM 63
 At1g77200 ---DPGYRGVRRM-TWGWVSEIREPR-----KKSRIWLGTFSTPEMAARAHDAAL 87
 At3g16280 ---HPVYRGVRRM-SWGWVSEIRQPR-----KKTRIWLGTFVTADMAARAHDAVAAL 105

At3g11020 ---HCSFRGVRQR-IWGKWVAEIREPK-----IGTRLWLGTFFPTAEKAASAYDEAAT 122
 At5g05410 ---RCSFRGVRQR-IWGKWVAEIREPN-----RGSRLWLGTFFPTAQEAASAYDEAAK 123
 CaERF10 ---RCNYRGVRQR-TWGWKVAEIREPN-----RGSRLWLGTFFGTAVEAALAYDEAAR 126
 At2g40350 ---ICDYTGVRQR-TWGWKVAEIREPG-----RGAKLWLGTFFSSSYEAALAYDEASK 111
 At2g40340 ---ICDYRGVRQR-RWGWKVAEIREPD-----GGARLWLGTFFSSSYEAALAYDEAAK 116
 CaERF11 ---RHNYRGVRQR-TWGWKVAEIREPK-----KGSRLWLGTFFGTAIEAALAYDKAAR 125
 At2g38340 ---VCRFRGVRQR-VWGWKVAEIREPVSHRGANSRKRSLWLGTFFATAAEEAALAYDRAAS 122
 At1g75490 ---SCTYKGVQR-TWGWKVAEIREPN-----RGARLWLGTFFDTSREAALAYDSAAR 77
 At5g18450 ---TCTFRGVRQR-TWGWKVAEIREPN-----RGTRLWLGTFFNTSVEAAMAYDEAAK 78
 At3g57600 ---LCQYRGVRQR-TWGWKVAEIREPK-----KRARLWLGSFATAEEAAMAYDEAAL 72
 CaERF12 ---SCAYKGVQR-TWGWKVAEIREPN-----RGSRLWLGTFFDTSHEAAIAYDAAAR 84
 At5g11190 ----RKFRGVRQR-QWGSWVSEIRHPL-----LKRRVWLGTFFETAEEAARAYDQAAL 51
 At5g25390 ----KKFRGVRQR-QWGSWVSEIRHPL-----LKRRVWLGTFFDFTAETAARAYDQAAV 51
 At1g15360 ----KKFRGVRQR-HWGSWVAEIRHPL-----LKRRVWLGTFFETAEEAARAYDEAAV 51
 At5g25190 ----QRFRGVRQR-HWGSWVSEIRHPL-----LKTRVWLGTFFETAEDAARAYDEAAR 52
 At5g19790 ---TKFVGVQR-PSGKWAAEIKDPT-----QKIRMWLGTFFETAEEAARAYDEAAC 66
 At3g15210 -AKEIRYRGVQR-PWGRYAAEIRDPP-----KKTRVWLGTFFDFTAEEAARAYDTAAR 69
 At1g53170 AAKEIRYRGVQR-PWGRYAAEIRDPV-----KKTRVWLGTFFDFTAQQAARAYDAAAR 75
 At1g28370 --NGVRYRGVQR-PWGRYAAEIRDPF-----KKSrvWLGTFFDTPEEAARAYDKRAI 64
 CaERF22 AVKEVHFRGVRQR-PWGRYAAEIRDPP-----KKSrvWLGTFFDFTAEEAARAYDAAAR 69
 CaERF23 AVKEVHFRGVRQR-PWGRYAAEIRDPP-----KKSrvWLGTFFDFTAEEAARAYDAAAR 69
 CaERF25 AVKEVHFRGVRQR-PWGRYAAEIRDP-----KKSrvWLGTFFDFTAEEAARAYDAAAR 69
 At1g03800 KSKEVRYRGVRRR-PWGRYAAEIRDPV-----KKKRVWLGSFNTGEEAARAYDSAAI 97
 At5g44210 MRKEVHFRGVRQR-PWGRYAAEIRDPP-----KKTRVWLGTFFDFTAEEAARAYDTAAR 77
 At1g50640 SVKEIRFRGVRQR-PWGRFAAEIRDPW-----KKARVWLGTFFDSAEAAARAYDSAAR 72
 At3g20310 SVKEPRYRGVRRR-PWGRFAAEIRDPL-----KKSrvWLGTFFDSAVDAARAYDTAAR 71
 CaERF24 GSKEIRFRGVRQR-PWGRFAAEIRDPW-----KKTRVWLGTFFDSAEAAARAYDAAAR 77
 CaERF26 VVKESHFRGVRQR-PWGRYAAEIRDPW-----KKTRVWLGTFFDFTAEEAALAYDEAAR 80
 At1g28360 -AREVHYRGVRRR-PWGRYAAEIRDPW-----KKTRVWLGTFFDTPEEAALAYDGAAR 56
 At1g12980 -SSTTRYRGVRRR-PWGRYAAEIRDPM-----SKERRWLGTFFDFTAEEAALAYDGAAR 101
 At5g13910 QVG-TRFLGVRRRR-PWGRYAAEIRDPT-----TKERHWLGTFFDFTAEEAALAYDRAAR 64
 At5g18560 QAEPRFLGVRRRR-PWGRYAAEIRDPT-----TKERHWLGTFFDFTAEEAALAYDRAAL 99
 At1g28160 HIEEIKYVGVRRR-PWGRYAAEIRNPT-----TKERYWLGTFFDFTAEEAALAYDRAAR 83
 At1g12890 ----KYLGVRRR-PWGRYAAEIRNPF-----TKERHWLGTFFDFTAEEAAFAYDVAAR 65
 At1g24590 GVSTIRYRGVRRR-PWGRYAAEIRDPL-----SKERRWLGTFFDFTAEEAALAYDCAAR 102
 At2g47520 RERKNLYRGIRQR-PWGWAAAEIRDPP-----SKGVRVWLGTFFKTADEAARAYDVAAI 94
 At1g72360 KKQSSRYKGIIRR-PWGWAAAEIRDPP-----IKGVRVWLGTFFNTAEEAARAYDLEAK 120
 CaERF14 KPRKNKYRGIRQR-PWGWAAAEIRDPP-----QKGVVWLGTFFNTAEEAARAYDKAAK 120
 CaERF15 KPRKNKYRGIRQR-PWGWAAAEIRDPP-----QKGVVWLGTFFNTAEEAARAYDKAAK 120
 CaERF16 KPRKNKYRGIRQR-PWGWAAAEIRDPP-----QKGVVWLGTFFNTAEEAARAYDKAAK 120

At3g16770	RKRKNVYRGIRKR-PWGKWA AEIRD P-----RKGVRVWLGT FNTAEEAAMAYDVA AK	123
CaERF17	RQRKNLYRGIRQR-PWGKWA AEIRD P-----XKGVRVWLGT FNTAEEAARAYDKEAR	131
CaERF18	RQRKNLYRGIRQR-PRGKWA AEIRD P-----RKGVRVWLGT FNTAEEAARAYDKEAR	68
At1g53910	RKRKNQYRGIRQR-PWGKWA AEIRD P-----REGARIWLGT FKTAEAAARAYDAAAR	169
At3g14230	RKRKNQYRGIRQR-PWGKWA AEIRD P-----RKGSRVWLGT FDTAEEAARAYDAAAR	172
CaERF19	RKRKNHYRGIRQR-PWGKWA AEIRD P-----RKGVRVWLGT FNTAEEAARAYDTEAR	169
CaERF20	RKRKNHYRGIRQR-PWGKWA AEIRD P-----RKGVRVWLGT FNTAEEGAARAYDTEAR	93
CaERF21	RKRKNQYRGIRQR-PWGKWA AEIRD P-----RKGVRVWLGT FNTAEEAARAYDTEAR	176
At5g61890	DLRRRH YRGVRQR-PWGKWA AEIRD P-----KKAARVWLGT FETAESAALAYDEAAL	134
At5g07310	LLRKRHYRGVRQR-PWGKWA AEIRD P-----QKAARVWLGT FETAEEAALAYDNAAL	136
CaERF36	NQRRRH YRGVRQR-PWGKWA AEIRD P-----KKAARVWLGT FETAEEAALAYDEAAL	99
At5g13330	QPRRRH YRGVRQR-PWGKWA AEIRD P-----KKAARVWLGT FETAEEAALAYDRAAL	83
At1g43160	LERPKKYRGVRQR-PWGKWA AEIRD P-----HKATRVWLGT FETAEEAARAYDAAAL	105
At5g50080	REKKRRYRGVRQR-PWGKWA AEIRD P-----HRAARVWLGT FDTAEEAARAYDEAAL	121
CaERF37	GERRRRYRGVRQR-PWGKWA AEIRD P-----HKAARVWLGT FDTAEEAARAYDEAAL	86
At5g64750	DQPRRRYRGVRQR-PWGKWA AEIRD P-----FKAARVWLGT FDNAESAARAYDEAAL	229
At2g33710	NSRQRNYRGVRQR-PWGKWA AEIRD P-----NKAARVWLGT FDTAEEAALAYDKAAF	114
At2g31230	----RSYRGVRKR-PWGKF AA EIRDST-----RNGIRVWLGT FDKAEEAALAYDQAAF	131
At1g06160	----KSYRGVRKR-PWGKF AA EIRDST-----RKGIRVWLGT FDTAEEAALAYDQAAF	127
At3g23240	----KSYGVRRR-PWGKF AA EIRDST-----RNGIRVWLGT FESAEEAALAYDQAAF	127
At5g43410	-----YRGVRRR-PWGKY AA EIRDSR-----KHGERVWLGT FDTAEEAARAYDQAAY	60
At1g04370	-----YRGVRRR-PWGKY AA EIRDSR-----KHGERVWLGT FDTAEDAARAYDRAAY	65
At3g23220	-----YRGVRKR-PWGKY AA EIRDSA-----RHGARVWLGT FNTAEDAARAYDRAAF	59
At3g23230	-----FRGVRRR-PWGKF AA EIRDPS-----RNGARLWLGT FETAEEAARAYDRAAF	65
At4g18450	--KTRKYRGVRKR-PWGKF AA EIRDST-----RNGVRVWLGT FQTAEAAAMAYDKAAV	155
At4g17500	--KGKHYRGVRQR-PWGKF AA EIRDPA-----KNGARVWLGT FETAEDAALAYDRAAF	193
At5g47220	--KAKHYRGVRQR-PWGKF AA EIRDPA-----KNGARVWLGT FETAEDAALAYDIAAF	162
CaERF27	QPKGRHFRGVRQR-PWGKF AA EIRDPA-----KNGARVWLGT YETAEEAALAYDRAAY	165
CaERF28	QPKGRHFRGVRQR-PWGKF AA EIRDPA-----KNGARVWLGT YETAXXAALAYDRAAY	165
CaERF29	QPKGRHFRGVRQR-PWGKF AA EIRDPS-----KNGARVWLGT YETAEEAALAYDRAAY	165
At2g44840	--KGMQYRGVRRR-PWGKF AA EIRDPK-----KNGARVWLGT YETPEDAAVAYDRAAF	137
CaERF30	---KKHYRGVRRR-PWGKY AA EIRDST-----RNGTRVWLGT FQTAEAAALAYDRAAF	106
CaERF31	---KKHYRGVRRR-PWGKY AA EIRDST-----RNGTRVWLGT FQTAEAAALAYDRAAF	106
At5g47230	EEEKHHYRGVRQR-PWGKF AA EIRDPN-----KRGSRVWLGT FDTAIEAARAYDEAAF	201
At4g17490	AEEKRH YRGVRMR-PWGKF AA EIRDPT-----RRGTRVWLGT FETAIEAARAYDKEAF	182
At5g61600	EEEERH YRGVRRR-PWGKY AA EIRDPN-----KKGCR IWLGT YDTAVEAGRAYDQAAF	133
At5g51190	D--QRHYRGVRRR-PWGKY AA EIRDPN-----KKGVRVWLGT FDTAMEAARGYDKAAF	117
CaERF32	SGEKHHYRGVRQR-PWGKF AA EIRDPN-----KKGSRVWLGT FDTAVEAAKAYDKAAF	205
CaERF34	SGEKHHYRGVRQR-PWGKF AA EIRDPN-----KKGSRVWLGT FDTAVEAAKAYDKAAF	204
CaERF33	SGEKHHYRGVRQR-PWGKF AA EIRDPN-----KKGSRVWLGT FDTAVEAAKAYDKAAF	204

At5g07580	EVDARHYRGVRRR-PWGKFAAEIRDPA-----KKSRIWLGTFESDVDAARAYDCAAF	222
At5g61590	T---RHYRGVRRR-PWGKFAAEIRDPA-----KKSRIWLGTFESDIDAARAYDYAAF	152
At1g80580	PQKTGGFMGVRKR-PWGRWSAEIRDRIGR-----CRHWLGTFTDAEEAARAYDAAAR	160
CaERF35	PEKQ--YIGVRKR-PWGNNAEIRDSTRNG-----IRVVRGTFTDAEEGAVAYDHCRK	144
At2g20350	--KEKKHKGVRKK-PSGKWAAEIWDPS-----LKVRRWLGTFPTAEMAAYNDAAA	131
At5g67000	--KEKKYKGVRKK-PSGKWAAEIWDPR-----SKSRRWLGTFLLTAEMAAQSYNDAAA	165
At5g67010	--QEKKHKGVRKK-PSGKWSAEIWDPS-----TRTRRWLGTFTPTAEMAADAYDEAAA	134
At1g25470	TRQFKYPVGVRPR-PSGKFAAEILNPF-----TKTKKWLGTYETPAEAEKAYVDKKV	125
At1g68550	AVPRKKPVGVRQR-KWGKWAAEIRDPI-----KKTRTWLGTFTDLEEAAYDAKKL	157
At3g25890	STTVSKPVGVRQR-KWGKWAAEIRHPI-----TKVRTWLGTYETLEQAADAYATKKL	175
	* * : . * : : * *:: . .	
At4g11140	QLRGPNAELNF-----PPPPVTENVEEASTEVEKGVSD-----	163
At4g23750	QLRGPDALTNFSVTPPTTATEKKAPPPSPVKKKKKKNNKSKKSVTASSSI-----	215
CaERF13	KLKGPDAVTNF-----PVESIKERSVGRVRATKVVKAP-----	198
At5g53290	RLRGPDALTNFSIPPQEEEEEE--EPEPVIEEKPVIMTTPTPTTSSSE-----	215
At4g27950	KLKRPDALTNFTVQP-----EPEPVQEQE---QEPESNMSVSIS-----	198
At2g46310	RIKGHNAQTNFLT-----PPSPPTTEVLPEPVIDLETVSGCD-----	181
AT2G20880	KLRG-ETAR-----LNFPELFLNKQE-----	251
AT4G28140	KLRG-HSAT-----LNFPEHFVNKE-----	206
AT1g22190	KLRG-DFAR-----LNFPLDRHN-----DE-----	146
AT1g78080	KLRG-DFAR-----LNFPNLRHNGSHIGGDFGE-----	223
AT1G36060	KLRG-DSAR-----LNFPALRYQTGSSPSDTGE-----	214
CaERF01	KLRG-DFTR-----LNFPHLRHNGSLIGGDFGE-----	165
CaERF02	KLRG-DFTR-----LNFPHLRHNGSLIGGDFGE-----	117
CaERF05	KLRS-DFTR-----AELP-----	219
AT2g22200	QLRG-DIAK-----LNFPNLIHED-----	133
AT4g39780	KLRG-EFAR-----LNFPPQFRHEDGYGGG-SC-----	163
CaERF04	KLRG-EFAR-----LNFPNLRHQLSQDFG---D-----	236
CaERF03	KLRG-EFAR-----LNFPNLRHQLSQDFG---D-----	80
AT5g65130	KIRG-DNAR-----LNFDPDIVRQGHY-----	175
At4g31060	CIRG-EKGV-----YNFPTDKKPKQLPEGSVRPL-----	98
At2g23340	YLRG-PSAR-----LNFPELLAGLTVSNGGGR-----	98
At4g36900	YLRG-PSAR-----LNFPELLAGVTVTGGGGGG-----	101
At5g67190	YLRG-PTAR-----LNFPELLPGEKF-----	85
At3g50260	YLRG-PTAT-----LNFPELLPCT-----	83
CaERF06	YLRG-PTAR-----LNFPELLVGDTTAAN-----	87
At4g06746	YLRG-PSAR-----LNFPEEVFKDNGGEG-----	104
At1g46768	YLRG-PSAR-----LNFPDLLLQEEDHLSAAT-----	101
At1g21910	CLKG-PQAN-----LNFPTSSSSHLLLDNLLD-----	120

At1g77640	CLKG-PKAN-----LNFPNITTTSPFLMN-ID-----	112
At1g44830	CLKG-SSANN-----LNFPEISTSLYHIINNGD-----	106
At1g22810	CLHQ-PDSL--ES-----LNFPHLLNPS-----	78
At1g71520	CLHR-PSSLDDDES-----FNFPHLLTTS-----	80
At1g19210	CLRG-NNAK-----FNFPDNPPVIS-----	75
At1g74930	CLRG-GDAN-----FNFPNPPSIS-----	84
At5g21960	CLRG-PGAR-----FNFPDNPEIP-----	70
At2g40220	YLYG-SRAQLNLT-----PSSPSSVSSSSSSVS-AA-----	128
At1g71450	HFRG-REAR-----LNFPELASSLP-----	87
At1g33760	HLRG-RGTN-----LNFPELVDSFP-----	83
At1g63040	ALKG-GDTL-----LNFDPDSLGSYP-----	152
At1g12630	ALKG-REAV-----LNFPGSVGSYP-----	81
At5g52020	ALKG-PDAV-----LNFPGALALTYV-----	127
CaERF08	ALKG-PDAI-----LNFPELAPTYP-----	136
CaERF07	ALKG-PDAI-----LNFPELAPTYP-----	136
At2g36450	ALKG-SQAE-----LNF PNSVSSLP-----	79
At4g25470	ALRG-RSAC-----LNFADSAWRLR-----	114
At4g25480	ALRG-RSAC-----LNFADSAWRLR-----	114
At4g25490	ALRG-RSAC-----LNFADSAWRLR-----	111
At5g51990	ALRG-RSAC-----LNFADSAWRLR-----	117
At1g63030	ALRG-RSAC-----LNFSDSAWRLP-----	93
At1g12610	ALRG-RSAC-----LNFADSAWRLP-----	93
At3g60490	AIKG-NSGF-----LNFPELSGLLP-----	134
CaERF09	AIKG-HSAY-----LNFPHLAHELP-----	187
At2g35700	AIKG-GSAH-----LNFPELAYHLP-----	108
At4g16750	AIKG-GSAH-----LNFPELAYHLP-----	103
At2g44940	AIKG-TTAY-----LNF PKLAGELP-----	163
At5g25810	SIKG-ASAI-----LNF PDLAGSFP-----	99
At5g11590	SIKG-TAAI-----LNFPELADSF P-----	114
At2g25820	SIKG-SSAI-----LNFPELADFLP-----	89
At4g32800	SIKG-TSAI-----LNFPELSKLLP-----	82
At1g77200	TIKG-TSAV-----LNFPELATYLP-----	106
At3g16280	TIKG-SSAV-----LNFPELASLFP-----	124
At3g11020	AMYG-SLAR-----LNF PQSV-----	137
At5g05410	AMYG-PLAR-----LNF PRSD-----	138
CaERF10	SMYG-PCAR-----LNL PNYG-----	141
At2g40350	AIYG-QSAR-----LNL PLLP-----	126
At2g40340	AIYG-QSAR-----LNLPEITNRSSTAATATV-----	143
CaERF11	VMYG-PSAQ-----LNL PNYN-----	140
At2g38340	VMYG-PYAR-----LNF PEDLG-----	138

At1g75490	KLYG-PEAH-----LNLPESLR-----	93
At5g18450	KLYG-HEAK-----LNLVHPQQ-----	94
At3g57600	KLYG-HDAY-----LNLPHLQRN---TRPSLSN-----	96
CaERF12	KLYG-TAAK-----LNLPHLYR----RP-----	102
At5g11190	LMNG-QNAK-----TNFPVVKSEEGSD-----	72
At5g25390	LMNG-QSAK-----TNFPVIKSN-GS-----	70
At1g15360	LMSG-RNAK-----TNFPLNNNNTGETSEGKTDI-----	79
At5g25190	LMCG-PRAR-----TNFPYNPNA-----	69
At5g19790	LLRG-SNTR-----TNFANHFPPNNSQLSLKIR-----	92
At3g15210	DFRG-AKAK-----TNFPTFLELSQKVPTGFAR-----SPSQSS	103
At1g53170	DFRG-VKAK-----TNFGVIVGSS-----PTQSS	98
At1g28370	EFRG-AKAK-----TNFPCYNINA-----HCL	85
CaERF22	DFRG-PKAK-----TNFPSPSEIQ-----SPSNSS	93
CaERF23	DFRG-PKAK-----TNFPSPSEIQ-----SPSNSS	93
CaERF25	EFRG-PKAK-----TNFPSPSEIQ-----SPSNSS	93
At1g03800	RFRG-SKAT-----TNFPLIGYYG---ISSATPV-----NNNLSE	128
At5g44210	EFRG-SKAK-----TNFPLPGE-----ST	95
At1g50640	NLRG-PKAK-----TNFPIDSSPPPPNLRFNQIRNQN-----QNQVD	109
At3g20310	NLRG-PKAK-----TNFPIDCS-PSSPLQPLTYLHNQNLCSPPVIQNQID	114
CaERF24	TLRG-PKAK-----TNFPLPSTSSYHHPHQFSNYQNPN--NPNFNNDPA	119
CaERF26	SLRG-PKAK-----TNFGSPALTPMSFTP-----	103
At1g28360	FLRG-IKAK-----TNFSPSLSLDLN-----	76
At1g12980	AFRG-AKAR-----TNFTYPTAVIMPEPRFSFS-----	128
At5g13910	SMRG-TRAR-----TNFVYSDMPPS-----SSVTS	88
At5g18560	SMRG-TQAR-----TNFVYTPTDVH-----TILTN	123
At1g28160	SIRG-LTAR-----TNFVYSDMPRG-----SSVTS	107
At1g12890	SISG-SLAT-----TNFFYTENTSL-----ERHP-	88
At1g24590	AMRG-LKAR-----TNFVYPMPSLD-----SYHHR	126
At2g47520	KIRG-RKAK-----LNFPN-----	107
At1g72360	RIRG-AKAK-----LNFPNESSGKRKAKA-----	143
CaERF14	RIRG-DKAK-----LNFVADDQPPP-----	139
CaERF15	RIRG-DKAK-----LNFVADDQPPP-----	139
CaERF16	RIRG-DKAK-----LNFVADDQPPP-----	139
At3g16770	QIRG-DKAK-----LNFDDLHPPPPNYT-----	146
CaERF17	KIRG-KKAK-----VNFPNE-----	145
CaERF18	KIRG-KKAK-----VNFPNE-----	82
At1g53910	RIRG-SKAK-----VNFPEE-NM-----	185
At3g14230	RIRG-TKAK-----VNFPEEKNP-----	189
CaERF19	RIRG-KKAK-----LNFPDD-----	183
CaERF20	RIRG-KKAK-----LNFPDE-----	107

CaERF21	RIRG-KKAK-----VNFPDX-----	190
At5g61890	KFKG-SKAK-----LNFPERVQLGSNS-----	155
At5g07310	KFKG-SKAK-----LNFPERAQLASNTSTTT-----	161
CaERF36	RFKG-NKAK-----LNFPERVQ GKTEY-----	120
At5g13330	KFKG-TKAK-----LNFPERVQGP TTTTTIS-----	108
At1g43160	RFRG-SKAK-----LNFPENVTQT IQRN-----	128
At5g50080	RFRG-NKAK-----LNFPEDVRILPP-----	141
CaERF37	RFRG-NRAK-----LNFPENVRLLP LQASSQSAQLHAV-----	119
At5g64750	RFRG-NKAK-----LNFPENVKLV RPASTEAQP VHQTA-----	262
At2g33710	EFRG-HKAK-----LNFP EHIRVNPTQLYP-----	138
At2g31230	ATKG-SLAT-----LNFPVEVVRESLKKMENVNLH-----	160
At1g06160	ALKG-SLAV-----LNFPADVVEESLRKMENVNLN-----	156
At3g23240	SMRG-SSAI-----LNFS AERVQESLSEIKYT-YE-----	155
At5g43410	SMRG-QAAI-----LNFPHEY--NMGSGVSSSTAM-----	87
At1g04370	SMRG-KAAI-----LNFPHEY--NMGT-----	84
At3g23220	GMRG-QRAI-----LNFPHEY--QMMKDGPN-----	82
At3g23230	NLRG-HLAI-----LNFPNEYYP RMD DYSLRPPYA-----	94
At4g18450	RIRGTQKAH-----TNFQLETVIKAMEMDCNP NYR-----	186
At4g17500	RMRG-SRAL-----LNFP LRVNSGEPDPVRIKSKRS-----	223
At5g47220	RMRG-SRAL-----LNFP LRVNSGEPDPVRI TSKRS-----	192
CaERF27	RMRG-SKAL-----LNFP HRVGLNEPEPVRVTAKRR-----	195
CaERF28	XMRG-XKAL-----LN-----SLTEX-----A-----	181
CaERF29	RMRG-SKAL-----LNFP HRVGLNEPEPVRVTAKRR-----	195
At2g44840	QLRG-SKAK-----LNFP H LIGSCKYEPVRIRPRRR-----	167
CaERF30	RMRG-AKAL-----LNFP SEIVTASESSILRQRVHH-----	136
CaERF31	RMRG-AKAL-----LNFP SEIVTASESSILRQRVHH-----	136
At5g47230	RLRG-SKAI-----LNFP LEV GKWK PRADEGE-----	227
At4g17490	RLRG-SKAI-----LNFP LEVDKWN PRAEDGRGL-----	210
At5g61600	QLRG-RKAI-----LNFP L DVRTSETCSGEGVIGL-----	163
At5g51190	KLRG-SKAI-----LNFP LEAGKHEDLGDNKKTISL-----	147
CaERF32	RLRG-SKAI-----LNFP HEVGNNTLPESELATTSG-----	235
CaERF34	RLRG-SKAI-----PHFP HXVEDNTLPERELAMTSG-----	234
CaERF33	RLRG-SKAI-----LIX-HTEXGNILX-PKVS WRRL-----	232
At5g07580	KLRG-RKAV-----LNFP L DAGKYEAPANSGR-----	248
At5g61590	KLRG-RKAV-----LNFP L DAGKYDAPVNSCR-----	178
At1g80580	RLRGTAKACTN-----FVIPPLFPKEIA-----	182
CaERF35	WDVKTWVEN-----PQLQP-----	158
At2g20350	EFVG-----	135
At5g67000	EYR-----	168
At5g67010	ALVE-----	138

At1g25470	EYDALASSG-----SAVSSSVVT--VTSQCLRSP-----	152
At1g68550	EFDIVAGN-----VSTTKRDVSSSETSQCRRSPVVPV-----	191
At3g25890	EFDALAAAT-----SAASSVLSN-----	193
At4g11140	-----FIIGGGECLRSPVSVLESFSGESTAVKEEFVGVST---AEIVV	204
At4g23750	-----SRSSNDCLCSPVSVLRSPFA-----VDEFSGISSSPVAVVV	253
CaERF13	-----RQNMSLCRPRQCSVMMMR-----LV	218
At5g53290	-----STEEDLQHLSSPTSVLNHR-----SEE-IQQVQQPFKSA	248
At4g27950	-----ESMDDSQHLSSPTSVLNYQTY-----VSEEPIDSLIKPVKQE	235
At2g46310	-----SARESQISLCSPTSVLRFSHN-----DETEYR	208
AT2G20880	-----PTPVHQKQCETGTTSEDSSRRGEDDSSTALAVGGVSEE-----	289
AT4G28140	-----SELHDSNSSDQKEPETPQPSEVNLESKELPVIDVGRE-----	243
AT1g22190	-----YQPLQSSVDAKLEAICQNLAEETTQKQVRSTKKSSSRKR-----	184
AT1g78080	-----YKPLHSSVDAKLEAICKSMAET-QKQDKSTKSSKKREKVS-----	263
AT1G36060	-----YGPIQAAVDAKLEAI---LAEPKNQPGKTERTSRKRAK-----	249
CaERF01	-----YKPLHSSVDAKLQAICQSLAEGKGIDGKSKGTRRCAA-----	203
CaERF02	-----YKPLHSSVDAKLQAICQSLAEGKGIDGKSKGTRRCAA-----	155
CaERF05	-----	
AT2g22200	-----MNPLPSSVDTKLQAICKSLRKTEEICSVSDQTKEYSVYSVSDKTELF	180
AT4g39780	-----FNPLHSSVDAKLQEICQSLRKTEDIDLPCSET-----	195
CaERF04	-----FKPLHSSVDAKLEAICQSLADSQKQGNTGEQHQQKVG-----VQT	276
CaERF03	-----FKPLHSSVDAKLEAICQSLADSQKQGNTGEQHQQKVG-----VQT	120
AT5g65130	-----KQILSPSINAKIESICNSSDLPLPQIEKQNK-----	207
At4g31060	-----SKLDIQTITATNYASSVVHVPSHATTLPATTQVPSEVPAS-----	137
At2g23340	-----GGDLAAYIRRKAAEVGAQVDALGATVVVN-----TGGENRG--	135
At4g36900	-----VNGGGDMSAAYIRRKAAEVGAQVDALEAAGAGGNRHHHHHQHQRG--	146
At5g67190	-----SDEDMSAATIRKKATEVGAQVDALG-TAVQNN-RHRVFGQNRDSD	128
At3g50260	-----SAEDMSAATIRKKATEVGAQVDAIGATVVQNNKRRRVFSQKRD--	126
CaERF06	-----GGDLAAAIRKKAIEVGSRVDA-----	109
At4g06746	-----LGGDMSPTLIRKKAEEVGARVDAEL-----	129
At1g46768	-----TADMPAALIREKAAEVGARVDALLASAAPS-----	131
At1g21910	-----ENTLLSPKSIQRVAAQAAN-SFNHFAP-----	146
At1g77640	-----EKTLSPKSIQKVAAQAANSSSDHFTPPSDENDHDHDDGLDHHPS	157
At1g44830	-----NNNDMSPKSIQRVAAAAAANTDPSSSS-----VS	136
At1g22810	-----LVSRTSPRSIQQAASNAGMAIDAGIVH-----	105
At1g71520	-----LASNISPKSIQKAASDAGMAVDAGFHG-----	107
At1g19210	-----GGRNLSRSEIREAAARFAN-----SA	96
At1g74930	-----VEKSLTPPEIQEAAARFANTFQDIV-----KG	111
At5g21960	-----GGRSLTPQIQVVASRFACEEELLPEQ-----HHPSPPRGDHNT	110

At2g40220 -----SSPSTSSSSTQTLRPLLPRPAAATVGGGANFGPYGIPFNNNIFLN 173
 At1g71450 -----RPADSSSDSIRMAVHEATLCRTTEGTESAMQVDSSSSSNVAPTMV 132
 At1g33760 -----RPESSSSEHIQAAAQDAALMFK-PGRLSEPALESGQ----GLSRV 123
 At1g63040 -----IPLSSSAHIRCAAAAAAATRGAAG--AAVKVGQKK----- 186
 At1g12630 -----VPESTSAADIRAAAAAAMKGC EE EEEKKAKEKKSSSSKSRAR 126
 At5g52020 -----APVSN SAADIRAAASRAAEMKQPDQGGDEKVL EPVQPG----- 165
 CaERF08 -----IPASLSAADVRAAAAASAAAAREPLQEPPVLDQAPATSPRD----- 176
 CaERF07 -----IPASLSAADVRAAAAASAAAAREPPQEPAVLDQAPATSPRD----- 176
 At2g36450 -----APTSMSPADIQAAAASAAAFAAARD AIVMANNNSQTSG----- 118
 At4g25470 -----IPESTCAKEIQKAAAEAAALNFQDEMCHMTTDAHGLDMEETLVEAI 159
 At4g25480 -----IPESTCAKDIQKAAAEAAALAFQDEMCDATTD-HGFDMEETLVEAI 158
 At4g25490 -----IPESTCAKDIQKAAAEAAALAFQDETCDTTTTNHGLDMEETMVEAI 156
 At5g51990 -----IPETTCPKEIQKAAASEAAMAFQNETTTEGSK-TAAEAEAAEGEV 161
 At1g63030 -----VPASTDPDTIRR TAAEAEMFRPPEFSTGITVLP SA----- 129
 At1g12610 -----VPESNDPDVIRRVAAEAEMFRPVDLESGITVLP CAGDDVDLGF G 138
 At3g60490 -----RPVSCSPKDIQAAATKAAEATTW----HKPVIDKKLADEL SHSEL 175
 CaERF09 -----RPATTSPKDIQAAA AKAATAAT-----FPGENGGGAEAEPIQAE 225
 At2g35700 -----RPASADPKDIQAAAAAAA AV-----AI 131
 At4g16750 -----RPASADPKDIQEAAAAAAA----- 122
 At2g44940 -----RPVTNSPKDIQAAASLAAVNWQDSVNDVSNSEVAEIVEAEPSRAV 208
 At5g25810 -----RPSSLSPRDIQVAALKA AHMET--SQSFSS----- 127
 At5g11590 -----RPVLSLSPRDIQTAALKA AHMEP--TTSFSS----- 142
 At2g25820 -----RPVLSLQQDIQAAAAEAALMDFK-TVPFHLQDDSTPLQTRCDTEK 133
 At4g32800 -----RPVLSLSPRDVRAAATKAALMDFD-TTAFRSDTETSETTTSNKMSE 126
 At1g77200 -----RPASSSPRDVQAAAAVAAMDFS-PSSSSLVSDP----TTVIAP 146
 At3g16280 -----RPASSSPHDIQTAAAEAAAMVVE-EKLLEK----- 153
 At3g11020 -----GSEFTSTSSQSEVCTVENKAVVCGDVCVKHEDTDCE SNPFSQILDVR 184
 At5g05410 -----ASEVTSTSSQSEVCTVETPGCVH----VKTEDPDCE SKPFSG--GVE 179
 CaERF10 -----PQQEASIESSSFPATSAS-----DSTTTSGLSEVC 171
 At2g40350 -----LCQAR-----LLH----- 134
 At2g40340 -----SGSVTAFSDESEVCAREDTNASSGFGQVKLEDCSDEYVLLDSSQCIK 190
 CaERF11 -----VVKEASTEPSLELTMSIP-----ESTLTTS----- 165
 At2g38340 -----GGRKKDEEAESSGGYWLE----- 156
 At1g75490 -----SYPKTASSPASQTT PSSN----- 111
 At5g18450 -----QQQVVVNRNLSFSGHSGSWAYNKKLDMVHGLDLGLGQASCSRGSCS 141
 At3g57600 -----SQRFKWVPSRKFI SMFPSC----- 115
 CaERF12 -----QHIPVYSTLNTTTPIDTG----- 120
 At5g11190 -----HVKDVNSPLMSPKLSSELLNAKLRKCKDLTPSLTCLRLD TDSSHIGVWQ 122
 At5g25390 -----NSLEINSALRSPKLSSELLNAKLRKNCKDQTPYLTCLRLDNDSSHIGVWQ 120
 At1g15360 -----SASSTMSSTSSSSLSSILSAKLRKCKSPSPSLTCLRLDTASSHIGVWQ 129

At5g25190 -----IPTSSSKLLSATLTAKLHKCY-----MASLQMTKQTQTQTQTQ 107
 At5g19790 -----NLLHQKQSMKQQQQQHKPVSSLTDCNINYIISTATSLTTTTTTTTTTAIP 142
 At3g15210 TLDCASPPTLVV---PSATAGNVPPQLELSLGGGGGGSCYQIPMSR-----PV 148
 At1g53170 TVVDS-----PTAARFITPPHLELSLGGGG-ACRRKIPLVH-----PV 135
 At1g28370 SLTQS-----LSQSSTVESSFPNLNLGSDS--VSSRFPPFK-----IQ 121
 CaERF22 TVESSSGDG-----PTHAPVELDLTRRLGVAAGGGGAGSEIPILHHPQAMAVLPNGQPV 147
 CaERF23 TVESSSGDG-----PTHAPVELDLTRRLGVXAGGGGAGSEIPILHHPQAMAVLPNGQPV 147
 CaERF25 TVESSSGDG-----STHAPVELDLTRRLGVAAGGGGAGSEIPILHHPQAMAVLPNGQPV 147
 At1g03800 TVSDGNANL-----PLVGDDGNALASPVNNTLSETARDGTLPSDCHDMLSPGVAEAVAG 182
 At5g44210 TVNDG-----GENDSYVNRTTVTTAREMTRQ-RFPFACHRERKVVGGYASAG 141
 At1g50640 PFMDH---RLFT---DHQQQFPIVNRPTSSSMSSTVESFSGPRPT-----TM 150
 At3g20310 PFMDH---RLYGGGNFQEQQQQIIISRPASSSMSSTVKSCSGPRPM-----EA 159
 CaERF24 MFIDHPQNRLYH---PEPHQILHPQRPTSSSLSSSTVESFSGPRPQ-----TT 163
 CaERF26 -----PVMVQVPSVNPTGLAENFPHWCSPAFF-----AT 132
 At1g28360 -----HLPSAPSAATAAANNQPHQHQLWFAAPPPVPPS-----SDHH 114
 At1g12980 -----NKKSSPSARCPLPSLPLDSSTQNFYAGAPAAQRIYNTQSIFLRDAS 173
 At5g13910 -----IVSPDDPPPPPPAPPSPNDPVDYMMMFVNYQYSSSTDSP-----MLQP 129
 At5g18560 PNLHSLIVSPYNNNQSFPLPNSSP-QFVIDHHPHYQNYHQPKHTLPQTVLPAASFKTP 182
 At1g28160 -----FVSPDESQRFISELFNPPSQLEATNSNNNNNNNLYSS-----TNN 147
 At1g12890 -----QQSLEPHMTWGS-----SSLCLLQDQ-----PFE 112
 At1g24590 IFSSPPMNMFLLRDVLNSQSLSPLTTFAYPPCNLSNVNDVVHE-----SFT 172
 At2g47520 -----TQVEE----- 112
 At1g72360 -----KTVQQVEENHEADLDVAVVSSAPSSSCLD-----FLWEENNPDT 182
 CaERF14 -----TTPSQPTPPRQPL-AKRRCVTPEPPPAYY-----PQPDQEILES 177
 CaERF15 -----TTPSQPTPPREPL-AKRRCVTPEPPPAYY-----PQPDQEILES 177
 CaERF16 -----TTPSQPTPPRQPL-AKRRCVTPEPPPAYY-----PQPDQEILES 177
 At3g16770 -----PPSSPRSTDQPP-AKKVCVVS-----QSESELSQ- 175
 CaERF17 -----DSTSHPT-----CIPS-----QPQYQVH- 164
 CaERF18 -----DSTSHPT-----CIPS-----QPQYQ- 98
 At1g53910 -----KANSQKR-SVKAN-LQKPVAKPN-----PNPSPALVQN 216
 At3g14230 -----SVVSQKRPSAKTNNLQKSVAKPNKSVTLV-----QQPTHLSQY 228
 CaERF19 -----TQLSASRRTVVAN-AQKVRVKEDQSQ-----AQPSL NESMT 218
 CaERF20 -----TQLSASRRTVVAN-PQKVRVKEDQSQ-----VQPSL NESMT 142
 CaERF21 -----TQGSASRRPVAVN-PQKLQVKENPGP-----VQPTLSESM 225
 At5g61890 -----TYYSSNQ-----IPQMEPQSI PNYNQYYH-----DASSGDMLSFN 190
 At5g07310 -----GPPNYSSNNQIYYSNPQTNPQTIPIFYFNQYYYNQYLHQGGNSNDALSYS 210
 CaERF36 -----GYLTTTRQ-----DARVVA----- 133
 At5g13330 -----HAPRGVSESMNSPPPRPGPSTTTTTSWPMTYN-----QDILQYAQL 149
 At1g43160 -----SHFLQNS-----MQPSLTYIDQCPTLLS----- 151
 At5g50080 -----PPLLRSPADTVANKA---EEDLIN-----YWSYTKLLQSSG- 175

CaERF37	-----SSTPPATILPVTTTPSPQARMFQPQPIQSSDIAR----	DYWEYSQLLRSSGD	166
At5g64750	-----QRPTQSRNSGSTTTLLPIRPASNQSVHSQPLMQS---	YNLSYSEMARQQQQ	310
At2g33710	-----SPATSHDRIIVTPPPPPPPIAPDILLDQYG-----		168
At2g31230	-----DGGSPVMALKRKHSLRNRPRGKKRSSSSS-----	SSSSNSSS	197
At1g06160	-----DGESPVIALKRKHSMRNRPRGKKKSSSSST-----	LTSSPSSSSS	196
At3g23240	-----DGCSPVVALKRKHSMRRRMTNKK-----	TKDSD	183
At5g43410	-----AGSSASASASSS-----	SRQVFE	106
At1g04370	-----GSSSTAANSSSS-----	SQQVFE	102
At3g23220	-----GSHENAVASSSSGYRGGGGDD-----	GREVIE	110
At3g23230	-----SSSSSSSGSTSTNVSQN-----	QREVFE	119
At4g18450	-----MNNSNTSDPLRSSRKIGLRTGKEAVKAYDEV----	VDGMVENHCA	227
At4g17500	-----S--FSSSN-----ENGAPKRR-----	TVAA	242
At5g47220	-----S--SSSSSSSSTSSSENGKLR-----	KAEN	219
CaERF27	-----SPEXAASSPSSVSSASESGSPSP-----	KREE	224
CaERF28	-----		
CaERF29	-----SPEQAASSSSVSSASESGSPSP-----	KRRR	224
At2g44840	-----SPEPSVSDQLTSEQKRESHVDDGESS-----	LVPPEL	199
CaERF30	-----NHAKLNLDRQESNSSSSSTEVSST-----	GLGLAS	168
CaERF31	-----NHAKLNLDRQESNSSSSSTEVSST-----	GLGLAS	168
At5g47230	-----KKRKR-DDEKVTVVEKVLKTEQSVDVNGG----	ETF-----P	260
At4g17490	-----YNKRKRDEEEEEVTVVEKVLKTEESYDVSGG----	EN-----	243
At5g61600	-----GKRKRDKGSPPEEKAARVKVEEESNTSET----	TEAEVPPVP	204
At5g51190	-----KAKRKRQVTEDESQILSRKAVKREEAQVQAD-----	ACP	181
CaERF32	-----RKRAREPDVVSSEDIVKKEVKREEVSPASVRK----	GEERGLPTDP	276
CaERF34	-----SKRAREPDCVSEDM-----		248
CaERF33	-----LXRKGQXSGCCE-----		244
At5g07580	-----KRRKRSVHEELQRTQSNSSSSSCDAF-----		274
At5g61590	-----KRRRTDVPQPQGTSTSTSSSSSN-----		201
At1g80580	-----QAQEDNMRMQKQKKKKKVVSRKCVKVT-----		212
CaERF35	-----		
At2g20350	-----RRSARRGTKNGEEASTKKTTEKN-----		158
At5g67000	-----ARRGKTNGE--GIKRRWR-----		184
At5g67010	-----KRSARRGSKKGE-GSIHQEVGGDD-----		162
At1g25470	-----TSAS-VSCVS-ADDLSKEK-TSLNKDVAASGDSTTKEVFTTDFSD		195
At1g68550	-----EQDDTSASALTCVNNPDDVSTVAPTPTPNVPAGGN---	KETLDFDFFTN	238
At3g25890	-----ESGSMISASGSSIDLKKLVDSTLDQQAGESK---	KASFDFDFAD	235
At4g11140	KKEPSFNG-----SDFSAPLFSDDDVFGFSTSMS-ESFGDLFGDNLF-	ADMSFGSGFG-	256
At4g23750	KEEPSMTTVSETFSDFSAPLFSDDDVDFRSSVVPDYLGGDLFGEDLFTADMCTDMNFG-		312

CaERF13	KRRQRFRLLELELATPFHQPLR-----FSVAATA-----	246
At5g53290	KPEPGVSNAPWWHTGFNTGLGESDDSFPLDTPFLDNYFNESPPEMSIFDQPMQIFCEND	308
At4g27950	FLEPEQEPISWHLGEGNT--NTNDDSFPLDITFLDNYFNESLPDISIFDQPMSPIQPTEN	293
At2g46310	TEPTEEQNPFFLPDLFRSGDYFWDSEITPDPLFLDEFHQSLLPNINNNNTVCDKDTNLS-	267
AT2G20880	-----TGWAEAWFNAIPEEWG	305
AT4G28140	-----EGMAEAWYNAITSGWG	259
AT1g22190	---SSTVAVKLPPEEDYS-SAGSSPLLTESYG-----SGGSSSPLSE-LTFGDT-	227
AT1g78080	---SPDLSEKVKAEENSVSIGGSPPVTEFEE-----STAGSSPLSD-LTFADP-	307
AT1G36060	-----AAASSAEQPS----APQQHSGSGE-----SDGSGSPTS-VMVQEM-	285
CaERF01	-----EEEKAAAKTVDAGSESETGSGH-----LTDGSSPVSD-LTVPDFT	242
CaERF02	-----EEEKAAAKTVDAGSESETGSGH-----LTDGSSPVSD-LTVPDFT	194
CaERF05	-----SSSP-----	223
AT2g22200	LPKAELFLPKREHLETNELSNESPRSDETS-----LLDE-SQAEYSSSD--KTFLDIFS	230
AT4g39780	----ELFPPKTEYQES---EYGFRLRSDENS-----FSDE-SHVESSSPESGITTFLDIFS	241
CaERF04	IPKVENTFDDSLNSEYVYPGNEDVKVEETSSSLSPSRSDETSSLEMSSPESE-ITFLDFA	335
CaERF03	IPKVENTFDDSLNSEYVYPGNEDVKVEETSSSLSPSRSDETSSLEMSSPESD-ITFLDFA	179
AT5g65130	----EEVLSGFASKPEK-----EPEFGEIYG-----CGYSGSSPESD-ITLLDIFS	246
At4g31060	-----SDVSASTEITEMVDEYYL	155
At2g23340	-DYEKIEN-----C----R-----KSGNGSLERVDLNKLDPENSDGDD	169
At4g36900	-NHDYVDNHSDYRINDDLME----C----SSKEGFKRCNGSLERVDLNKLDPETSDDD-	196
At5g67190	VDNKNFHRNYQNGEREEEEEE---D---EDDKRLRSGRLLDRVDLNKLDPPESSDEEW	180
At3g50260	-----FGGGLLELVDLNKLDPENLDDDL	150
CaERF06	-----	
At4g06746	-----RLENRMVENLDMNKLPEAYGL----	150
At1g46768	-----MAHSTPPVIKPDNLNQPESGDI----	153
At1g21910	-TSSAVSSPSDHDHHDHDDG-----MQSLMGSFVDN--HVSLMDSTSSWYDD	189
At1g77640	ASSSAASSPPDDHHDHDDG-----DLVSLMESFVDYNEHVSLMDP--SLYEF	203
At1g44830	TSSPLLSSPSED-----LYDVVMSQYDQQVSLSES--SSWYNC	173
At1g22810	--STSVNSGCGD-----TTTTYYENGADQVEP-----	129
At1g71520	--AVSGSGGCEE-----RSSMAN--MEEEDK-----	129
At1g19210	EDDSSGGAGYEIRQESA-----STSMVDVSEFLSMLPTVG-SGNFASEFGL	141
At1g74930	EEESGLVPGSEIRPESP-----STASAVATSTVDY-----DFSFLDLLPM	151
At5g21960	EEEVIIISARGEINSGSGGPTLGQVG----EDNNNEGNSNDTSSYWPLIWEENFVGGPPNS	166
At2g40220	GGTSMCLCPSYGFPPQQQQQQNQMVQ---MGQFQHQQYQNLHSNTNNNKISDIELTDVPV	229
At1g71450	RLSPREIQAINESTLGSF-----TTMMHSTYDPM-----FANDVE	168
At1g33760	GLSPDQIQAINESPLDSPR-----MGWMQDLEVADYEELYGQFFGQHDR	167
At1g63040	EDKVYDTAESSTMG-----FVDEEELLNMPGLLADMAKG	220
At1g12630	ECHVDNDVGSSSWCG-----TEFMDEEEVLNMPNLLANMAEG	163
At5g52020	KEEELEEVSNCNSCS-----LEFMDEEAMLNMPNLLTEMAEG	201
CaERF08	PVEINEDINASTSGDNQM-----GQDQFVDEEELFDMPNLLVDMMAEG	218

CaERF07 PVEITR---TSTRR-----LP-VIIKWDRS 197
 At2g36450 VACMNSSYDNTNMNG-----FMDEDLVFDMPNVLMMNMAEG 153
 At4g25470 YTP-----QSQDAFY-----MDEEAMLGMSLLDNMAEG 189
 At4g25480 YTAE-----QSENAFY-----MHDEAMFEMPSLLANMAEG 188
 At4g25490 YTP-----QSEGAFY-----MDEETMFGMPTLLDNMAEG 186
 At5g51990 REGERRAEEQNGGVFY-----MDDEALLGMPNFFENMAEG 196
 At1g63030 -----SEFDTSD-----EGVAGMMMLAE 149
 At1g12610 SGSGSGSGSEERNSSS-----YFGFDYEEVSTTMMRLAEG 173
 At3g60490 LSTAQSSTSSSFVFSDD-----TSETSSTDKESNEETVFDLPDLFTD--GL 219
 CaERF09 LTNCHSSTDFSLHNMQE-----SSNSPSTD-DDIDDTFFDLPLSFDVTGA 270
 At2g35700 DMDVETSSPSPSPTVTE-----TSSPAMIA--LSDDAFSDLPDLN--VN- 173
 At4g16750 ---VDWKAPE-SPSSTV-----TSSP-----VADDAFSDLPDLN--VND 157
 At2g44940 VAQLFSSDTSTTTTTQSQEYSEA-----SCASTSACTDKDSEEEKLFDLPDLFTDENEM 262
 At5g25810 SSSLTFSSSQSSSLES-----LVSS--SATGSEELGEIVELPSLGSSYDG- 171
 At5g11590 STSSSSLS-STSSLES-----LVLVMDLSRTESEELGEIVELPSLGASYDVD 189
 At2g25820 IEKWSSSSSSASSSSS-----SSSS--SSMLSGELGDIVELPSLEN--NV- 176
 At4g32800 SSESNETVSFSSSSWSS-----VTSIE--ESTVSDDLDEIVKLPSLGTSLNE- 171
 At1g77200 AETQLSSSSYSTCTSSS-----LSPSSEEAATAEELSEIVELPSLETSYDES 194
 At3g16280 ---DEAPEAPSSSESS-----YVAEE--SEDEERLEKIVELPNIEEGSYDE 194
 At3g11020 EESCGRPDSCTVGHQDMNSSLNYD--LLEFEQYWGQVLQEKEKPKQEEEEIQQQQE 242
 At5g05410 PMYCLN-----GAEEMKRGVKADKHWLSEFEHNYWSDILKEKEK--QKEQGIVETCQQ 231
 CaERF10 PP-----ADRDAEPNSN-----VKEEDG--EGESRIQDTR-- 199
 At2g40350 -----FLMN-----LKF-VHVRIQ 147
 At2g40340 EELKGKE-----EVREEHNLAVGFGIGQDSKRETLDAWLMGNGNEQEPLEFGVDETFD 243
 CaERF11 -----HAFAKAEDSKVTSETSYH 184
 At2g38340 -----TNKAGNGVIETEGGKDYVVY 176
 At1g75490 -----TGGKSSSDSESPCSSNEMSS 131
 At5g18450 ERSSFLQEDDDHSHNRCSSSSGSNLCWLLPKQSDSQDQETVNATTSYGGEGGGSTLTFS 201
 At3g57600 -----GMLNVNAQPSVHI IQQRLEELKK 138
 CaERF12 -----GENQTGLRP-VPFQQSPFRNFAS 142
 At5g11190 KRAGSKTSPTWVMRLELG-----NVVNESAVDLGLTTMKNQVKEKEE 165
 At5g25390 KRAGSKTSPNWKLVLELG-----DKVNARPG--GDIETNKMKVRNEDV 161
 At1g15360 KRAGSKSDSSWVMTVELGPA-----SSSQETTSKASQDAILAPTTEVEIGGS 176
 At5g25190 TARSQSADSDGVT-----ANESHLNRGVTETTEIKWEDGNA 143
 At5g19790 LNNVYRPDSSVIGPETEG-----LQLPYSWPLVSGFNHQIPLAQAGGET 187
 At3g15210 YFLDLMGIGNVGRGQPPPVTSAFRSPVVHVATKMACGAQSDSDSSSVVDF---EGGMEKR 205
 At1g53170 YYYNMATY-----PKMT-----TCGVQSESETSSVVDV---EGGAGKI 170
 At1g28370 VKAGMMVF-----DERSESDSSSVVMD---VVRYEGR 150
 CaERF22 VFFQPMALNRADYPYRFGPVAMDYANVVGRG-RGGGAVHSDSDTSSVDDNISDGDNMKR 206
 CaERF23 VFFQPMALNRADYPYRFGPVAMDYANVVGRG-RGGGAVHSDSDTSSVDDNISDGDNMKR 206

CaERF25	VFFQPMALNRADYPYRFGPVAMDIYANVVG-----	176
At1g03800	FFLDLPEVIALKEELDRVCPDQFES-----IDMGLTIGPQTAVEEPETSSAVDCKL	233
At5g44210	FFFDPSPRAASLRAELSRVCPVRFDP-----VNIELSIG-----IRETVKVEP	183
At1g50640	KPATT-----KRYPRTPPVVP-----EDCHSDCDSSSSVIDDDD-DIASSSR	191
At3g20310	AAASSSVAKPLHAIKRYPRTPPVAP-----EDCHSDCDSSSSVIDDGD-DIASSSS	209
CaERF24	AAAATSMIPS--SSRRHPRSPPVVP-----DDCHSDCDSSFSVIDDGDGCDIASSSI	212
CaERF26	DEADR-----STERAARS-----EYTGKLETIGVMMSEREMD-----	165
At1g28360	HQHHRIFLRTGVLNDKTSYSS-TEAPLYFTSSPNTATSSPGYQVVGFPMMNSSPSPVTV	173
At1g12980	CSSRKTTTPYNNSFNGSSSSYSASKTACVSYSENNENESFFPEESDGTGLLQEVVQEFLLK	233
At5g13910	HCDQ-VDSYMF-----GSQSS-----NSYCYS-----NDSSNELPPLPSDLSNS	168
At5g18560	VRHQSVDIQAFGNPQNSSSSNGSLSSSLDEENNFFSLTSEEHNKSNNNNGYLDCIVPNH	242
At1g28160	QNQNSIEFSYNGWPQEAECGYQSITSN---AEHCDHELPLPPSTCFGAELRIPETDSYW	204
At1g12890	NNHFVADPIS--SSFSQKQESSTNLTN-----TFSHCYNDGDHVGQSKEISLPNDMSNS	164
At1g24590	NVNDVCEDLSPKAKRSSTIENESLISN-----IFEPEPASSGLLQEI VQGFLPKPISQH	226
At2g47520	-----EADTKPGGNQN-----ELISENQVESLSEDLMALEDYMRFY-	148
At1g72360	LLIDTQWLEDIIMGDANKKHEPNDS-----EEANNVDASLLSEELLAFENQTEYFS	233
CaERF14	IGPTPPVVDPGFQPQPYYP--SQVT-----DQDYQYGDQFSSLESFGLGLEPETT-TT	226
CaERF15	IGPTPPVVDPGFQPQPYYP--SQVT-----DQDYQYGDQFSSLESFGLGLEPETTPQ	227
CaERF16	IGPTPPVVDPGFQPQPYYP--SXVT-----DQDYQYGESFSSLESFGLGLDLRRH--H	225
At3g16770	--PSFPVEICIGFNGDEFQNL SYGF-----EPDYDLKQQISSLESFLELDGNTA--E	223
CaERF17	-IPSPPTVC-----GPSSSNASSISSVLGRLIAMPVTAITR	199
CaERF18	-----	
At1g53910	SNISFEN----MCFMEEKHQVSNNNN-----NQFGMTNSVDAG-CNGYQYFSSDQGSNS	265
At3g14230	CNNSFDNSFG-DMSFMEEKPQMYNN-----QFGLTNSFDAGGNGYQYFSSDQGSNS	279
CaERF19	FMSNIENDHCNSLNLVEEKPPAKQFSSVDAYPTVGDMEFKSYFQSDCANLCFNSDQGSSS	278
CaERF20	FMSNIENDHCNSLNLVEEKPPAKQFSSVDAYPTVGDMEFKSYFQSDCANPCFNSDQRSSS	202
CaERF21	FTSSMENDYYNSLNLGERKPPTKQX-----WLC-----	253
At5g61890	LGGGYGSGTGYSMSHDNSTTTAATT-----SSSSGGS-----SR	224
At5g07310	LAGGETGGSMY-----NHQTLSTTN-----SSSSGGS-----SR	239
CaERF36	-----DQQVSHSH-----SNRSFPP-----PH	150
At5g13330	LTSNNEVDLSYYTSTLFSQPFTSPS-----SSSSSQ-----QT	183
At1g43160	-----YSRCMEQQQ-----PLVGMLQ-----PT	169
At5g50080	---QR--SFLERGOEESNIFEHSP-----MEQPLPSSSGP-----	207
CaERF37	FQGGQPAGLLEQMYASSLALLHSHSLPSTTTTTFAASPTPSSSASHPLLLPGQQRNNTF	226
At5g64750	FQQHHQQSLDLYDQMSFPLRFGHTGG-----SMMQSTSSSSSHSRPLFSPA AVQPPP	362
At2g33710	-----HFQSRSSDSSAN-----LSMNMLSS-----SS	190
At2g31230	CSSSSSTSSTSRSSSKQSVVK-QESG-----TLVV	226
At1g06160	YSSSSSSSLSSRSRKQSVVMTQESNT-----TLVV	227
At3g23240	FDHRSVKLDN-----VVV	196
At5g43410	FEYLDSSV-----LEE	117

At1g04370	FEYLDDSV-----LDE	113
At3g23220	FEYLDDSL-----LEE	121
At3g23230	FEYLDDKV-----LEE	130
At4g18450	LSYCSTKEHSETRGLRGSEETWFDLRKRRRSNEDSMCQEVEMQKTVTGEETVCDVFGLFE	287
At4g17500	GGGMDKGLTV-----KCE	255
At5g47220	--LTSEVVQV-----KCE	230
CaERF27	KSVXAEISAK-----PRA	237
CaERF28	-----	
CaERF29	KVVAEELSAE-----PEP	237
At2g44840	DFTVDQFYFD-----GSL	212
CaERF30	FESENAAGDL-----SNA	181
CaERF31	FESENAAGDL-----SNA	181
At5g47230	FVTSNLTELCDW-DLTGFLNFPLLSP-----LSP	288
At4g17490	-VESGLTAIDDW-DLTEFLSMPLLSP-----LSP	270
At5g61600	LTPSSWMGFWDV-GAGD--GIFSIPP-----LSP	230
At5g51190	LTPSSWKGFWDG-ADSKDMGIFSVPL-----LSP	209
CaERF32	LTPSSWTAVWDCGDVKGIFEIPPLSP-----PSP	305
CaERF34	-----	
CaERF33	-----	
At5g07580	-----	
At5g61590	-----	
At1g80580	-----VAQLFDDANFINSSSIKGNVISS	235
CaERF35	-----	
At2g20350	-----	
At5g67000	-----	
At5g67010	-----	
At1g25470	VKIPDLRFLAAEEDSMVSNANGAELDFDCFLTDSNILLDDYSLLENDINFSRFENSLPSE	255
At1g68550	LQIPDFGFLAAEEQQ-----DLDFDCFLADD--QFDDFGLLD---DIQGFEDNGPSA	284
At3g25890	LQIPEMGCFFIDDSFIP----NACELDFLLTEENNNQMLDDYCGID-DLDIIGLECDGPSE	290
At4g11140	-----FGSGSGFSSWHVEDHFQDIGDLFGSDPVLTV-----	287
At4g23750	-----FDFGSGLSSSHMEDHFQDIGDLFGSDPLLAV-----	343
CaERF13	-----	
At5g53290	DIFNDMLFLGGETMNIIEDELTSSSIKDMGSTFSDFDDSLISDLLVA-----	354
At4g27950	DFFNLDLMLFDSN----AEEYSSSEIKEIGSSFNLDLDDSLISDLLLV-----	335
At2g46310	-----DSFPLGVIGDFSSWDVDEFFQDHLLDK-----	294
AT2G20880	PGSPLWDDYHFPI SNHKDDL DATQNSSSDTI-----	336
AT4G28140	PESPLWDDL DSSHQFSSESSSSSPLSCPMRPF-----	292
AT1g22190	-EEEIQPPWEN-ALEKYPSYEIDWDSILQCSSLVN-----	261

AT1g78080	-EPPQ--WNETFSLEKYPSYEIDWDSILA-----	334
AT1G36060	-CQEPMPWNENFMLGKCPSEIDWASILS-----	314
CaERF01	EEEGSWEAYPENLMLEKYPSYEIDWAAL-----	270
CaERF02	EEEGSWEAYPENLMLEKYPSYEIDWAAL-----	222
CaERF05	-----	
AT2g22200	-DTEFEEIGSF--GLRKFPSVEIDWDAISKLANS-----	261
AT4g39780	-DSGFDEIGSF--GLEKFPSVEIDWDAISKLSES-----	272
CaERF04	-EPTFDESENF--LLQKFPSVEIDWAAL-----	360
CaERF03	-EPTFDESENF--LLQKYPSEIDWAAL-----	204
AT5g65130	-SDCVKEDESFLMGLHKYPSLEIDWDAIEKLF-----	277
At4g31060	PTDATAESIFSVEDLQLDSFLMMDIDWINNLI-----	187
At2g23340	DECVKRR-----	176
At4g36900	-----	
At5g67190	ESKH-----	184
At3g50260	VGK-----	153
CaERF06	-----	
At4g06746	-----	
At1g46768	-----	
At1g21910	HNG--MFLFDNGAPFNYSQQLNSTTMLDEYFY-----	219
At1g77640	GHN--EIFFTNGDPFDYSPQLHSSEATMDDFY-----	233
At1g44830	FDGDDQFMFINGVSAPY-----LTTSLSDDF-----	200
At1g22810	-----LNISVYDYL-----	139
At1g71520	-----LSISVYDYL-----	138
At1g19210	FPGFDDFSDEYS---GDRFREQLSPTQDYQYL-----	170
At1g74930	NFGFDSFSDDFSGFSGGDRFTEIL-PIEDY-----	180
At5g21960	DHEFGFFTDDSTNLYFPTQQQQHQLSSDFYYDG-----	200
At2g40220	TNSTSFHHEVALGQEQGGSGCANNSSMEDLNLSLAGSVGSSLSITHPPPLVDP-----	281
At1g71450	-----	
At1g33760	-----	
At1g63040	-----MMVAPP-----WMSGPPS-----	233
At1g12630	-----MMVAPPS-----WMSGRPS-----	177
At5g52020	-----MLMSPPR-----MMIHPTME-----	216
CaERF08	-----MLVSPPR-----MKPAVSG-----	232
CaERF07	-----FX-----	199
At2g36450	-----MLLSPPR-----PTVFDAAY-----	168
At4g25470	-----MLLPSP-----SVQWNYN-----	202
At4g25480	-----MLLPLP-----SVQWNHN-----	201
At4g25490	-----MLLPPP-----SVQWNHN-----	199
At5g51990	-----MLLPPP-----EVGWNHN-----	209
At1g63030	-----PLMSPPRSYID-MN-TSVY-----	166

At1g12610	-----PLMSPPRSYMEDMTPTNVY-----	192
At3g60490	-----MNPNDAFCLCNGTFTW--QLYGE--ED-----	242
CaERF09	-----TDHLDGFNCFHGSSSWWPQLHPA--PA-----	295
At2g35700	-----HN-IDGFWDS-----	182
At4g16750	-----HNKNDGFWDS-----	167
At2g44940	-----MIRNDAFCYYSSTWQLCGADAG-----	284
At5g25810	LTQLGNEFIFSDSADLWPYPQWSEG DYQM---I-----	202
At5g11590	SANLGNEFVFYDSVDYCLYPPPWGQSS EDN---Y-----	220
At2g25820	----KYDCALYDSLEGLVSMPPWLDATEND-FRY-----	205
At4g32800	----SNEFVIFDSLEDLVYMPRWLSGTEEEVFTY-----	201
At1g77200	----LSEFVYVDS-AYPPSSPWYINN CYSFYYS-----	223
At3g16280	SVTSRADLAYSEPFDCWVYPP--VMDFYEE-----	222
At3g11020	QQQQQLQPDLLTVADYGWPWSNDIVNDQTSWDPNECFDINELLGDLN-----	289
At5g05410	QQQ-----DSL SVADYGWP--NDV--DQSHLDSSDMFVD ELLRDLNGDDVF-----	274
CaERF10	-----PLTXVDAGTP-----LCTVKG-----	215
At2g40350	MQ-----DLVLRSD-----	157
At2g40340	IN-----ELLGILNDNNVSGQETMQYQVDRHPNF SYQTQFPNSNLLG-----	285
CaERF11	-----SASVPSFAGK-----EEMVTSFS DMVRKGX-----	209
At2g38340	N-----EDAIELGHDKTQNPMTDNEIVNPAVKSEEGYSYDRFK-----	214
At1g75490	CGR-----VTEEISWEHINVDLPVMD DSSIWEEATMSLGFPWVHEG-----	172
At5g18450	TNLKPKNLMSQNYGLYNGAWSRFLV GQEKKTEHDVSSSCGSSDNKES-----	248
At3g57600	TGLLSQSYSSSSSSTESKTNTSFLDEKTSKGETDNMFEGGDQKKPEIDLTEFLQQLGILK	198
CaERF12	S-----SNSSYPPQCQPNERLFFQTETNATTTGYLAG-----	174
At5g11190	EEEE--IISDEDQLAMEMIEELLNWS-----	189
At5g25390	QE-----DDQMAMQMIEELLNWTCPGSGSIAQV-----	189
At1g15360	REE---VLDEEEKVALQMIEELLNTN-----	199
At5g25190	NMQQ-NFRPLEEDHIEQMIEELLHYGSIELCSVLPTQTL-----	181
At5g19790	HGHLNDHYSTDQHLGLAEIERQISASLYAMNGANSYYDNMNAEY AIFDPTDPIWDLPSLS	247
At3g15210	SQLLDLNLNLPPPSEQA-----	222
At1g53170	SPPLDLNLNLAPP AE-----	185
At1g28370	RVVLDLNLNFPPPPEN-----	166
CaERF22	GGGLDLNLNFPPPVEA-----	222
CaERF23	GGGLDLNLIFHHLWKLEERLMR-----	228
CaERF25	-----	
At1g03800	RMEPDLNLNASP-----	245
At5g44210	RRELNLNLNLAPPVVDV-----	200
At1g50640	RRNPPFQFDLNFPPPLDCVDL FNG-ADDLHCTDLRL-----	225
At3g20310	RRKTPFQFDLNFPPPLDGVDLFAGGIDDLHCTDLRL-----	244
CaERF24	RX-AXIAIDLN-----CQRRW-----	227
CaERF26	-----	

At1g28360	RRGLAIDLNEPPPLWL-----	189
At1g12980	NRGVPPSPPTPPPVTSHHDNSGYFSNLTIIYSENMVQETKETLSSKLDRYGNFQANDDGVR	293
At5g13910	CYSQPQWTWTGDDYSSEYVHSPMPFSRMPVSDSFPQGFN-----YFGS-----	211
At5g18560	CLKPPPEATTTQNQAGASFTTPVASKASEPYGGFSNSYFEDGEMMMNHHEFGSCDLSAM	302
At1g28160	NVAHASIDTFAFELDGFVDQNSLQSGTEGFNSLPSTFFYQ-----	245
At1g12890	LFG----HQDKVGEHDNADHMKFGSVLSDEPLCFEYDYIG-----NYLQSF	206
At1g24590	ASIPPKSNQQSVGVFPTMPESGFQTDVRLADFHVEGNGFGQVKYHGELGWADHENGFDSDA	286
At2g47520	QIPVADD--QSATDIGNLWSYQDSN-----	171
At1g72360	QMPFTEGNCDSSSTLSLSSLDGGMGLWS-----	262
CaERF14	ATSQFGGITESDPVDLWMMDDLAVATAQQPNLLY-----	260
CaERF15	HLSSVGSLSNRTR-----LISG-----	243
CaERF16	STSQLGGINRTG-----	238
At3g16770	QPSQLD--ESVSEVDMWMLDDVIASYE-----	248
CaERF17	TASMIVSRSTI-----	210
CaERF18	-----	
At1g53910	FDCSEFGWSDQAPITPDISSAVINNNNSALFFEEANPAKCLK-----SMDFETPYNNTEW	320
At3g14230	FDCSEFGWSDHGPKTPEISSMLVNNNE-ASFVEETNAAKCLKPNSDESDDLMAAYLDNALW	338
CaERF19	FGCSDFGWEEQCPKTPEISSIFHLQ-----	303
CaERF20	F-----WLF-----	206
CaERF21	-----	
At5g61890	QQE-EQDYARFWRFGDSSSSPHSGY-----	248
At5g07310	QQDDEQDYARYLRFGDSSP-PNSGF-----	263
CaERF36	EYPHILQYAE LLRSGDQNTCMMR-----	173
At5g13330	QQQLQQQQQQREEEKKNYGYNYNYNPRE-----	212
At1g43160	EEENHFFKWPTEYDQYNYSSFG-----	192
At5g50080	-----SSSNFPAPSLPNT-----	220
CaERF37	QTQGSQDQGGSSSFPAPSWTDSGHYPSSS-----	256
At5g64750	ESASETGYLQDIQWPSDKTSNNYNNSSPS-----	391
At2g33710	SSLNHQGLRPNLEDGENVKNISIHKRRK-----	218
At2g31230	FEDLGAEYLEQLLMSSC-----	243
At1g06160	LEDLGAEYLEELMRSCS-----	244
At3g23240	FEDLGEQYLEELLGSSSENSGTW-----	218
At5g43410	LLEEGEKPNK-----GKKK-----	131
At1g04370	LLEYGENYNKTHNINMGKRQ-----	133
At3g23220	LLDYGERSNQ-DNCNDANR-----	139
At3g23230	LLDSEERKR-----	139
At4g18450	FEDLGSDYLETLLSSF-----	303
At4g17500	VVEVARG-----DRLLVL-----	268
At5g47220	VGDETRV-----DELLVS-----	243
CaERF27	RVESHM-----	244

CaERF28	-----	
CaERF29	EVESQSNAYRVNCQKTHVPVGEQLLVS-----	264
At2g44840	LMDQSECSYS-----DNRI-----	226
CaERF30	ESEEYAATFSVF-----	193
CaERF31	ESEEYAATFSVF-----	193
At5g47230	HPPFGYSQLTVV-----	300
At4g17490	HPPFGYPQLTVV-----	282
At5g61600	T-SPNFSVISVT-----	241
At5g51190	CPSLGHSQLVVT-----	221
CaERF32	HPSLGYSQLMVI-----	317
CaERF34	-----	
CaERF33	-----	
At5g07580	-----	
At5g61590	-----	
At1g80580	IDNLEKMGLELDLSLGLLSRK-----	256
CaERF35	-----	
At2g20350	-----	
At5g67000	-----	
At5g67010	-----	
At1g25470	LPDCDFTEME-----FQLDDFKFAYTDHLTTP-----PLGLV-----	287
At1g68550	LPDFDFADVED-----LQLADSSFGFLDQLAPINISCPLKSFAAS-----	324
At3g25890	LPDYDFSDVEIDLGLIGTTIDKYAFVDHIATTTTP-TPLNIACP-----	332
At4g11140	-----	
At4g23750	-----	
CaERF13	-----	
At5g53290	-----	
At4g27950	-----	
At2g46310	-----	
AT2G20880	-----	
AT4G28140	-----	
AT1g22190	-----	
AT1g78080	-----	
AT1G36060	-----	
CaERF01	-----	
CaERF02	-----	
CaERF05	-----	
AT2g22200	-----	
AT4g39780	-----	

CaERF04	-----	
CaERF03	-----	
AT5g65130	-----	
At4g31060	-----	
At2g23340	-----	
At4g36900	-----	
At5g67190	-----	
At3g50260	-----	
CaERF06	-----	
At4g06746	-----	
At1g46768	-----	
At1g21910	-----	EDADI 224
At1g77640	-----	DDVDI 238
At1g44830	-----	EEGDI 205
At1g22810	-----	GHDHV 144
At1g71520	-----	EDDLV 143
At1g19210	-----	GEETYADGSM 180
At1g74930	-----	GGESLLDESL 190
At5g21960	-----	ACEDDFSHYNI 211
At2g40220	-----	VCSMGLDPGYMVG DGSS 298
At1g71450	-----	MNAWETYQSD 178
At1g33760	-----	DEFFEMQQFQ 177
At1g63040	-----	DDSPENS DGE 243
At1g12630	-----	DDSPENS NDE 187
At5g52020	-----	DDSPENHEGD 226
CaERF08	-----	ESPENS DAD 241
CaERF07	-----	
At2g36450	-----	DADGFPGGDD 178
At4g25470	-----	FDVEGDDD-V 211
At4g25480	-----	HEVDGDDDDV 211
At4g25490	-----	YDGE GDG-D-V 208
At5g51990	-----	DFDGVGD-V 217
At1g63030	-----	VDEEMCYEDL 176
At1g12610	-----	TEEEMCYEDM 202
At3g60490	-----	VG--FRFEED 250
CaERF09	-----	AEPVFRLEED 305
At2g35700	-----	FPYEED 188
At4g16750	-----	FPYEED 173
At2g44940	-----	FRLEED 290
At5g25810	-----	PASLSQDWDL 212

At5g11590	-----GHGISPNFGH	230
At2g25820	-----GD-DSVLLDP	214
At4g32800	-----NNNDSSSLNYS	211
At1g77200	-----DENGISMAEP	233
At3g16280	-----ISEFNF	228
At3g11020	-----EPGPHQSQDQ-----NHVNSGSYDLHPLHLEPHDG-	319
At5g05410	-----AGLNQDRYPGNSVANGSYRPESQQSGFDPLQSLNYGIPPFQLEGKDG-	322
CaERF10	-----RAERWAHGSXRYEYSSRGAC	235
At2g40350	-----	
At2g40340	-----SLNPMEIAQPGVDYGCPYVQPSDMENYGID	315
CaERF11	-----LVSAFLCEVSRAEIVSQLHD	229
At2g38340	-----LDNGLLYNEPQSSSYHQGG	233
At1g75490	-----DNDISRFDTCISGGYSNWD	191
At5g18450	-----MLVPSCGGERMHRPELEERTGY	270
At3g57600	DENEAEPSVAECHSPPPWNEQEETGSPFRTENFSWDTLIEMPRSETTTMQFDSSNFGSY	258
CaERF12	-----TTTAGVTSVPNDNN	188
At5g11190	-----	
At5g25390	-----	
At1g15360	-----	
At5g25190	-----	
At5g19790	QLFCPT-----	253
At3g15210	-----	
At1g53170	-----	
At1g28370	-----	
CaERF22	-----	
CaERF23	-----	
CaERF25	-----	
At1g03800	-----	
At5g44210	-----	
At1g50640	-----	
At3g20310	-----	
CaERF24	-----	
CaERF26	-----	
At1g28360	-----	
At1g12980	AVADGGLSLGSNEWGYQEMLMYGTQLGCTCRRSWG-----	328
At5g13910	-----	
At5g18560	ITNYGAAAASMSMEDYGMMEPQDLSSSSIAAFGDVVADTTGFYSVF-----	348
At1g28160	-----	
At1g12890	LKDVNDDAPQFLM-----	219
At1g24590	KMQQNGNGGMFYQYCFHDDY-----	306

At2g47520	-----	
At1g72360	-----	
CaERF14	-----	
CaERF15	-----	
CaERF16	-----	
At3g16770	-----	
CaERF17	-----	
CaERF18	-----	
At1g53910	DASLD---FLNEDAVTTQDNGANPMDLWSIDEIHSMIGGVF-----	358
At3g14230	DTPLEVEAMLGADAGAVTQEEENPVELWSLDEINFMLEGDF-----	379
CaERF19	-----	
CaERF20	-----	
CaERF21	-----	
At5g61890	-----	
At5g07310	-----	
CaERF36	-----	
At5g13330	-----	
At1g43160	-----	
At5g50080	-----	
CaERF37	-----	
At5g64750	-----	
At2g33710	-----	
At2g31230	-----	
At1g06160	-----	
At3g23240	-----	
At5g43410	-----	
At1g04370	-----	
At3g23220	-----	
At3g23230	-----	
At4g18450	-----	
At4g17500	-----	
At5g47220	-----	
CaERF27	-----	
CaERF28	-----	
CaERF29	-----	
At2g44840	-----	
CaERF30	-----	
CaERF31	-----	
At5g47230	-----	
At4g17490	-----	

At5g61600 -----
At5g51190 -----
CaERF32 -----
CaERF34 -----
CaERF33 -----
At5g07580 -----
At5g61590 -----
At1g80580 -----
CaERF35 -----
At2g20350 -----
At5g67000 -----
At5g67010 -----
At1g25470 -----
At1g68550 -----
At3g25890 -----

At4g11140 -----
At4g23750 -----
CaERF13 -----
At5g53290 -----
At4g27950 -----
At2g46310 -----
AT2G20880 -----
AT4G28140 -----
AT1g22190 -----
AT1g78080 -----
AT1G36060 -----
CaERF01 -----
CaERF02 -----
CaERF05 -----
AT2g22200 -----
AT4g39780 -----
CaERF04 -----
CaERF03 -----
AT5g65130 -----
At4g31060 -----
At2g23340 -----
At4g36900 -----
At5g67190 -----

At3g50260	-----	
CaERF06	-----	
At4g06746	-----	
At1g46768	-----	
At1g21910	PLWSFN-----	230
At1g77640	PLWSFS-----	244
At1g44830	RLWNFC-----	211
At1g22810	-----	
At1g71520	-----	
At1g19210	FLWNF-----	185
At1g74930	ILWDF-----	195
At5g21960	NLWNF-----	216
At2g40220	TIWPFGEYSHNWGSIWDFIDPILGEFY-----	328
At1g71450	FLWDP-----	183
At1g33760	SIWNSNN-----	184
At1g63040	SLWSY-----	248
At1g12630	DLWGY-----	192
At5g52020	NLWSYK-----	232
CaERF08	NLWSYP-----	247
CaERF07	-----	
At2g36450	YLWNFP-----	184
At4g25470	SLWSY-----	216
At4g25480	SLWSY-----	216
At4g25490	SLWSY-----	213
At5g51990	SLWSFDE-----	224
At1g63030	SLWSY-----	181
At1g12610	SLWSYRY-----	209
At3g60490	FNWQND-----	256
CaERF09	FLWEWY-----	311
At2g35700	FLSQSY-----	194
At4g16750	FFLENY-----	179
At2g44940	FFLSE-----	295
At5g25810	QGLYNY-----	218
At5g11590	GLSWDL-----	236
At2g25820	CLKESFLWNYE-----	225
At4g32800	SVFESWKHFP-----	221
At1g77200	FDSSNFGPLFP-----	244
At3g16280	VELWSFNH-----	236
At3g11020	HEFNGLSSLDI-----	330
At5g05410	GFFDDLSYLDLEN-----	335

CaERF10	-----	
At2g40350	-----	
At2g40340	LDHRRFNDLDIQDLDFGGDKDVHGST-----	341
CaERF11	ELSHG-----	234
At2g38340	GFDSYFEYFRF-----	244
At1g75490	SFHSPL-----	197
At5g18450	LEMDDLLEIDDLGLLIGKNGDFKNWCCEEFQHPWNWF	307
At3g57600	DFEDDVSFPSIWDYYGSLD-----	277
CaERF12	NFVDDNGNEAV-----	199
At5g11190	-----	
At5g25390	-----	
At1g15360	-----	
At5g25190	-----	
At5g19790	-----	
At3g15210	-----	
At1g53170	-----	
At1g28370	-----	
CaERF22	-----	
CaERF23	-----	
CaERF25	-----	
At1g03800	-----	
At5g44210	-----	
At1g50640	-----	
At3g20310	-----	
CaERF24	-----	
CaERF26	-----	
At1g28360	-----	
At1g12980	-----	
At5g13910	-----	
At5g18560	-----	
At1g28160	-----	
At1g12890	-----	
At1g24590	-----	
At2g47520	-----	
At1g72360	-----	
CaERF14	-----	
CaERF15	-----	
CaERF16	-----	
At3g16770	-----	
CaERF17	-----	

CaERF18	-----
At1g53910	-----
At3g14230	-----
CaERF19	-----
CaERF20	-----
CaERF21	-----
At5g61890	-----
At5g07310	-----
CaERF36	-----
At5g13330	-----
At1g43160	-----
At5g50080	-----
CaERF37	-----
At5g64750	-----
At2g33710	-----
At2g31230	-----
At1g06160	-----
At3g23240	-----
At5g43410	-----
At1g04370	-----
At3g23220	-----
At3g23230	-----
At4g18450	-----
At4g17500	-----
At5g47220	-----
CaERF27	-----
CaERF28	-----
CaERF29	-----
At2g44840	-----
CaERF30	-----
CaERF31	-----
At5g47230	-----
At4g17490	-----
At5g61600	-----
At5g51190	-----
CaERF32	-----
CaERF34	-----
CaERF33	-----
At5g07580	-----
At5g61590	-----

At1g80580	-----
CaERF35	-----
At2g20350	-----
At5g67000	-----
At5g67010	-----
At1g25470	-----
At1g68550	-----
At3g25890	-----

APÊNDICE B – Motivos conservados da família ERF de *Coffea arabica* e *Arabidopsis thaliana*

Grupo I

AP2

Sequence name	Start	P-value	Site			
CaERF-01	94	8.22e-73	SASPPKPTKL	YRGVQRHWGKWVAEIRLPKNRTRLWLGTFFDTAEAAALAYDKAAFKLRGDFTRLNFP	HLRHNGSLIG	
CaERF-02	46	8.22e-73	SASPPKPTKL	YRGVQRHWGKWVAEIRLPKNRTRLWLGTFFDTAEAAALAYDKAAFKLRGDFTRLNFP	HLRHNGSLIG	
CaERF-03	12	2.20e-73	VATPPKPAKL	YRGVQRHWGKWVAEIRLPKNRTRLWLGTFFDTAEAAALAYDKAAFKLRGDFARLNFP	NLRHQLSQDF	
CaERF-04	168	2.20e-73	VGTPPKPAKL	YRGVQRHWGKWVAEIRLPKNRTRLWLGTFFDTAEAAALAYDKAAFKLRGDFARLNFP	NLRHQLSQDF	
CaERF-05	163	7.20e-64	SASPPKPTKL	YRGVQRHWGKWVAEIRLPKNRTRLWLGTFFDTAEAAALAYDKAAFKLRSDFTRAELP	SSSP	
AT4g39780	93	4.07e-72	PMKNMTAQKL	YRGVQRHWGKWVAEIRLPKNRTRLWLGTFFDTAEAAAMAYDLAAYKLRGDFARLNFP	QFRHEDGYYG	
AT1G36060	143	7.76e-71	SGSVSKPAKL	YRGVQRHWGKWVAEIRLPKNRTRLWLGTFFDTAEAAALAYDRAAFKLRGDSARLNFP	ALRYQTGSSP	
AT2G20880	187	2.98e-68	PVQPFSAATKL	YRGVQRHWGKWVAEIRLPKNRTRLWLGTFFDTAEAAAMAYDREAFKLRGETARLNFP	ELFLNKQEP	
AT4g13620	232	1.96e-65	VVQTTTPTKL	YRGVQRHWGKWVAEIRLPKNRTRVWLGTFFDTAEAAAMAYDTAAYILRGEFAHLNFP	DLKHQLKSGS	
AT1g64380	136	2.62e-65	AGRVTKKKKL	YRGVQRHWGKWVAEIRLPQNRMRVWLGTFFDTAEAAAYAYDRAAYKLRGEYARLNFP	NLKDPSSELLG	
AT2g22200	71	1.47e-64	TPSTSKTKNL	YRGVQRHWGKWVAEIRLPKNRTRLWLGTFFDTAEAAALAYDQAAFQLRGDIAKLNFP	NLIHEDMNPL	
AT5g65130	111	7.90e-64	IDVATKPVKL	YRGVQRQWGKWVAEIRLPKNRTRLWLGTFFDTAEAAALAYDQAAHKIRGDNARLNFP	DIVRQGHYKQ	
AT4G28140	143	7.74e-61	AVQPYIATKL	YRGVQRQWGKWVAEIRKPRSRARLWLGTFFDTAEAAAMAYDRQAFKLRGHSATLNFP	EHFVNKESEL	
AT1g78080	152	5.39e-74	SGVPSKPTKL	YRGVQRHWGKWVAEIRLPKNRTRLWLGTFFDTAEAAALAYDKAAFKLRGDFARLNFP	NLRHNGSHIG	
AT1g22190	83	5.39e-74	GTSAAKPTKL	YRGVQRHWGKWVAEIRLPKNRTRLWLGTFFDTAEAAALAYDKAAFKLRGDFARLNFP	DLRHNDHEYQP	
		Multilevel consensus sequence		YRGVQRHWGKWVAEIRLPKNRTRLWLGTFFDTAEAAALAYDKAAFKLRGDFARLNFP		R E M R Y E T

CMI-1

Sequence name	Start	P-value	Site			
CaERF-01	250	1.58e-21	DFTEEEGSWE	AYPENLMLEKYPSEIDWAAL		
CaERF-02	202	1.58e-21	DFTEEEGSWE	AYPENLMLEKYPSEIDWAAL		
CaERF-03	184	1.22e-22	TFLDFAEPTF	DESENFLLQKYPSEIDWAAL		
CaERF-04	340	4.07e-22	TFLDFAEPTF	DESENFLLQKYPSEIDWAAL		
AT4g39780	246	3.93e-24	TFLDFSDSGF	DEIGSFGLKFPSEIDWDAL	SKLSES	
AT2g22200	235	1.56e-22	TFLDFSDTEF	EEIGSFGLRKFPSVEIDWDAL	SKLANS	
AT1G36060	292	2.86e-22	VQEMCQEP	PWNENFMLGKCPSEIDWASI	LS	
AT1g78080	312	3.74e-21	LTFADPEEPP	QWNETFSLEKYPSEIDWDSI	LA	
AT5g65130	253	1.43e-19	LDFSSDCVKE	DESFLMGLHKYPSEIDWDAL	EKLF	
AT1g22190	233	1.46e-18	TFGDTEEEIQ	PPWENALEKYPSEIDWDSI	LQCSSLVN	
		Multilevel consensus sequence		DEIENFGLKYPSEIDWDAL		SGS M F Y ASL

CMI-2

Sequence name	Start	P-value	Site			
CaERF-01	162	7.90e-34	LRHNGSLIGG	DFGEYKPLHSSVDAKLQAICQSLAEGKGI	DGKKSkgTRR	
CaERF-02	114	7.90e-34	LRHNGSLIGG	DFGEYKPLHSSVDAKLQAICQSLAEGKGI	DGKKSkgTRR	
CaERF-03	77	3.00e-32	FPNLRHQLSQ	DFGDFKPLHSSVDAKLEAICQSLADSQKQ	GNTGEQHqKK	
CaERF-04	233	3.00e-32	FPNLRHQLSQ	DFGDFKPLHSSVDAKLEAICQSLADSQKQ	GNTGEQHqKK	
AT1g78080	220	1.41e-34	LRHNGSHIGG	DFGEYKPLHSSVDAKLEAICKSMAETQKQ	DKSTKSSkKR	
AT1g22190	143	1.63e-25	FARLNFPDLR	HNDEYQPLQSSVDAKLEAICQSLAETTQK	QVRSTKkSSS	

AT4g39780	160	4.21e-25	QFRHEDGYYG	GGSCFNPLHSSVDAKLQEICQSLRKTEDI	DLPCSETELF
AT2g22200	130	2.96e-23	DIAKLNFPNL	IHEDMNPLPSSVDTKLQAICKSLRKTEEI	CSVSDQTKY
AT1g36060	211	3.60e-19	LRYQTGSSPS	DTGEYGPQAQAAVDAKLEAILAEPKNQPGK	TERTSRKRAK
			Multilevel	DFGEYKPLHSSVDAKLEAICQSLAETQKI	
			consensus	DFN Q Q K MRDGEGQ	
			sequence	KSK K	

CMI-3

Sequence name	Start	P-value	Site		
CaERF-01	17	1.23e-25	YSSQSTYGLV	DQGGSLGLNQLTPSQIHDIQSRFSLQN	ENRHHMISGP
CaERF-04	91	8.97e-22	SFDEMGFDEQ	SGGGSIGLNHLTPYQILQVEAQIKLQN	QQQYLAAS
CaERF-05	86	1.23e-25	YSSQSTYGLV	DQGGSLGLNQLTPSQIHDIQSRFSLQN	ENRHHMISGP
AT2g22200	18	2.79e-32	FPVPTSTFGV	NKSMPLGLNQLTPYQIHQIQNLNHR	STISNLSPNR
AT4g39780	44	1.06e-29	PTYNTVFGF	NQTTSLGLNQLTPYQIHQIQNLNQR	NIISPNLAPK
AT5g65130	53	3.84e-26	PNNQVLPVS	NQTGPIGLNQLTPTQILQIQTELHLRQ	NQSRRRAGSH
AT1g78080	80	4.65e-20	QPFYSGSDLQ	QTGSLIGLNNLSSSQIHQIQSIIHPL	PPTHNNNNS
			Multilevel	NQGGSLGLNQLTPYQIHQIQNLNLRN	
			consensus	DKQMPI S LD SRFHHQR	
			sequence	S IS	
				T	

CMI-4

Sequence name	Start	P-value	Site		
CaERF-04	1	1.91e-13	.	MKALEPFMKGASSL	PSSPTTTSSP
CaERF-05	1	1.78e-14	.	MEALEPFMKGASS	PSYSSSYNPN
AT1g36060	14	7.51e-16	LFGGGHGGEL	MEALQPFYKSASTS	ASNPAFASSN
AT5g65130	13	2.47e-15	LNMNAYVDEF	MEALEPFMKVTSSS	STSNSSNPKP
AT1g78080	22	1.58e-13	RSFQDSGGEL	MDALVPFIKSVSDS	PSSSSAASAS
AT1g22190	24	1.42e-12	QSDPFGGEL	MEALLPFIKSPSND	SSAFASLPA
AT4g39780	19	3.10e-10	SNTDPFQEEL	MKALQPYTTNTDSS	SPTYSNVTFG
			Multilevel	MEALEPFMKSASSS	
			consensus	K Q I T T	
			sequence	Y	

CMI-5 não contém

CMI-6 não contém

CMI-7

Sequence name	Start	P-value	Site		
CaERF-01	206	1.69e-24	KGTRRKAEEE	EKAAAKTVDAGSESETGSGHLTDGSSPVSDL	TVPDFTEEEG
CaERF-02	158	1.69e-24	KGTRRKAEEE	EKAAAKTVDAGSESETGSGHLTDGSSPVSDL	TVPDFTEEEG
AT2G20880	291	1.61e-39	LAVGGVSEET	GWAEAWFNAIPEEWGPGSPLWDDYHFPISNH	KDDLDTQNS
AT4G28140	245	2.37e-24	LPVIDVGREE	GMAEAWYNIAITSGWGPESPLWDDLSSHQFS	SESSSSSPLS
			Multilevel	GKAEAWFNAIGSEWGPESPLWDDGHSPISDH	
			consensus	EW A KTVDAPEGSETE GHLT YSFSVQNL	
			sequence	M Y T LD H FS	

GrupoII

AP2

Sequence name	Start	P-value	Site
CaERF-06	20	9.28e-67	GAGKKKSERL YKGIRMRKWKWVAEIREPNKRSRIWLGSYSTPVAARAYDTAVYYLRGPTARLNFP ELLVGDTTAA
At4g36900	30	2.25e-69	AVATRKRDKP YKGIRMRKWKWVAEIREPNKRSRIWLGSYSTPEAAARAYDTAVFYLRGPSARLNFP ELLAGVTVTG
At5g67190	21	6.11e-68	VPGTRKRDRP YKGIRMRKWKWVAEIREPNKRSRIWLGSYSTPEAAARAYDTAVFYLRGPTARLNFP ELLPGEKFS
At2g23340	28	3.69e-67	IGTRKRDLKP YKGIRMRKWKWVAEIREPNKRSRIWLGSYATPEAAARAYDTAVFYLRGPSARLNFP ELLAGLTVSN
At4g06746	36	4.68e-65	RNLISSNEKR YKGIRMRKWKWVAEIREPNKRSRIWLGSYKTAVAAARAYDTAVFYLRGPSARLNFP EEVFKDGNNG
At3g50260	21	9.81e-65	VAVKMKRERP FKGIRMRKWKWVAEIREPNKRSRIWLGSYSTPEAAARAYDTAVFYLRGPTATLNFP ELLPCTSAED
At1g21910	50	1.57e-60	NKNKSKIKK YKGVMRKSWGSWVSEIRAPNQKTRIWLGSYSTAEAAARAYDVALLCLKGPQANLNFP TSSSSHLLD
At1g46768	31	3.91e-60	VPNHVATRP YRGIRRRKWKWVAEIREPNKRSRIWLGSYTTDIAAARAYDVAVFYLRGPSARLNFP DLLLQEEEDHL
At1g77640	43	7.66e-60	HQSCKNKIKK YKGVMRKSWGSWVTEIRAPNQKTRIWLGSYSTAEAAARAYDAALLCLKGPKANLNFP NITTTSPFLM
At1g19210	12	1.19e-59	EGSSSSMQSK YKGVRRKWKWVSEIRLPSRERIWLGSYDTPEKAARAFDAALYCLRGNNAKFNFP DNPPVISGGR
At1g74930	21	1.44e-58	KKRNTAMQSK YKGVRRKWKWVSEIRLPHSRERIWLGSYDTPEKAARAFDAALYCLRGNNAKFNFP NNPPSISVEK
At5g21960	7	1.80e-56	MDASPK YTGVRKRKWKWVAEIRLPSNRDRIWLGSFDSAEEKAARAFDAALYCLRGPGARFNFP DNPPEIPGGR
At1g44830	35	2.11e-52	TSSSSIRMKK YKGVMRKSWGSWVSEIRAPNQKTRIWLGSYSTAEAAARAYDAALLCLKGSSANLNFP PEISTSLYHI
At4g31060	27	3.79e-51	GAHEDRKFCK YRGVRRKSWGWVSEIRVPKTRRIWLGSYDAPEKAARAYDAALFCIRGEKGVYVNF TDKKPQLPEG
At1g22810	14	5.44e-49	RESTGESQSK YGIRRRKWKWVSEIRVPGTRDRLWLGSFSTAEGAAVAHDVAFVCLHQPDSESLN FPHLLNPSLV
At1g71520	14	3.27e-48	RDTGETDQSK YGIRRRKWKWVSEIRVPGTRQRLWLGSFSTAEGAAVAHDVAFVCLHRPSSLDDES FNFPHLLTTS
		Multilevel consensus sequence	YKGIRMRKWKWVAEIREPNKRSRIWLGSYSTPEAAARAYDAAVFCLRGPSARLNFP V K S S L D A K T L Y Y N V

CMII-1

Sequence name	Start	P-value	Site
CaERF-06	88	1.59e-21	LLVGDTTAAN GGDLSAAAIRKKAIEVGSRVDA
At4g36900	105	7.01e-27	TGGGGGVNG GGDMSAAYIRKKAIEVGAQVDA LEAAGAGGNR
At2g23340	99	6.17e-25	GLTVSNGGGR GGDLSAAYIRKKAIEVGAQVDA LGATVVVNTG
At4g06746	106	1.97e-23	FKDNGGEGE GGDMSPTLIRKKAIEVGARVDA ELRLNRMVE
At5g67190	87	2.88e-23	PELLPGEKFS DEDMSAATIRKKAIEVGAQVDA LGTAVQNNRH
At3g50260	85	3.18e-23	NFPPELLPCTS AEDMSAATIRKKAIEVGAQVDA IGATVVQNNK
At1g46768	102	7.92e-21	QEEDHLSAAT TADMPAALIREKKAIEVGARVDA LLASAAPSMA
		Multilevel consensus sequence	GGDMSAAYIRKKAIEVGAQVDA E L L R T R T

CMII-2 não contém

CMII-3 não contém

GrupoIII

AP2

Sequence name	Start	P-value	Site
CaERF-07	82	5.32e-71	PHFRGVRERS GKWVSEIREPRKTTRIWLGTYPTPEMAAAAYDAATLALKGPDAILNFPPELAPTYPIPASLSAADVRAAAAASAAAAR EPPQEPAVL
CaERF-08	82	5.32e-71	PHFRGVRERS GKWVSEIREPRKTTRIWLGTYPTPEMAAAAYDAATLALKGPDAILNFPPELAPTYPIPASLSAADVRAAAAASAAAAR EPLQEPPVL
CaERF-09	133	2.43e-80	TYRVRKRNRW GKWVSEIREPRKRSRIWLGSYPTAEMAARAHVAALAIKGHSAAYLNFPHLAHELPRPATTSPKDIQAAAASAAAAT FPGENGGGAE

At4g25480	60	1.17e-78	IYRGVRRNS	GKWVCEVREPNNKTRIWLGTFTQTAEMAARAHDAALALRGRSACLNFDASAWRLRIP	ESTCAKDIQKAAAEALAF	QDEMCDATTD
At2g25820	35	1.37e-78	VYRGARMSW	GKWVSEIREPRKKSRIWLGTFTAEMAARAHDAALSIKSSAILNFPPELADFLPRPVLSLQDDIQAAAAEALMD	FKTVPFHLQD	
At4g25490	57	3.98e-78	IYRGVQRNS	GKWVSEVREPNNKTRIWLGTFTAEMAARAHDAALALRGRSACLNFDASAWRLRIP	ESTCAKDIQKAAAEALAF	QDETCDTTTT
At2g35700	54	1.13e-77	NFRGVRMRQW	GKWVSEIREPKKSRIWLGTFTAEMAARAHDAALAIKGGSAHLNFPPELAYHLPRPASADPKDIQAAAAAAAV	AIDMDVETSS	
At5g11590	60	1.99e-76	VYRGVMRNW	GKWVSEIREPRKKSRIWLGTFTPEMAARAHDAALSIKGTAAILNFPPELADSFPRPVLSLSPRDIQTAALKAHME	PTTSFSSSTS	
At4g16750	49	2.63e-76	NFRGVRMRQW	GKWVSEIREPRKKSRIWLGTFTPEMAARAHDAALAIKGGSAHLNFPPELAYHLPRPASADPKDIQAAAAAAAVD	WKAPESPSST	
At3g60490	80	5.23e-76	SYRGVMRMSW	GKWVSEIREPRKKSRIWLGTFTAEMAARAHDAALAIKGNISGFLNFPPELSGLLPRPVSCSPKDIQAAATKAAEAT	TWHKPVIDKK	
At2g44940	109	8.61e-75	IYRGVMRMSW	GKWVSEIREPRKKSRIWLGTFTAEMAARAHDAALAIKGTAYLNFPKLAGELPRPVTNSPKDIQAAASLAAVNW	QDSVNDVSNS	
At5g51990	63	3.12e-74	IYRGVQRNS	GKWVCEVREPNNKKSRIWLGTFTVEMAARAHDAALALRGRSACLNFDASAWRLRIP	ETTCPKEIQKAAEAAAMAF	QNETTTTEGSK
At5g25810	45	4.02e-74	VYRGVQRNS	GKWVSEIREPRKKSRIWLGTFTPEMAARAHDAALSIKGSAILNFPDLAGSFPRPSSLSPRDIQVAALKAHME	TSQSFSSSSS	
At1g77200	52	4.02e-74	GYRGVMRWTW	GKWVSEIREPRKKSRIWLGTFTPEMAARAHDAALAIKGTSAVLNFPPELATYLPASSPRDVQAAAAVAAAMF	FSPSSSSLVV	
At4g25470	60	4.56e-74	IYRGVQRNS	GKWVCEVREPNNKTRIWLGTFTAEMAARAHDAALALRGRSACLNFDASAWRLRIP	ESTCAKEIQKAAAEALNF	QDEMCHMTTD
At4g32800	28	2.04e-73	VYRGVMRMSW	GKWVSEIREPRKKSRIWLGTFTAEMAMRAHDVAAMS	IKGTSAILNFPPELSKLLPRPVLSLSPRDVRAAATKAALMD	FDTTAFRSDT
At3g16280	70	8.43e-72	VYRGVMRMSW	GKWVSEIRQPRKTRIWLGTFTVADMAARAHDAALAIKGTSAVLNFPPELASLFPASSPHDIQTAAEAAAMV	VEEKLLEKDE	
At1g12630	27	7.47e-71	PVYRGIRCRS	GKWVSEIREPRKTRIWLGTFTPEMAAAAAVDVAAMALKGREAVLNFPVSGVSYVPPESTSAADIRAAAAAAAMK	GCEEGEEEEK	
At5g52020	73	2.66e-68	SIFRGIRLRN	GKWVSEIREPRKTRIWLGTFTPEMAAAAAVDVAALALKGPDAVLNFPGLALTYVAPVNSAADIRAAASRAEMK	QPDQGGDEKV	
At1g01250	54	1.18e-66	VYHGVRKRRW	GKWVSEIREPRKKSRIWLGSFVPEMAAKAYDVAAFCLKGRKAQLNFPPEIEDLPRPSTCTPRDIQVAAAKAANAV	KI IKMGDDDV	
At1g63040	98	7.04e-66	PVYRGIRCRS	GKWVSEIREPKKTRVWLGTFTPEMAAAAAVDVAALALKGGDTLLNFPDLSGYPVPLSSAAHIRCAAAAAAATR	GAAGAANKVG	
At2g36450	25	7.76e-66	YRQVQRKNS	NKWVSEIREPRKPNRIWLGTFTPEMAAIIAYDVAALALKGSQAEELNFPNSVSSLPAPTSMSPADIQAAAAAAAF	GAARDAIVMA	
At1g12610	39	1.15e-65	VYRGIRRRNG	DKWVCEVREPHTQRIWLGTFTPADMAARAHDAVAVLALRGRSACLNFDASAWRLPVPEESNDPDIIRRVAAEAAEMF	RPVDLESGIT	
At1g33760	29	4.47e-65	SYRGIRRRKW	GKWVSEIREPGKTRIWLGSFVPEMAAKAYDVAAFCLKGRKAQLNFPPEIEDLPRPSTCTPRDIQVAAAKAANAV	KPGRLESEPAL	
At1g63030	39	3.31e-64	IYRGVRRDG	DKWVCEVREPIHQRRVWLGTFTPADMAARAHDAVAVLALRGRSACLNFDASAWRLPVPASTDPTIRRATAEAAEMF	RPPEFSTGIT	
At1g71450	33	3.16e-58	AYRGVQRKRW	GKWVSEIREPGTKNRIWLGSFVPEMAAATAYDVAAFHFRGREARLNFPPELASSLPADSSSDSIRMAVHEATLCR	TTEGTESAMQ	

Multilevel consensus sequence
 T Y P A Y IR I DS R I E L A A R K L A
 V

CMIII-1

Sequence name	Start	P-value	Site
CaERF-07	139	6.07e-14	PELAPTYPIP ASLSAADVRAAAAASAAAAR EPPQEPVLD
CaERF-08	139	6.07e-14	PELAPTYPIP ASLSAADVRAAAAASAAAAR EPLQEPVLD
CaERF-09	190	8.82e-18	PHLAHELPRP ATTSKDIQAAAAKAAAT FPGENGGGAE
At4g25480	117	7.79e-20	ADSAWRLRIP ESTCAKDIQKAAAEALAF QDEMCDATTD
At4g25490	114	7.79e-20	ADSAWRLRIP ESTCAKDIQKAAAEALAF QDETCDTTTT
At2g25820	92	3.01e-17	PELADFLPRP VLSLQDDIQAAAAEALMD FKTVPFHLQD
At4g25470	117	5.73e-17	ADSAWRLRIP ESTCAKEIQKAAAEALNF QDEMCHMTTD
At3g60490	137	7.36e-17	PELSGLLPRP VSCSPKDIQAAATKAAEAT TWHKPVIDKK
At3g16280	127	1.20e-16	PELASLFPFP ASSSPHDIQTAAEAAAMV VEEKLLEKDE
At5g11590	117	3.45e-16	PELADSFPRP VLSLPRDIQTAALKAHME PTTSFSSSTS
At2g35700	111	3.45e-16	PELAYHLPRP ASADPKDIQAAAAAAAV AIDMDVETSS
At1g12630	84	4.33e-16	PGSVGYPVP ESTSAADIRAAAAAAAMK GCEEGEEEEK
At2g36450	82	7.53e-16	PNSVSSLPAP TMSMPADIQAAAAASAAA F GAARDAIVMA
At1g77200	109	1.16e-15	PELATYLPFP ASSSPRDVQAAAAVAAAM FSPSSSSLVV
At5g51990	120	1.43e-15	ADSAWRLRIP ETTCPKEIQKAAEAAAMAF QNETTTTEGSK
At4g32800	85	1.96e-15	PELSKLLPRP VLSLPRDVRAAATKAALMD FDTTAFRSDT
At5g25810	102	3.28e-15	PDLAGSFPRP SLSLPRDIQVAALKAHME TSQSFSSSSS
At4g16750	106	2.25e-14	PELAYHLPRP ASADPKDIQAAAAAAAVD WKAPESPSST
At2g44940	166	2.47e-14	PKLAGELPRP VTNSPKDIQAAASLAAVNW QDSVNDVSNS
At1g01250	111	2.96e-14	PEEIEDLPRP STCTPRDIQVAAAKAANAV KI IKMGDDDV

At5g52020	130	4.25e-14	PGLALTYVAP	VNSAADIRAAASRAAEMK	QPDQGGDEKV
At1g33760	86	4.23e-13	PELVDSFPRP	ESSSSEHIQAAAQDAALMF	KPGRLESEPAL
At1g12610	96	4.97e-13	ADSAWRLPVP	ESNDPDVIRRVAAEAAEMF	RPVDLESGIT
At1g63030	96	5.38e-13	SDSAWRLPVP	ASTDPDTIRRTAAEAAEMF	RPPEFSTGIT
At1g63040	155	3.98e-12	PDSLGSYPPI	LSSSAHIRCAAAAAAATR	GAAGAAVKVG
			Multilevel	ASTSPKDIQAAAAEAAAF	
			consensus	E L AA R K LM	
			sequence	V	

CMIII-2

Sequence name	Start	P-value	Site		
CaERF-08	188	1.07e-31	VEINEDINAS	TSGDNQMGQDQFVDEEELFDMPNLLVDM	PAVSGESPEN
At2g36450	123	8.10e-29	NSQTSGVACM	NSSYDNTNMNGFMDEDLVFDMPNVLMNMAEGMLLSPPRPT	VFDAAYDADG
At1g63040	190	1.43e-28	VKVGQKKEDK	VYDTAESSTMVGFVDEEELLNMPGLLADMAKGMVAPPWMG	SPPSDDSPEN
At4g25490	156	1.71e-45	LDMEETMVEA	IYTPEQSEGAFYMDEETMFGMPDLLDNMAEGMLLPPPSVQ	WNHNYDGEDG
At4g25470	159	4.38e-42	LDMEETLVEA	IYTPEQSQDAFYMDEEAMLGMSLLDNMAEGMLLPPPSVQ	WNYNFDVEGD
At4g25480	158	3.03e-37	FDMEETLVEA	IYTAEQSENAFYMHDEAMFEMPSLLANMAEGMLLPLPSVQ	WNHNHEVDGD
At5g51990	166	8.68e-33	AAGEGVREGE	RRAEEQNGGVFYMDDEALLGMPNFFENMAEGMLLPPPEVG	WNHNDFDGVG
At5g52020	171	1.05e-32	PVQPGKEEEL	EEVSCNSCSLEFMDEEAMLNMPDLLTEMAEGMLMSPPRMM	IHPTMEDDSP
At1g12630	133	8.92e-32	SRARECHVDN	DVGSSSWCGTEFMDEEEVLNMPNLLANMAEGMMVAPPSSW	GSRPSDDSP
			Multilevel	IYTPEQSCGAFFMDEEAMLGMPNLLANMAEGMLLPPPSVQ	
			consensus	SGS N ED EYV D ELFN SS DD MVSS RMG	
			sequence	G G V D T A M	
				Q	

CMIII-3 não contém

CMIII-4

Sequence name	Start	P-value	Site		
CaERF-08	232	1.44e-11	SPPRMKPAVS	GESPENSADNLWSY	P
At3g16280	220	1.33e-09	CWVYPPVMDF	YEEISEFNFVELWSF	NH
At2g36450	169	2.61e-09	PRPTVFDAAV	DADGFPGGDDYLWNF	P
At5g25810	176	5.40e-09	GSSYDGLTQL	GNEFIFSDSADLWPY	PPQWSEGDYQ
At1g12610	193	3.65e-16	MEDMTPTNVY	TEEEMCYEDMSLWSY	RY
At1g63030	167	7.71e-15	SYIDMNTSVY	VDEEMCYEDLSLWSY	
At4g25490	199	1.10e-13	LPPPSVQWNH	NYDGEEDGDVSLWSY	
At4g25470	202	1.61e-13	LPSPSVQWNY	NFDVEGDDDVSLWSY	
At5g52020	217	8.58e-13	PRMIIHPTME	DDSPENHEGDNLWSY	K
At5g51990	208	1.87e-12	LLPPPEVGVN	HNDFDGVDVSLWSF	DE
At1g63040	234	2.58e-12	APPWMGSPPS	DDSPENS DGESLWSY	
At4g25480	202	4.35e-12	PLPSVQWNHN	HEVDGDDDDVSLWSY	
At1g12630	178	7.24e-12	PPSWMGSRPS	DDSPENSDEDLWGY	
			Multilevel	DEEPENSDDVSLWSY	
			consensus	DDEMCDE D F	
			sequence	S GY	

CMIII-5 não contém

CMIII-6

Sequence name	Start	P-value	Site
CaERF-09	255	1.62e-11	NSPSTDDID DTFDDLPLSFDVTG ATDHLDFNC
At2g35700	160	8.89e-14	SSPAMIALSD DAFSDLPLLLNVNH NIDGFWSFP
At5g25810	157	1.64e-13	SSSATGSEEL GEIVELPSLGSSYDG LTQLGNEFIF
At4g32800	157	3.32e-13	IEESTVSDDL DEIVKLPSLGTSLNE SNEFVIFDSL
At1g77200	179	4.16e-13	EEAATAEEL SEIVELPSLETSYDE SLSEFVYVDS
At4g16750	143	8.94e-13	STVTSSPVAD DAFSDLPLLLDVND HNKNDGFWS
At2g25820	164	1.23e-12	SSSSMLSGEL GDIVELPSLENNVKY DCALYDSLEG
At2g44940	247	5.97e-12	SACTDKDSEE EKLFDLPLDFTDENE MMIRNDAFCY
At3g60490	206	8.69e-11	TSSTDKESNE ETVFDLPLDFTDGLM NPNDAFCLCN
At3g16280	179	1.16e-09	AAESEDEERL EKIVELPNIEEGSYD ESVTSRADLA
		Multilevel	DEIVDLPLDGTSVNE
		consensus	GAFFE S ELDYD
		sequence	E S L N

CMIII-7

Sequence name	Start	P-value	Site
CaERF-09	294	9.38e-13	SSWWPQLHPA PAAEPVFRLEEPFLW EWY
At2g44940	279	6.17e-19	FCYSSSTWQL CGADAGFRLEEPFFL SE
At4g16750	162	4.83e-14	LLDVNDHNKN DGFWDSFPYEDPFFL ENY
At2g35700	177	9.63e-14	DLLLNVNHNH DGFWDSFPYEEPFLS QSY
At3g60490	239	1.24e-13	CNGFTWQLY GEEDVGFRFEEPFW QND
		Multilevel	CGADAGFRLEEPFFL
		consensus	D FWDS PY LW
		sequence	

Grupo IV

AP2

Sequence name	Start	P-value	Site
CaERF-10	60	2.06e-91	DEEGKVIRKV PAKGSKKCMKGGPENARCNYRGVQRQTWGWVVAE IREPNRGSRLLWLGTFGTAVEAALAYDEAARSMYGPCLRNLP NYGPQQEASI
CaERF-11	59	1.85e-81	SINGKLVIRKA PGRGSKKCMKGGPENFRHNYRGVQRQTWGWVVAE IREPCKGSRLLWLGTFGTAEAAALAYDKAARVMYGPSAQLNLP NYNVVKEAST
CaERF-12	18	8.00e-81	NGNGRPGKKP AQASSRKGCMRGGPENASCAKGVQRQTWGWVVAE IREPNRGSRLLWLGTFDTSHEAAIAYDAAARKLYGTAAKLNLPL HLYRRPQHIP
At5g05410	57	1.26e-86	EEVSTKKRKY PAKGSKKCMKGGPENSRCSFRGVRQRIWGWVVAE IREPNRGSRLLWLGTFPTAQEAASAYDEAAKAMYGPLARLNLP RSDASEVTST
At5g18450	12	6.86e-83	EEEQPPAKKR NMGRSRKCMKGGPENATCTFRGVRQRQTWGWVVAE IREPNRGTRLLWLGTFNTSVEAAMAYDEAAKLYGHEAKLNLV HPQQQQVVV
At2g40340	50	1.01e-82	GGGPKSIRKP PPKGSRKCMKGGPENICDYRGVQRQRWGWVVAE IREPDPGARLLWLGTFSSSYEAALAYDEAAKAIYQSARLNLP EITNRSSTA
At3g11020	56	4.13e-82	KEGEKPKRKY PAKGSKKCMKGGPDNSHCSFRGVRQRIWGWVVAE IREPKIGTRLLWLGTFPTAEKAASAYDEAATAMYGLARLNLP QSVGSEFTST
At1g75490	11	4.36e-81	MVGANKKQRT VQASSRKGCMRGGPDNASCTYKGVQRQTWGWVVAE IREPNRGSRLLWLGTFDTSREAAALAYDSAAKLYGPEAHLNLP ESLRSYPKTA
At2g40350	45	1.38e-79	GGGSKPIRKA PPKRSRKGCMKGGPENICDYTGVRQRQTWGWVVAE IREPGRGAKLWLGTFSSSYEAALAYDEASKAIYQSARLNLP LLPLCQARLL
At3g57600	6	2.57e-73	MEKSS SMKQWKKGPARGKGGPNALCQYRGVQRQTWGWVVAE IREPCKRRLWLGSFATAEAAAMAYDEAAKLYGHDAYLNLP HLQRNTRPSL
At2g40220	33	2.55e-60	NNPQSDSTTD SSTSSAQRKRKGGPDNSKFRYRGVQRQRWGWVVAE IREPCKRTRKWLGTAFATAEDAAARAYDRAAVLYGSRAQLNLT PSSPSSVSSS
At2g38340	48	2.72e-49	EEGLARARV QAKGSKKCMRGGPENPVCFRGVRQRIWGWVVAE IREPVSHRGANSRKRLLWLGTFATAEAAALAYDRAASVMYG PYARLNFPED
		Multilevel	PAKGSKKCMKGGPENARCxYRGVQRQTWGWVVAE IREPNRGARLLWLGTFxTAEAAALAYDEAAKALYGPSARLNLP
		consensus	S R R D S F KK S S RKM
		sequence	T

CMIV-1 não contém

CMIV-2

Sequence name	Start	P-value	Site
CaERF-10	22	7.30e-22	SVPMDYSRKR KSRRRDGTNNVADTLAKWKEYNDKLD S LDEEGKVIRK
CaERF-11	23	7.95e-22	LTLGPFGRKN SKRRSDASKHVAETLNKWKKEYFDKIES INGLVRKAP
At2g40340	8	1.67e-31	MPSEIVD RKRKSRGTRDVAEILRWREYNEQIEA ESCIDGGGPK
At2g40350	3	2.51e-29	MP RKRKSRGTRDVAEILRWREYNEQTEA DSCIDGGGSK
At3g11020	16	2.45e-24	QTGTEQPKKR KSRARAGGLTVADRLKKWKEYNEIVEA SAVKEGKPK
At5g05410	22	1.43e-21	RTQIDTSRKR KSRSRGDGTTVAERLKRWKEYNETVEE VSTKKRKVPA
		Multilevel	KKRKSRTDVAEILRWKEYNEQIEA
		consensus	RS RRD G T DR KQ R DKV S
		sequence	T

CMIV-3 não contém

GrupoVI

AP2

Sequence name	Start	P-value	Site
CaERF-13	119	1.13e-60	PPAEPDVSSR KKFRGVRQRPWGRWAAEIRDPTKRKRVLGTYDTPPEEASVYDRAAVKLKGPDAVTFNPVES IKERSVGRVVR
At4g27950	116	1.58e-73	QKPSVSGQNO KKFRGVRQRPWGWAAEIRDPEQRRRIWLGTFATAEEAAIVYDAAIKLRGPDALTNFTVQP EPEPVQEQEQ
At5g53290	123	4.90e-72	NNKVSQGI KKFRGVRQRPWGWAAEIRDPEQRRRIWLGTFATAEEAAIVYDAAIKLRGPDALTNFSIPP QEEEEEEPE
At4g23750	120	2.16e-70	VSATTTTTGE KKFRGVRQRPWGWAAEIRDPLKRVRLWLGTYNTAAEAMVYDAAIQLRGPDALTNFSVTP TTATEKKAPP
At4g11140	85	1.98e-68	TVPVVVTAT RKFRGVRQRPWGWAAEIRDPSRRVRVWLGTFATAEEAAIVYDAAIQLRGPNALNFPPPP VTENVEEAST
At2g46310	97	3.25e-61	ENGGDLVKS V KYRQVRQRPWGWAAEIRDPSRRTRLWLGTFATAEEAAIGYDRAAIRIKGHNAQTNFLTPP PSPTTEVLPE
At1g22985	76	3.10e-43	RNTTANKKKF RGVQRQRPWGWAAEIRCGRVKGRPERIWLGTFTATAEEAALAYDAAIQLRGPDAPTNFGRPD VDSAVVKKQD
At1g71130	82	2.20e-41	TAKTEKKYRG VRQRPWGWAAEIRCGRGACKRRDRWLWLGTFNTAAEALAYDNASIKLIGPHAPTDFGLPA ENQEDKTVIG
		Multilevel	KKFRGVRQRPWGWAAEIRDPEERRRIWLGTFATAEEAAIVYDAAIKLRGPDALTNFGVPP
		consensus	R Y G KGK V L YD LA R Q I N P P
		sequence	V SQ V E R K S
			N

CMVI-1

Sequence name	Start	P-value	Site
CaERF-13	49	4.27e-20	SGQSLTVVRQ KVVRIIYTDADATDSSDDEEE TRNVRRVKRH
At2g46310	26	5.40e-20	ITSTPFDGFP KIVKIIVTDPCATDSSSDEEN DNKSVAPRVK
At4g27950	41	8.65e-20	ERKTSVRDSA RLVRVSMTRDRATDSSSDEEE FLFPRRRVKR
At1g22985	3	1.38e-19	MK RIVRISFTDMEATDSSSSEDE SPPSSRRRGK
At1g71130	3	2.18e-19	MK RIIRISFTDAEATDSSSDEDT EERGASQTR
At5g53290	37	3.19e-19	SEKLSGGKSL KKVSICTDPDATDSSSDEDE EDFLFPRRRV
At4g11140	11	1.48e-23	ME'EKVSLP RILRISVTDPYATDSSSDEEE EVDFDALSTK
At4g23750	35	2.24e-21	VSELTNTHQT RILRISVTDPPDATDSSSDEEE EEHQRFVSKR
		Multilevel	RIVRISVTDPPDATDSSSDEEE
		consensus	K L IF AE D
		sequence	Y Y

CMVI-2

Sequence name	Start	P-value	Site
CaERF-13	73	6.03e-13	SSDDEEETRN VRRVKRHVEEINFGP PKTTKIEQP

At2g46310	52	1.86e-11	SDEENDNKS	APRVKRYVDEIRFC	EDDEPKPARK
At5g53290	64	2.35e-16	DEDEEDFLFP	RRRVKRFVNEITVEP	SCNNVVTGVS
At1g71130	33	8.97e-16	TEERGGASQT	RRRGKRLVKEIVIDP	SDSADKLDVC
At4g27950	66	4.71e-15	SSDEEEFLFP	RRRVKRLINEIRVEP	SSSSTGDVSA
At1g22985	29	9.92e-15	SSEDESPPSS	RRRGKLVKEIVIDH	SDPPEVGKTR
At4g23750	65	2.61e-14	EEHQRFVSK	RRRVKRFVNEVYLD	GAVVTGSCGQ
At4g11140	42	3.10e-14	EVDFDALSTK	RRRVKRYVKEVYLD	VVSDKEKPMK

Multilevel
consensus
sequence

RRRVKRFVNEIVVDP
G KL K VRFES
Y TI
L

CMVI-3 não contém

CMVI-4 não contém

CMVI-5 não contém

Grupo VII

AP2

Sequence name	Start	P-value	Site		
CaERF-14	62	3.25e-87	NKHPKKGAGN	QKSEKISQKPRKNKYR	GIRQRPWGKWAEEIRD
CaERF-15	62	1.09e-86	NKHPKKGAGN	QKAEKISQKPRKNKYR	GIRQRPWGKWAEEIRD
CaERF-16	62	3.25e-87	NKHPKKGAGN	QKSEKISQKPRKNKYR	GIRQRPWGKWAEEIRD
CaERF-17	73	4.17e-84	DPHDEAGGKQ	QVLPKGAKRQRKNLYR	GIRQRPWGKWAEEIRD
CaERF-18	10	8.62e-83	MCKQTGGKQ	QVLPKGAKRQRKNLYR	GIRQRPWGKWAEEIRD
CaERF-19	111	8.27e-86	TKCGNSSDSS	DESGKYLKRKRKNHYR	GIRQRPWGKWAEEIRD
CaERF-20	35	3.51e-86	TKCGNSSDSS	DESGKYLKRKRKNHYR	GIRQRPWGKWAEEIRD
CaERF-21	118	7.93e-88	TKYKSSDSS	DESEKYSKRKRKNQYR	GIRQRPWGKWAEEIRD
At3g16770	65	2.10e-77	NVKEEAVKKE	QATEPGKRRKRKNVYR	GIRKRPWGKWAEEIRD
At3g14230	114	2.19e-76	AYAKKTVESA	EQAEKSKRKRKNQYR	GIRQRPWGKWAEEIRD
At1g53910	111	1.11e-73	VFGKVTGLD	GDAEKSANRKRKNQYR	GIRQRPWGKWAEEIRD
At2g47520	36	2.78e-71	SSRKKRKPVS	VSEERDGRERKNLYR	GIRQRPWGKWAEEIRD
At1g72360	62	3.65e-65	VFSSTHKHHA	SGSASDGKQSSRYKGI	RRRPWGRWAAEIRD

Multilevel
consensus
sequence

QESEKYSKRKRKNKYR
DKA IAQKP L Q TE KK D V VNDTS PAPT
Y Q A T

CMVII-1

Sequence name	Start	P-value	Site		
CaERF-14	1	3.01e-20	. MCGGAIISGGSPINPGRKLT	PKELWAEFDT	
CaERF-15	1	3.01e-20	. MCGGAIISGGSPINPGRKLT	PKELWAEFDT	
CaERF-16	1	3.01e-20	. MCGGAIISGGSPINPGRKLT	PKELWAEFDT	
CaERF-17	1	2.43e-16	. MCGGAIILEGLIPPRNNDGT	RRLLSADLWP	
CaERF-19	1	8.67e-20	. MCGGAIISDYIQPSRSSHRL	SADLLWGIGG	
CaERF-21	1	3.88e-21	. MCGGAIISDFIPPSRSSRRL	TADLLWGTGA	
At3g16770	1	7.94e-17	. MCGGAIISDYAPLVTKAKGR	KLTAEEELWSE	
At2g47520	1	4.73e-16	. MCGGAIISDFIWSKSESEPS	QLGSVSSRKK	

At1g53910	1	2.84e-25	. MCGGAIISDFIPPPRRRVT	SEFIWPDLLK
At3g14230	1	1.39e-24	. MCGGAIISDFIPPPRLRVT	NEFIWPDLLK
At1g72360	1	5.59e-17	. MCGGAVISDYIAPEKIARSS	GKSSWRSNGV
		Multilevel	MCGGAIISDFIPPNRSRLT	
		consensus	GGG IPPGSKV	
		sequence	Y	

CMVII-2

Sequence name	Start	P-value	Site	
CaERF-19	81	8.32e-17	EFSDQAVEEL	EDVKPFGFSASKHSA ILSRGTKCGN
CaERF-20	5	8.32e-17	MEEL	EDVKPFGFSASKHSA ILSRGTKCGN
CaERF-22	88	8.32e-17	DHSDDEAEEL	EDVKPFGFSASKHSA IPPRGTKYK
At1g72360	46	4.36e-16	YDFDGNFDEL	ESDEPFVFSSTHKHH ASGSASDGKK
At1g53910	78	3.63e-14	DDDFDVGDF	ADVKPFVFTSTPKPA VSAAAEGSVF
At3g14230	75	8.20e-14	FDCEDDDVF	VNVKPFVFTATTKPV ASAFVSTGIY
		Multilevel	EDVKPFVFSATKKSA	
		consensus	SDE G TSSHHHH	
		sequence	P	

CMVII-3 não contém

CMVII-4 não contém

CMVII-5

Sequence name	Start	P-value	Site	
CaERF-14	239	3.18e-10	SQFGGITESD	PVDLWMMD DLVATAQQPN
At3g14230	362	2.58e-09	AGAVTQEEEN	PVELWSLD EINFMLEGDF
At1g53910	341	1.00e-08	AVTTQDNGAN	PMDLWSID EIHSMIGGVF
At2g47520	161	2.63e-07	PVADDQSATD	IGNLWSYQ DSN
At3g16770	234	7.39e-07	QPSQLDESVS	EVDMMWLD DVIASYE
		Multilevel	PVELWSLD	
		consensus	IGDM MMQ	
		sequence	EV I	
			Y	

CMVII-6

Sequence name	Start	P-value	Site	
CaERF-19	50	5.59e-26	IKKKKSTSNF	YSKQLRSEIIDLDDEFEADFREFSDQ AVEELEDVKP
CaERF-21	54	7.14e-28	KNNKNTSDL	YSKPLRSEVVDLDEFETDFQEFNDH SDDEAELED
At3g14230	37	5.03e-25	DLKKNVKASK	KRSNKRSDFFDLDDDEFADFGQFKDD SAFDCEDDDD
At1g53910	37	1.01e-24	DLKKNLKGSK	KSSKNRSNFFDFDAEFEADFGQFKDD SSIDCDDDFD
		Multilevel	KSKKLRSEFFDLDDDEFADFGQFKDD	
		consensus	YRSNK DII F AD T RG N H	
		sequence	PN NVV S Q	
			Q	

CMVII-7

Sequence name	Start	P-value	Site	
---------------	-------	---------	------	--

CaERF-19	268	1.08e-29	KSYFQSDCAN	LCFNSDQGSSSFGCSDFGWEEQCPKTPEISSI	FHLQ
At3g14230	269	3.61e-41	SFDAGGNNGY	QYFSSDQGSNSFDCSEFGWSDHGPKTPEISSM	LVNNNEASFV
At1g53910	255	1.90e-34	NSVDAGCNGY	QYFSSDQGSNSFDCSEFGWSDQAPITPDISSA	VINNNSALF
			Multilevel	QYFSSDQGSNSFDCSEFGWSDHGPKTPEISSM	
			consensus	LC N S G D EEQA I D A	
			sequence		C I

CMVII-8 não contém

Grupo VIII

AP2

Sequence name	Start	P-value	Site		
CaERF-22	20	2.88e-72	VKPSTGKANA	VKEVHFRGVRKRPWGRYAAEIRDPGKKS	SRVWLGTFDTAEEAARAYDAAARDFRGPKAKTNF
CaERF-23	20	2.88e-72	VKPSTGKANA	VKEVHFRGVRKRPWGRYAAEIRDPGKKS	SRVWLGTFDTAEEAARAYDAAARDFRGPKAKTNF
CaERF-24	28	7.96e-67	AAGETNGSGG	SKEIRFRGVRKRPWGRFAAEIRDPWKK	TRVWLGTFDSDAEDAARAYDAAARTLRGPKAKTNF
CaERF-25	20	9.75e-71	VKPSTGKANA	VKEVHFRGVRKRPWGRYAAEIRDPGKKS	SRVWLGTFDTAEEAARAYDAAAREFRGPKAKTNF
CaERF-26	31	5.37e-69	VQVTPSSLRV	VKESHFRGVRKRPWGRYAAEIRDPWKK	TRRWLGTFDTAEEAALAYDEAARS
At3g15210	20	6.50e-71	PATTNQTHNN	AKEIRYRGVRKRPWGRYAAEIRDPGKKS	TRVWLGTFDTAEEAARAYDAAARDFRGAKAKTNF
At5g44210	28	2.30e-69	GKGTMTMTM	RKEVHFRGVRKRPWGRYAAEIRDPGKKS	TRVWLGTFDTAEEAARAYDAAAREFRGSKAKTNF
At1g50640	23	7.86e-68	PTVVAANGS	VKEIRFRGVRKRPWGRFAAEIRDPWKK	KARVWLGTFDSDAEEAARAYDAAARNLRGPKAKTNF
At1g53170	26	7.96e-67	TNPTHHESNA	AKEIRYRGVRKRPWGRYAAEIRDPVKK	TRVWLGTFDTAQQAARAYDAAARDFRGVKAKTNF
At1g28360	7	9.15e-67	MASTTC	AREVHYRGVRKRPWGRYAAEIRDPWKK	TRVWLGTFDTPEEAALAYDGAARFLRGIKAKTNF
At1g24590	53	4.97e-64	KAASGAGASG	VSTIRYRGVRRRPWGRYAAEIRDPLSK	ERRWLGTFDTAEEAACAYDCAARAMRGLKARTNF
At3g20310	22	1.63e-62	PALPVTAGGS	VKEPRYRGVRKRPWGRFAAEIRDPVKK	SRVWLGTFDSDAVDAARAYDAAARNLRGPKAKTNF
At1g28370	15	6.96e-62	VKTAAVKTNE	GNGVRYRGVRKRPWGRYAAEIRDPFKK	SRVWLGTFDTPEEAARAYDKRAIEFRGAKAKTNF
At1g12980	52	2.87e-61	DTTVTIAGAG	SSTTRYRGVRRRPWGRYAAEIRDPMSK	ERRWLGTFDTAEQAACAYDAAARAFRGAKARTNF
At5g13910	15	5.48e-61	SSKSKKKQDD	QVGTFLGVRRRPWGRYAAEIRDPTTK	ERHWLGTFDTAEEAALAYDRAARS
At1g03800	48	2.41e-60	ALAVNDGGEK	SKEVRYRGVRRRPWGRYAAEIRDPVKK	RVWLGSFNTGEEAARAYDSAAIRFRGSKATTNF
At1g28160	34	6.15e-60	QQPQPQQH	IEEIKYGVRRRPWGRYAAEIRNPTTK	KERYWLGTFDTAEEAALAYDRAARS
At1g80580	111	1.57e-58	GGVKGGRKMP	QKTGGFMGVRKRPWGRWSAEIRDRI	GRHWLGTFDTAEEAARAYDAAARLRGTRKAKTNF
At5g18560	50	1.36e-57	GERRGRRSKQ	AEPGRFLGVRRRPWGRYAAEIRDPTTK	ERHWLGTFDTAHEAALAYDRAALS
At1g12890	16	1.22e-56	NKRKSKEEKK	LQEGKYLGVRRRPWGRYAAEIRNPFTK	ERHWLGTFDTAEEAAFAYDVAARS
			Multilevel	VKEVFRGVRKRPWGRYAAEIRDPGKK	KERVWLGTFDTAEEAARAYDAAARS
			consensus	A IHY R	WT T H L DL R
			sequence		S

CMVIII-1

Sequence name	Start	P-value	Site		
CaERF-23	205	6.25e-17	DDNISDGDNM	KRGGGLDLDLNFPPP	VEA
CaERF-24	205	1.36e-12	DDNISDGDNM	KRGGGLDLDLIFHHL	WKLEERLMR
At1g28370	149	6.60e-15	SVVMDVRYE	GRRVVLDDLNFPPP	PEN
At1g53170	169	9.47e-14	SVVDFEGGAG	KISPPLDLDLNLAPP	AE
At3g20310	210	4.51e-13	DGDDIASSSS	RRKTPFQFDLNFPP	DGVDLDFAGGI
At1g50640	192	4.51e-13	DDDDIASSSR	RRNPPFQFDLNFPP	DCVDLDFNGAD
At3g15210	204	3.75e-17	SVVDFEGGME	KRSQLDLDLNLPPP	SEQA
At5g44210	182	2.88e-11	SIGIRETVKV	EPRRELNLDLNLAPP	VVDV
At1g28360	172	6.17e-11	PMMNSSPSPV	TVRRGLAIDLNEPPP	LWL

Multilevel
consensus
sequence

KRRGGLDLLNFPPP
R SPPFQF LA L
GQL
R

CMVIII-2

Sequence name	Start	P-value	Site		
CaERF-22	184	3.12e-13	VVGRGRGGGA	VHSDSDTSSVDDNI	SDGDNMKRGG
CaERF-23	184	3.12e-13	VVGRGRGGGA	VHSDSDTSSVDDNI	SDGDNMKRGG
CaERF-24	189	1.52e-13	PRSPVPVDD	CHSDCDSSFSVIDDG	DCDIASSSIR
At3g20310	187	8.91e-15	PRTPPVAPED	CHSDCDSSSSVIDDG	DDIASSSRR
At3g15210	186	2.81e-14	VHVATKMACG	AQSDSDSSSVDFEG	GMEKRSQLLD
At1g53170	151	5.66e-14	ATYPKMTTCG	VQSESETSSVDFEG	GAGKISPLLD
At1g50640	169	8.28e-13	PRTPPVVPED	CHSDCDSSSSVIDDG	DDIASSSRR
At1g28370	131	2.96e-10	QVKAGMMVFD	ERSESDSSSVMDV	RYEGRRVLD
			Multilevel consensus sequence	VQSDSDSSSVDDDG CH E T FS IFNI AE S M D E V	

CMVIII-3 não contém

Grupo IX

AP2

Sequence name	Start	P-value	Site			
CaERF-27	116	4.79e-60	PPATAGAAQP	KGRHFRGVRQRPWGKFAAE	IRDPKNGARVWLGTYTEAEAAALAYDRAAYMRG	SKALLNFPHR
CaERF-28	104	4.79e-60	PPATAGAAQP	KGRHFRGVRQRPWGKFAAE	IRDPKNGARVWLGTYTEATAXAALAYDRAAYMRG	XKALLNSL
CaERF-29	116	3.24e-59	PPAAAGAAQP	KGRHFRGVRQRPWGKFAAE	IRDPKNGARVWLGTYTEAEAAALAYDRAAYMRG	SKALLNFPHR
CaERF-30	57	1.99e-61	KMLHRPSSHAT	GKKHYRGVRRRPWGKYAAE	IRDSTRNGTRVWLGTFFQTAEEAALAYDRAAFMRG	AKALLNFPSE
CaERF-31	156	4.07e-62	EQKKVQEGSG	EKKHYRGVRQRPWGKFAAE	IRDPNKRKGSRVWLGTFFDTAVEAAKAYDKAAFRLRG	SKAILNFPHE
CaERF-32	155	4.07e-62	EQKKVQEGSG	EKKHYRGVRQRPWGKFAAE	IRDPNKRKGSRVWLGTFFDTAVEAAKAYDKAAFRLRG	SKAILIXHTE
CaERF-33	155	4.07e-62	EQKKVQEGSG	EKKHYRGVRQRPWGKFAAE	IRDPNKRKGSRVWLGTFFDTAVEAAKAYDKAAFRLRG	SKAIPHFPHX
CaERF-34	95	1.66e-56	ENEIIFSLQK	EPEKQYIGVRKRPWGNNAE	IRDSTRNGIRVWLGTFFDTAEEGAVAYDHCWKWDV	KTWVENPQLQ
At5g47220	113	1.42e-61	KAIPTVETAV	KAKHYRGVRQRPWGKFAAE	IRDPKNGARVWLGTFFETAEDAALAYDIAAFMRG	SRALLNFPLR
At5g47230	152	2.04e-60	PEVTKPVSEE	EKKHYRGVRQRPWGKFAAE	IRDPNKRKGSRVWLGTFFDTAIEAARAYDEAAFRLRG	SKAILNFPLE
At5g61600	84	1.27e-58	TVSTKTEKEE	EERHYRGVRRRPWGKYAAE	IRDPNKKGCRIWLGTYDTAVEAGRAYDQAAFQLRG	RKAILNFPLE
At4g17490	133	1.84e-58	ELPVPVAAE	EKRHYRGVRRRPWGKFAAE	IRDSTRNGIRVWLGTFFETAIEAARAYDKEAAFRLRG	SKAILNFPLE
At5g51190	68	2.07e-58	PTTAPVVQEN	DQRHYRGVRRRPWGKYAAE	IRDPNKKGVRVWLGTFFDTAMEAARGYDKAAFRLRG	SKAILNFPLE
At5g43410	11	8.56e-57	MDQGGRGVGA	EHGKYRGVRRRPWGKYAAE	IRDSRKHGERVWLGTFFDTAEEAARAYDQAAYSMRG	QAAILNFPHE
At1g04370	16	1.07e-56	RSSGSGGGGA	EQGKYRGVRRRPWGKYAAE	IRDSRKHGERVWLGTFFDTAEDAARAYDRAAYSMRG	KAAAILNFPHE
At3g23240	78	1.66e-56	LPIKSVSRK	SEKSYRGVRRRPWGKFAAE	IRDSTRNGIRVWLGTFFESAEEAALAYDQAAFSMRG	SSAILNFSAE
At3g23220	10	2.07e-56	MERIESYNT	NEMKYRGVRRRPWGKYAAE	IRDSARHGARVWLGTFFNTAEDAARAYDRAAFMRG	QRAILNFPHE
At5g61590	103	3.96e-56	SNRKKTKRFE	ETRHYRGVRRRPWGKFAAE	IRDPKNGARVWLGTFFESDIDAARAYDYAAFRLRG	RKAVLNFPPLD
At5g07580	173	8.35e-56	KKAERFEEEV	DARHYRGVRRRPWGKFAAE	IRDPKNGARVWLGTFFESDIDAARAYDYAAFRLRG	RKAVLNFPPLD
At2g44840	88	8.35e-56	KASGSHAPRQ	KGMQYRGVRRRPWGKFAAE	IRDPKNGARVWLGTYTEPEDAAYDRAAFQLRG	SKAKLNFPHL
At4g18450	106	8.35e-56	SCSLSSRVSC	KTRKYRGVRRRPWGKFAAE	IRDSTRNGVRLVWLGTFFQTAEEAAMAYDKAAVRIRG	TQKAHTNFQL
At1g06160	78	1.74e-55	GKEATVAAEE	EKSYRGVRRRPWGKFAAE	IRDSTRNGIRVWLGTFFDTAEEAALAYDQAAFALKG	SLAVLNFPAD

At3g23230	16	1.36e-54	RSSNNQSQDD	KQARFRGVRRRP	WGKFAAEIRDPSRNGARLWLGTFETAEAAARAYDRAAFNLRG	HLAILNFPNE
At2g31230	82	2.47e-54	SEEEKVGID	EERSYRQVRRKRP	WGKFAAEIRDSTRNGIRVWLGTFDKAEAAALAYDQAAFATKG	SLATLNFVPE
			Multilevel	EKKHYRQVRRRP	WGKFAAEIRDPTKNGARVWLGTFDTAEAAARAYDRAAFRLRG	
			consensus	KER	Q Y SARK S E VD L K M	
			sequence		N Q	

CMIX-1 não contém

CMIX-2

Sequence name	Start	P-value	Site
CaERF-31	4	7.60e-30	MAS QTQITETSALELIKQHLLDDFAMEN YCVSDDRLSQ
CaERF-32	4	7.60e-30	MAS QTQITETSALELIKQHLLDDFAMEN YCVSDDRLSQ
CaERF-33	4	7.60e-30	MAS QTQITETSALELIKQHLLDDFAMEN YCASDDRLSQ
At5g47230	1	8.75e-24	. MATPNEVSALWFIKHLLEASPVAT DPWMKHESSE
At4g17490	1	3.60e-22	. MATPNEVSALFLIKKYLLEDELSPPT TATTNRWMND
At5g61600	1	9.72e-22	. MATKQEALAIQFISQHLITDFVSMET DHPSLFTNQL
At5g51190	7	1.38e-20	MASSHQ QQEQDQSQALDLITQHLLTDFPSLDT FASTIHHCTT
At5g61590	1	3.40e-20	. METFEESDLDVIQKHLFEDLMIPDG FIEDFVFDIT
			Multilevel
			consensus
			sequence
			MATINETSALIKQHLLDDFAMET
			QTQPT V D DF EK EELSPLDN
			FE S W Q T A SVAG
			Q

CMIX-3

Sequence name	Start	P-value	Site
CaERF-27	27	3.10e-50	EFPAYRRSS FSSLVPCLTETWGDPLKVDSEDMVIYWLLRDAVNVGWSP FNLTASSDVS
CaERF-28	27	3.10e-50	EFPAYRRSS FSSLVPCLTETWGDPLKVDSEDMVIYWLLRDAVNVGWSP FNLTASSDVS
CaERF-29	27	1.49e-49	EFPAYRGSCS FSSLVPCLTETWGDPLKVDPEMVIYWILRDAVNVGWSP FNLTASSDVS
CaERF-30	1	4.49e-54	MVPYPQAEPLNENDSQEMVLYEVLTEANAISPSN FPSSR
At4g17500	57	5.81e-47	IKPVYGRNPS FSKLYPCFTESWGDPLKENDSEDMVYGLNDAFHGGWEP SSSSDEDERS
At5g47220	34	2.49e-33	GGGENELRLN ESTPSSCFTESWGGPLKENDSEDMVYGLLKDAFHFDTS SDLSCLFDFP
At2g44840	20	1.28e-28	GVNSRMYFRN PSFNSVILNDNWSLPLSVDDSQDMAIYNTLRDAVSSGWTP SVPPVTSPAE
			Multilevel
			consensus
			sequence
			FSSLVPCLTESWGDPLKVDSEDMVYGLRDAVHVGWSP
			K Y F T EN VV WL N FNG E

CMIX-4 não contém

CMIX-5

Sequence name	Start	P-value	Site
CaERF-31	297	3.36e-21	DCGDVKGIFE IPPLSPSPHPPLGYSQLMVI
At4g17490	262	1.83e-27	DDWDLTEFLS MPPLSPLSPHPFPFGYPQLTVV
At5g47230	280	1.04e-25	CDWDLTGFLN FPPLSPLSPHPFPFGYSQLTVV
At5g51190	201	5.80e-16	WDGADSKDMG IFSVPLSPPSPHPGYSQLVVT
			Multilevel
			consensus
			sequence
			IPPLSPLSPHPFPFGYSQLTVV
			MFPVPLP C SL HP M I
			F S V T

CMIX-6

Sequence name	Start	P-value	Site
CaERF31	275	5.70e-21	RKGEERGLPT DPLTPSSWTAVWDCGDVKGIF EIPPLSPSP
At5g51190	180	1.72e-26	REEAQVQADA CPLTPSSWKGFDGADSKDMG IFSVPLLSPC
At5g61600	203	9.37e-21	ETTEAEVEFV VPLTPSSWMGFWDVGAGDGIF SIPPLSPTSP
		Multilevel	CPLTPSSWKGFDGADSKDIF
		consensus	D MAV CGAGDGMG
		sequence	V T V V

Grupo X

AP2

Sequence name	Start	P-value	Site
CaERF-35	51	7.35e-81	NRAQQEQGNQ RRRHYRGVRQRPWGKAAEIRDPKKAARVWLGTTFETAEEAALAYDEAALRFRGNKAKLNFPERVQG KTEYGYLTTR
CaERF-36	38	1.98e-77	PAEAAEETGE RRRRYRGVRQRPWGKAAEIRDPHKAARVWLGTTFETAEEAARAYDEAALRFRGNRAKLNFPENVRL LPPLQASSQS
At5g61890	86	1.68e-79	APTHQDQGD L RRRHYRGVRQRPWGKAAEIRDPKKAARVWLGTTFETAESAALAYDEAALRFRGNKAKLNFPERVQ GSNSTYYSSN
At5g13330	35	2.42e-78	DQHQPQQDQP RRRHYRGVRQRPWGKAAEIRDPKKAARVWLGTTFETAEEAALAYDRAALKFKGKAKLNFPERVQ PTTTTTISHA
At5g07310	88	7.54e-76	APPTQDQGLL RRRHYRGVRQRPWGKAAEIRDPKKAARVWLGTTFETAEEAALAYDNAALKFKGKAKLNFPERAQL ASNTSTTTGP
At5g50080	73	1.58e-75	SKAFPEEARE KRRRYRGVRQRPWGKAAEIRDPHRAARVWLGTTFETAEEAARAYDEAALRFRGNKAKLNFPEDVRI LPPPPPLLS
At5g64750	181	3.89e-74	SETSSFSGDQ PRRRYRGVRQRPWGKAAEIRDPFKAARVWLGTTFDNAESAARAYDEAALRFRGNKAKLNFPENVKL VRPASTEAPQ
At1g43160	57	4.17e-73	KECESSMSLE RPKKYRGVRQRPWGKAAEIRDPHKATRVWLGTTFETAEEAARAYDAAALRFRGSKAKLNFPENVGT QTIQRNSHFL
At2g33710	66	4.58e-72	SLTLQEKSNS RQRNYRGVRQRPWGKAAEIRDPNKAARVWLGTTFETAEEAALAYDKAAEFERGHKAKLNFPPEHIRV NPTQLYPSA
At4g34410	132	2.56e-61	GGGKIRKRKN KKNGYRGVRQRPWGKFAAEIRDPKKATRVWLGTTFETAEDAARAYDRAAIGFRGPRAKLNFPFDVYT SSVSSPVAAD
		Multilevel	RRRHYRGVRQRPWGKAAEIRDPKKAARVWLGTTFETAEEAALAYDEAALRFRGNKAKLNFPERVQL
		consensus	KK R HR T D E R R K K SR N RG
		sequence	S S T

CMX-1

Sequence name	Start	P-value	Site
CaERF-35	4	1.74e-12	MTA MVSALSQVIGS TDNNNSVAVV
At5g07310	44	7.62e-15	SARSQYDMRA MVSALTQVIGN QSSSHDNNQH
At5g61890	36	7.62e-15	SARSQHDMRV MVSALTQVIGN QSKSHDNIS
At5g50080	4	2.06e-12	MSA MVSALTQVISA RSQTEAEGAH
At5g13330	1	2.18e-11	. MVSALSRVIEN PTDPPVKQEL
At5g64750	76	3.97e-11	GFSREMMSA IVSALTHVVAG NVPQHQQGGG
At1g43160	1	4.27e-11	. MVSMLTNVVS ETEPSASATW
At4g34410	33	9.77e-11	QLSPEQEELSV IVSALQHVISG ENETAPCQGF
		Multilevel	MVSALTQVIGN
		consensus	I SH VSG
		sequence	

CMX-2 não contém

CMX-3 não contém

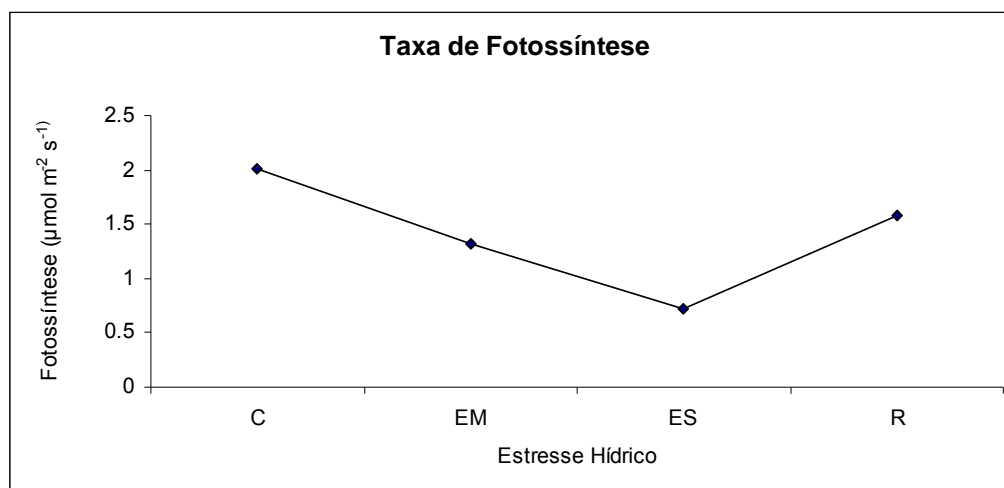
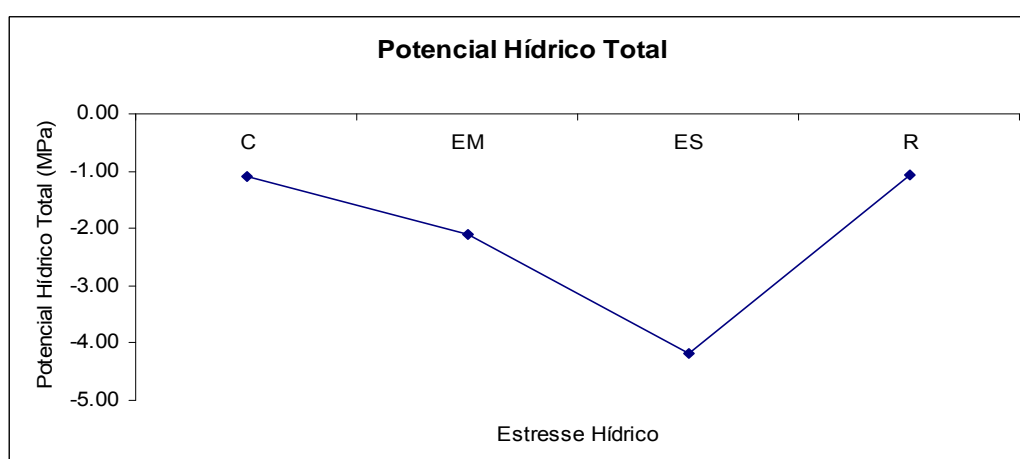
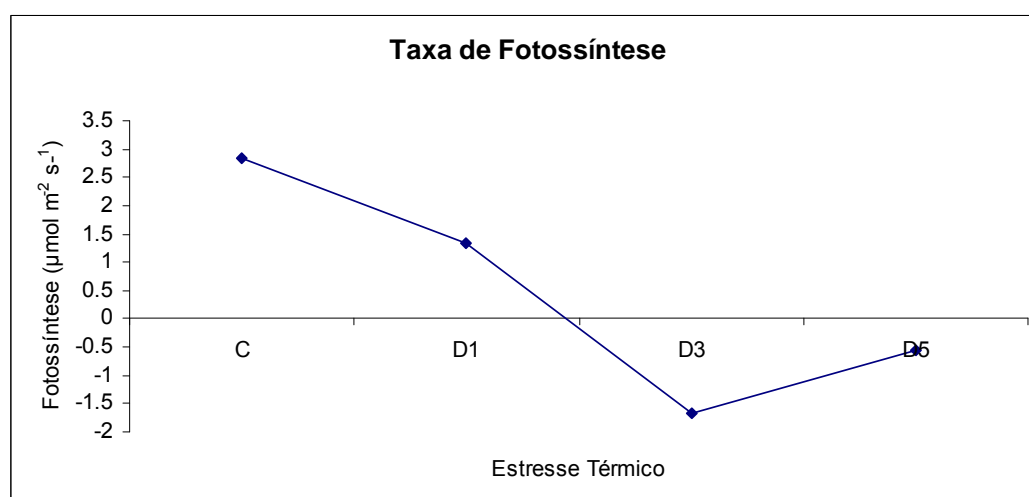
APÊNDICE C – Alinhamento entre as sequências completas de aminoácidos
CaERF17 de *Coffea arabica* e *CoERF* de *Coffea canephora*

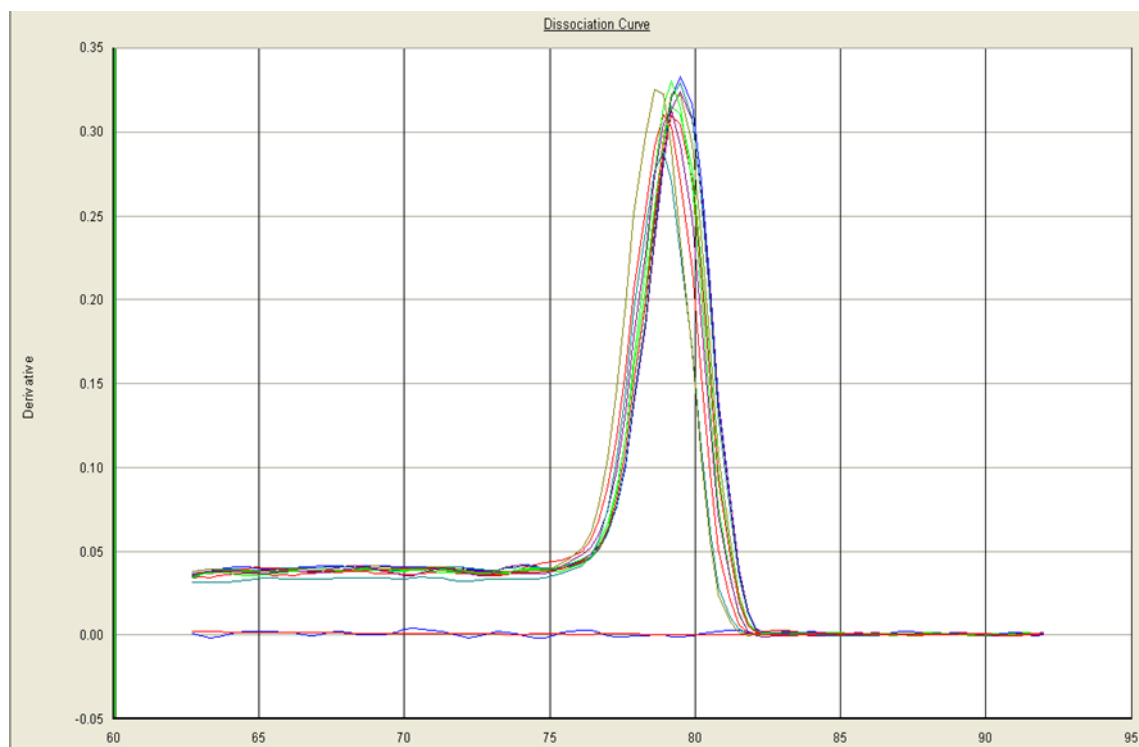
>lcl|63807 CoERF (Coffea canephora)
 Length=329

Score = 369 bits (948), Expect = 2e-107, Method: Compositional matrix adjust.
 Identities = 173/178 (97%), Positives = 174/178 (98%), Gaps = 0/178 (0%)

Query	1	MCGGAILEGLIPPRNNNDGTRRLSSADLWPSATDFWPNSHLAKPSRSNPTDAPHKATHET	60
		MCGGAILEGLIPPRNNNDGTRRLSSADLWPSA DFWPNSHLAKPSRSNPTDAPHKATHET	
Sbjct	1	MCGGAILEGLIPPRNNNDGTRRLSSADLWPSAADFWPNSHLAKPSRSNPTDAPHKATHET	60
Query	61	CHDPHDEAGGKQQVLPKGAKRQRKNLYRGIRQRPWGKWAAEIRDPKGVRVWLGTFTAE	120
		CHDPHDEAGGKQQVLPKGAKRQRKNLYRGIRQRPWGKWAAEIRDPKGVRVWLGTFTAE	
Sbjct	61	CHDPHDEAGGKQQVLPKGAKRQRKNLYRGIRQRPWGKWAAEIRDPRKGVRVWLGTFTAE	120
Query	121	EAARAYDKEARKIRGKKAKVNFNEDSTSHPTCIPSQPQYQQVHIPSPTVCGPSSSN	178
		EAARAYDKEARKIRGKKAKVNFNEDSTS+PTCIPSQ QYQQVHIPSPT CGPSSSN	
Sbjct	121	EAARAYDKEARKIRGKKAKVNFNEDSTSYPCTCIPSQTQYQQVHIPSPTFCGPSSSN	178

APÊNDICE D – Dados fisiológicos referentes aos estresses hídrico e térmico.

Figura 01 - Medidas da taxa de fotossíntese referente aos dias de estresse térmico**Figura 02** – Medidas de potencial hídrico referentes ao estresse hídrico.**Figura 03** - Medidas da taxa de fotossíntese referente aos dias de estresse térmico.

APÊNDICE E – Curvas de dissociação dos genes *RPL39* e *CaERF19***Figura 1** – Curva de dissociação do gene *RPL39*.**Figura 2** – Curva de dissociação do gene *CaERF19*.