



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE CIÊNCIAS**  
**DEPARTAMENTO DE BIOQUÍMICA E BIOLOGIA MOLECULAR**  
**CURSO DE PÓS-GRADUAÇÃO EM BIOQUÍMICA**

**KÁTIA DANIELLA DA CRUZ SARAIVA**

**CARACTERIZAÇÃO, ANÁLISE FILOGENÉTICA E PERFIL DE  
EXPRESSÃO DA FAMÍLIA MULTIGÊNICA DO FATOR DE  
ELONGAÇÃO 1 ALFA (EF1 $\alpha$ ) DE SOJA [*Glycine max* (L.) MERR.]:  
DETECÇÃO DE GENES NORMALIZADORES PARA qPCR**

**FORTALEZA – CE**

**2013**

**KÁTIA DANIELLA DA CRUZ SARAIVA**

**CARACTERIZAÇÃO, ANÁLISE FILOGENÉTICA E PERFIL DE  
EXPRESSÃO DA FAMÍLIA MULTIGÊNICA DO FATOR DE  
ELONGAÇÃO 1 ALFA (EF1 $\alpha$ ) DE SOJA [*Glycine max* (L.) MERR.]:  
DETECÇÃO DE GENES NORMALIZADORES PARA qPCR**

Dissertação submetida à coordenação do curso de Pós-Graduação *Stricto Sensu* em Bioquímica da Universidade Federal do Ceará, como parte dos requisitos para a obtenção do grau de Mestre em Bioquímica.

Orientador: Prof. José Hélio Costa, PhD.

**FORTALEZA – CE**

**2013**

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca de Ciências e Tecnologia

---

S246c

Saraiva, Kátia Daniella da Cruz

Caracterização, análise filogenética e perfil de expressão da família multigênica do fator de alongação 1 alfa (EF1 $\alpha$ ) de soja [*Glycine max* (L.) Merr.]: detecção de genes normalizadores para qPCR / Kátia Daniella da Cruz Saraiva. – 2013.

119 f. : il. color., enc. ; 30 cm

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Departamento de Bioquímica e Biologia Molecular, Programa de Pós-Graduação em Bioquímica, Fortaleza, 2013.

Área de Concentração: Bioquímica vegetal. Orientação:  
Prof. PhD. José Hélio Costa.

1. Expressão gênica. 2. Genes de controle endógeno. 3. Soja. I. Título.

---

CDD 574.192

**KÁTIA DANIELLA DA CRUZ SARAIVA**

**CARACTERIZAÇÃO, ANÁLISE FILOGENÉTICA E PERFIL DE  
EXPRESSÃO DA FAMÍLIA MULTIGÊNICA DO FATOR DE  
ELONGAÇÃO 1ALFA (EF1 $\alpha$ ) DE SOJA [*Glycine max* (L.) MERR.]:  
DETECÇÃO DE GENES NORMALIZADORES PARA qPCR**

Dissertação submetida à coordenação do curso de Pós-Graduação *Stricto Sensu* em Bioquímica da Universidade Federal do Ceará, como parte dos requisitos para a obtenção do grau de Mestre em Bioquímica. Área de concentração Bioquímica Vegetal.

Aprovada em 21/02/2013

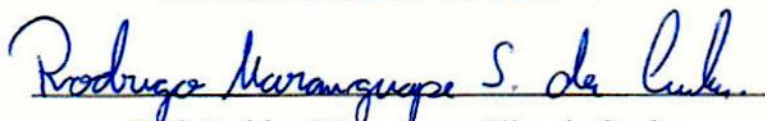
**BANCA EXAMINADORA**



Prof. José Hélio Costa  
Departamento de Bioquímica e Biologia Molecular  
Universidade Federal do Ceará - UFC



Prof. Dirce Fernandes de Melo  
Departamento de Bioquímica e Biologia Molecular  
Universidade Federal do Ceará - UFC



Prof. Rodrigo Maranguape Silva da Cunha  
Núcleo de Biotecnologia de Sobral-NUBIS  
Universidade Estadual Vale do Acaraú - UEVA

A minha família que sempre apoiou e incentivou meus estudos, não medindo esforços para a realização dos meus objetivos (Dedico).

Ao meu amado esposo Cicero Alisson dos Santos, por ser companheiro e estar presente, mesmo com a distância física que nos separa (ofereço).

## AGRADECIMENTOS

Nenhuma obra é fruto do trabalho individual. Por isso agradeço a todos que participaram direta ou indiretamente da realização deste trabalho, pois sem estas pessoas, este trabalho não se tornaria realidade.

Meus sinceros agradecimentos

A **Deus**, pois, sem sua ajuda, nada teria sido possível;

A **minha família**, pelo amor incondicional e apoio em todos os momentos da minha vida acadêmica e pessoal;

Ao meu orientador **Prof. Dr José Hélio Costa**, por todo o apoio, ensinamentos e confiança.

A Prof<sup>a</sup> **Dr<sup>a</sup> Dirce Fernandes de Melo** por ter me acolhido em seu laboratório e por todos os conhecimentos fornecidos.

Ao meu amado marido **Cícero Alisson dos Santos**, pelo amor, companheirismo e paciência. Apoiando-me mesmo estando distante.

Aos bolsistas de iniciação científica (Karine Thiers, Jamiley, Janaina, Leonardo) que ajudaram de forma direta na realização dos experimentos que compõem essa dissertação.

A todos os colegas de laboratório (Pedro, Geórgia, Nila, Flávio, Edson, Camila, Carine, Caroline, Lívia, Beatriz, Michele) pela acolhida e pelo apoio no decorrer do período de titulação.

## **AGRADECIMENTOS INSTITUCIONAIS**

Os procedimentos experimentais que compõem esta dissertação foram realizados com o apoio dos seguintes Programas/Instituições:

**UNIVERSIDADE FEDERAL DO CEARÁ** - através das atividades de capacitação no ensino e pesquisa em Bioquímica. E através do **Laboratório de Bioenergética dos sistemas vegetais**, onde os procedimentos experimentais foram realizados, sob a coordenação da professora Dra. Dirce Fernandes de Melo e Orientação do professor Dr. José Hélio Costa.

**PROGRAMA DE APOIO A PLANOS DE REESTRUTURAÇÃO E EXPANSÃO DAS UNIVERSIDADES FEDERAIS (REUNI)/ COORDENAÇÃO DE APERFEIÇOAMENTO PESSOAL DE NÍVEL SUPERIOR (CAPES)** - através da concessão de bolsa de mestrado e de fomentos à pesquisa.

**CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO (CNPq) e FUNDAÇÃO CEARENSE DE APOIO AO DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO (FUNCAP)** - pela contribuição para a melhoria nas condições experimentais e formação de recursos humanos dentro do Programa de Pós-Graduação em Bioquímica da UFC.

“O excelente mestre não é o que mais sabe, mas o que mais tem consciência do quanto não sabe. Não é o viciado em ensinar, mas o mais ávido em aprender. Não é o que declara seus acertos, mas o que reconhece suas próprias falhas. Não é o que deposita informações na memória, mas o que expande a maneira de ver, de reagir e de ser.”  
Augusto Cury.

“A persistência é o caminho do êxito”

(Charles Chaplin)

“O aprendizado nunca termina. Não existe parte da vida que não possua lições. Se você está vivo existem lições a aprender.” (Autor desconhecido).



## RESUMO

O fator de alongação 1alfa (EF1 $\alpha$ ) é responsável pela ligação do aminoacil-tRNA ao sítio A do ribossomo, atuando, assim, na fase de alongamento da síntese proteica. O EF1 $\alpha$  em plantas é codificado por uma família multigênica com número de genes variável entre espécies. O objetivo do presente trabalho foi caracterizar, analisar filogeneticamente e elucidar o perfil de expressão dos genes EF1 $\alpha$  em soja durante o desenvolvimento, bem como selecionar o(s) gene(s) de referência mais estáveis para ensaios de qRT-PCR no desenvolvimento e condições de estresse. Análises *in silico* nos genomas de soja e de outras leguminosas disponíveis em bancos de dados revelaram famílias gênicas de 3 a 6 membros com estrutura conservada de 2 éxons e 2 íntrons. Dentre as espécies analisadas (*Glycine max*, *Cajanus cajan*, *Vigna unguiculata*, *Medicago truncatula* e *Phaseolus vulgaris*) *G. max* foi a que apresentou o maior número de genes, classificados como EF1 $\alpha$  1a; 1b; 1c; 2a; 2b e 3 após análise filogenética. A expressão dos genes EF1 $\alpha$  foi estudada em diferentes tecidos (flores, vagens, sementes, cotilédones, folhas unifolioladas e trifolioladas, raízes, hipocótilos e epicótilos) durante o desenvolvimento da soja. A análise de expressão por qPCR revelou que os genes EF1 $\alpha$  são funcionais, sendo expressos em todos os tecidos, contudo apresentam expressão diferencial dependente do tecido e do estágio de desenvolvimento. A expressão relativa entre os genes membros mostrou predomínio de expressão para 5 dos 6 genes analisados: EF1 $\alpha$  1a: sem predomínio de expressão; EF1 $\alpha$  1b: Germinação e desenvolvimento do hipocótilo; EF1 $\alpha$  1c: Desenvolvimento da vagem; EF1 $\alpha$  2a: Desenvolvimento da vagem, folhas unifoliolada e trifoliolada; EF1 $\alpha$  2b: Desenvolvimento da vagem, folhas unifolioladas; EF1 $\alpha$  3: Germinação, flores, raízes, fases iniciais do desenvolvimento do hipocótilo e epicótilo. Esses dados, juntamente com os de folhas e de raízes submetidas a condições de estresse (PEG e AS) foram usados para avaliar a estabilidade de expressão dos diferentes genes EF1 $\alpha$  de soja através dos programas GeNorm e NormFinder. Para as análises durante o desenvolvimento, quatro genes de expressão estável (*UKNI*, *SKIP16*, *EF1 $\beta$*  e *MTP*) foram também utilizados. Os genes EF1 $\alpha$  foram mais estáveis em cotilédones (EF1 $\alpha$  3 e EF1 $\alpha$  1c); epicótilos (EF1 $\alpha$  1b, EF1 $\alpha$  2b e EF1 $\alpha$  1a); hipocótilos (EF1 $\alpha$  1a e EF1 $\beta$ ); vagens (EF1 $\alpha$  2a e EF1 $\alpha$  2b) e raízes (EF1 $\alpha$  2a e *UKNI*) e menos estáveis em folhas trifolioladas/unifolioladas (*UKNI* e *SKIP16*) e sementes em germinação (EF1 $\beta$  e *SKIP16*). Em folhas sob condições de estresse, análises através do GeNorm revelaram que quatro genes (EF1 $\alpha$  2a > 1b > 2b > 1a) apresentaram classificação de estabilidade semelhante nos estresses por AS e PEG. Já em raízes, distinto padrão de estabilidade foi encontrado: EF1 $\alpha$  2a > 1c > 2b > 1b > 1a > 3 para

PEG e *EF1α 1c > 3 > 2a > 1a > 2b > 1b* para AS. Tais resultados foram confirmados usando o programa NormFinder. Apesar das variações de expressão gênica em diferentes tecidos em desenvolvimento e condições de estresse, foram determinados genes *EF1α* que podem ser usados para normalizar ensaios de PCR em tempo real em soja e leguminosas (tribo Phaseoleae) vez que genes *EF1α* ortólogos foram identificados entre essas diferentes espécies.

Palavras - chave: Expressão gênica; genes de controle endógeno; Soja (*Glycine max*).

## ABSTRACT

The elongation factor 1alpha (EF1 $\alpha$ ) is responsible for binding of aminoacyl-tRNA to the A site of the ribosome, acting thus the elongation phase of protein synthesis. The EF1 $\alpha$  in plants is encoded by a multigene family with variable number of genes among species. The aim of this study was to characterize, analyze phylogenetic and elucidate the expression profile of EF1 $\alpha$  genes in soybean during development, as well as select (s) gene (s) of the most stable reference for qRT-PCR assays in the development and conditions stress. *In silico* analyses in the genomes of soybean and other leguminous plants available in databases revealed gene families 3-6 members with conserved structure of 2 exons and 2 introns. Among the analyzed species (*Glycine max*, *Cajanus cajan*, *Vigna unguiculata*, *Phaseolus vulgaris* and *Medicago truncatula*) *G. max* was the one with the largest number of genes classified as EF1 $\alpha$  1a, 1b, 1c, 2a, 2b and 3 after phylogenetic analysis. The EF1 $\alpha$  gene expression was studied in different tissues (flowers, pods, seeds, cotyledons, trifoliolate and unifoliolate leaves, roots, hypocotyls and epicotyls) during the development of soybean. Expression analyses by qPCR revealed that the EF1 $\alpha$  genes are functional and expressed in all tissues, however exhibit differential expression dependent on the tissue and developmental stage. Relative expression between genes members showed predominance of expression for 5 of the 6 genes analyzed: EF1 $\alpha$  1a: no predominance of expression; EF1 $\alpha$  1b: Germination and development hypocotyl; EF1 $\alpha$  1c: pods development; EF1 $\alpha$  2a: pods development, unifoliolate and trifoliolate leaves; EF1 $\alpha$  2b: pods development, unifoliolate leaves; EF1 $\alpha$  3: Germination, flowers, roots, early stages development (hypocotyl and epicotyl). These data, together with the leaves and roots subjected to stress conditions (PEG and AS) were used to assess the stability of expression of different EF1 $\alpha$  genes soybean through the programs GeNorm and NormFinder. For the analyses during development, four stable genes (*UKN1*, *SKIP16*, *EF1 $\beta$*  and *MTP*) were also used. EF1 $\alpha$  genes were more stable in cotyledons (EF1 $\alpha$  3 and EF1 $\alpha$  1c); epicotyls (EF1 $\alpha$  1b, 2b and 1a), hypocotyls (EF1 $\alpha$  1a, EF1 $\beta$  and EF1 $\alpha$  1c); pods (EF1 $\alpha$  2a and EF1 $\alpha$  2b) and roots (EF1 $\alpha$  2b and *UKN1*) and less stable in trifoliolate/unifoliolate leaves (*UKN1* and *SKIP16*) and germinating seeds (EF1 $\beta$  and *SKIP16*). In leaves under stress conditions, through GeNorm analyses revealed that four genes (EF1 $\alpha$  2a > 1b > 2b > 1a) showed similar stability in the classification of stress for AS and PEG. Already in roots, distinct pattern of stability was found: EF1 $\alpha$  2a > 1c > 2b > 1b > 1a > 3 for PEG and EF1 $\alpha$  1c > 3 > 2a > 1a > 2b > 1b to AS. These results were confirmed using the program NormFinder. In spite of changes in gene expression in different tissues in development and stress conditions were determined

*EF1 $\alpha$*  genes that can be used to normalize the qPCR assays in soybean and others leguminous plants (Phaseoleae tribe) seeing that *EF1 $\alpha$*  genes orthologous were identified between these different species.

**Keywords:** Endogenous control genes; gene expression; Soybean (*Glycine max*).

## LISTA DE FIGURAS

---

### REFERENCIAL TEÓRICO

---

- Figura 1** Planta de soja, mostrando as principais partes que a compõem.....5
- Figura 2** Mecanismo de ação da proteína EF1 $\alpha$  em eucariotos.....17

---

### CAPÍTULO I

---

- Figura 3** Etapas da extração de RNA total.....32
- Figura 4** Representação esquemática da estrutura dos seis genes *EF1 $\alpha$*  membros da família multigênica em *Glycine max* (soja).....41
- Figura 5** Análise filogenética dos genes *EF1 $\alpha$*  das fabales (*Mt: Medicago truncatula*, *Pv: Phaseolus vulgaris*, *Cc: Cajanus cajan*, *Vu: Vigna unguiculata* e *Gm: Glycine max*). Os genes *EF1 $\alpha$*  de *Gossypium hirsutum* (ordem malvales) foram utilizados como grupo externo.....42
- Figura 6** Expressão relativa dos diferentes membros da família gênica em soja.....44
- Figura 7** Análise da integridade do RNA extraído de vários tecidos de soja (cultivar BRS Pala) durante o desenvolvimento e submetidas a estresses em gel de agarose (1%), evidenciando as bandas do RNA ribossomal 18S e 28S. A definição das amostras está mostrada no gel.....49
- Figura 8** Eletroforese em gel de agarose (2%) dos produtos de amplificação do gene *Fator de alongação 1 alfa 1a IEF1 $\alpha$  1a*, para avaliação da qualidade do cDNA produzido na PCR semi - quantitativa.....54
- Figura 9** Curvas de dissociação obtidas por RT-qPCR usando *primers* específicos.....57

---

## CAPÍTULO II

---

- Figura 10** Média dos valores de *Ct* em vinte e quatro amostras de plantas submetidas a condições de estresse com ácido salicílico e polietileno glicol.....72
- Figura 11** Médias dos valores de *Ct* em amostras de diferentes tecidos durante o desenvolvimento da soja. As caixas representam as médias de *Ct* e as barras os desvios padrões para cada gene.....72
- Figura 12a** Classificação fornecida pelo GeNorm dos candidatos a genes de referência em diferentes tecidos durante o desenvolvimento da soja (*Glycine max*) .....75
- Figura 12b** Gráfico dos valores V fornecidos pelo GeNorm para os dez candidatos a gene de referência nas amostras de diferentes tecidos/órgãos durante o desenvolvimento de *Glycine max*.....78
- Figura 13** A: Classificação fornecida pelo GeNorm dos candidatos a genes de referência nas condições de estresse. Os genes de menores valores M são os mais estáveis. B: Gráfico dos valores V fornecidos pelo GeNorm para os seis candidatos a gene de referência nas amostras de *Glycine max* submetidas a condições de estresse.....79
- Figura 14** Expressão relativa do gene *FTL3*, usando os genes *EF1 $\alpha$ s* indicados pelo GeNorm.....84
- Figura 15** A: Expressão relativa do gene *LEA* em folhas B: Expressão relativa do gene *LEA* em raízes.....84
- Figura 16** A: Expressão relativa do gene *PR3* em folhas B: Expressão relativa do gene *PR3* em raízes.....84

## LISTA DE TABELAS

---

### REFERENCIAL TEÓRICO

---

|                 |  |   |
|-----------------|--|---|
| <b>Tabela 1</b> | Classificação botânica da soja (USDA).....                                       | 3 |
| <b>Tabela 2</b> | Estádios de desenvolvimento da soja proposto por (FEHR e CAVINESS, 1977).....    | 7 |
| <b>Tabela 3</b> | Principais Regiões Brasileiras produtoras de Soja e estimativas de produção..... | 9 |

---

### CAPÍTULO I

---

|                  |  |    |
|------------------|--|----|
| <b>Tabela 4</b>  | Órgãos e idade das plantas usadas para os ensaios durante o desenvolvimento.....   | 27 |
| <b>Tabela 5</b>  | Genes utilizados nas análises de expressão por RT – qPCR de <i>Glycine max</i> (L) MERR, sequência de iniciadores, tamanho do produto amplificado e temperatura de anelamento..... | 29 |
| <b>Tabela 6</b>  | Componentes e volumes usados nas reações de transcrição reversa.....   | 35 |
| <b>Tabela 7</b>  | Localização genômica dos genes <i>EF1<math>\alpha</math></i> de soja.....  | 39 |
| <b>Tabela 8</b>  | Caracterização e número de genes <i>EF1<math>\alpha</math></i> em espécies da ordem Fabales.....   | 45 |
| <b>Tabela 9</b>  | Identidade das sequências de cDNAs deduzidos dos genes <i>EF1<math>\alpha</math></i> em soja ( <i>Glycine max</i> ) obtidos do alinhamento no Clustal IW.....                      | 45 |
| <b>Tabela 10</b> | Identidade das sequências de aminoácidos deduzidos dos genes <i>EF1<math>\alpha</math></i> em soja ( <i>Glycine max</i> ) obtidos do alinhamento no Clustal IW.....                | 47 |
| <b>Tabela 11</b> | Identidade dos promotores dos genes <i>EF1A</i> em soja ( <i>Glycine max</i> ) obtidos do alinhamento no Clustal IW.....   | 47 |

|                  |   |
|------------------|---|
| <b>Tabela 12</b> | Valores médios das concentrações e relações de absorbâncias dos RNAs extraídos em vários tecidos de soja (cultivar BRS- Pala), durante o desenvolvimento e em condições de estresses.....51   |
| <b>Tabela 13</b> | Temperaturas de anelamento dos genes da família multigênica do <i>EF1 alfa</i> , <i>PR3</i> , <i>LEA</i> , <i>FTL3</i> , <i>MTP</i> , <i>UKN1</i> , <i>SKIP 16</i> e <i>EF1β</i> em soja, selecionadas após a realização do gradiente de temperatura em termociclador eppendorf realplex.....51 |
| <b>Tabela 14</b> | Cálculos da inclinação ( <i>slope</i> ) da curva e da eficiência de amplificação obtidos na RT-qPCR, usando iniciadores específicos.....52  |

---

## CAPÍTULO II

---

|                  |  |
|------------------|--|
| <b>Tabela 15</b> | Pares de primers utilizados nos ensaios de qRT-PCR .....68   |
| <b>Tabela 16</b> | Classificação dos genes de referência fornecida pelo programa NormFinder nas amostras de diferentes tecidos durante o desenvolvimento da soja.....82         |
| <b>Tabela 17</b> | Classificação dos genes de referência fornecida pelo programa NormFinder nas amostras de raiz e folha submetidas a condições de estresse com PEG e AS.....82 |



## LISTA DE ABREVIATURAS

1. BLAST – Ferramenta de busca de alinhamento local
2. cDNA – DNA complementar
3. Ct – Ciclo threshold
4. DAF – Dias após a floração
5. DAS – Dias após a semeadura
6. dNTP – Desoxirribonucleotídeo trifosfato
7. ddH<sub>2</sub>O – Água destilada
8. DNase - Desoxirribonuclease
9. EST – Expressed Sequence Tags
10. *EFl*  $\alpha$  – Fator de alongação 1 alfa
11. *EFl*  $\beta$  – Fator de alongação 1 beta
12. HKGs – Housekeeping genes
13. MgCl<sub>2</sub> – Cloreto de magnésio
14. mM – Milimolar
15. NCBI – Centro nacional de informação biotecnológica
16.  $\eta$ g – Nanograma
17. NaClO – Hipoclorito de sódio
18. Pb – Pares de bases
19. RPM – Rotações por minuto
20. RT – PCR – Transcrição reversa seguida da reação em cadeia da DNA polimerase
21. RT – qPCR - Transcrição reversa seguida da reação em cadeia da DNA polimerase quantitativa
22. Taq – *Thermus aquaticus*
23. T<sub>m</sub> – Temperatura de desnaturação
24. U – Unidades
25. USDA - Departamento de Agricultura dos Estados Unidos
26. WGS - Whole Genome Shotgun

## SUMÁRIO

|   |      |
|---|------|
| RESUMO.....   | viii |
| ABSTRACT.....   | x    |
| LISTA DE FIGURAS.....   | xii  |
| LISTA DE TABELAS.....   | xiv  |
| LISTA DE ABREVIATURAS.....  | xvi  |
| INTRODUÇÃO GERAL .....  | 1    |
| 1 REVISÃO DE LITERATURA.....  | 3    |
| 1.1 <i>Glycine max</i> (L.) MERR.....   | 3    |
| 1.1.1 Origem e classificação botânica.....  | 3    |
| 1.1.2 A soja no Brasil e no mundo .....   | 8    |
| 1.2 PCR quantitativa em tempo real (RT-qPCR).....                                 | 11   |
| 1.3 Genes constitutivos ou “ <i>Housekeeping</i> ” .....                          | 13   |
| 1.4 Fator de alongação.....   | 15   |
| 2 REFERÊNCIAS BIBLIOGRÁFICAS.....   | 18   |
| CAPÍTULO I.....   | 24   |
| 1 OBJETIVOS .....   | 25   |
| 1.1 Objetivo geral.....   | 25   |
| 1.2 Objetivos específicos.....  | 25   |
| 2 MATERIAIS E MÉTODOS .....   | 26   |
| 2.1 Material vegetal .....  | 26   |
| 2.2 Condições de crescimento .....  | 26   |
| 2.3 Análises <i>in silico</i> .....   | 27   |
| 2.3.1 Identificação <i>in silico</i> dos genes do fator de alongação 1 alfa ..... | 27   |
| 2.3.2 Alinhamento das sequências e análise filogenética.....                      | 28   |
| 2.3.3 Desenho dos oligonucleotídeos iniciadores.....                              | 28   |

|   |           |
|---|-----------|
| <b>2.4 Análise da expressão gênica pela Reação em Cadeia da DNA Polimerase quantitativa em tempo real (RT- qPCR).....</b> | <b>30</b> |
| 2.4.1 Extração de RNA total .....   | 30        |
| 2.4.2 Quantificação e pureza do RNA total.....  | 33        |
| 2.4.3 Integridade do RNA total .....  | 33        |
| 2.4.4 Síntese de DNA complementar (cDNA).....   | 33        |
| 2.4.5 Padronização para as reações de qPCR .....  | 35        |
| 2.4.5.1 Quantificação das amostras de RNA total e dos PRIMERS).....   | 35        |
| 2.4.5.2 Curva de eficiência dos primers.....  | 36        |
| 2.4.5.3 Gradiente de temperatura.....   | 36        |
| 2.4.5.4 PCR semi-quantitativa.....  | 36        |
| 2.4.5.5 RT – qPCR .....   | 36        |
| 2.4.6 Análise dos dados de expressão por RT-qPCR.....   | 37        |
| <b>3 RESULTADOS .....</b>   | <b>38</b> |
| 3.1 Qualidade do RNA .....  | 38        |
| 3.2 Qualidade do cDNA obtido pela reação de transcrição reversa .....   | 38        |
| 3.3 Escolha da melhor temperatura de anelamento .....   | 43        |
| 3.4 Especificidade e eficiência dos <i>primers</i> .....  | 43        |
| 3.3 Determinação da estrutura e número de genes EF1 $\alpha$ em espécies da ordem Fabales .....                           | 46        |
| 3.4 Análises das sequências de cDNAs, proteínas e promotores .....  | 50        |
| 3.5 Análises das relações filogenéticas entre os EF1 $\alpha$ de leguminosas.....   | 53        |
| 3.6 Perfil de expressão dos genes EF1 $\alpha$ em <i>Glycine max</i> .....  | 55        |
| <b>4 DISCUSSÃO .....</b>  | <b>58</b> |
| <b>5 CONCLUSÕES.....</b>  | <b>62</b> |
| <b>6 REFERÊNCIAS BIBLIOGRÁFICAS .....</b>   | <b>63</b> |
| <b>CAPÍTULO II.....</b>   | <b>65</b> |
| <b>1 OBJETIVOS .....</b>  | <b>66</b> |
| 1.1 Objetivo geral.....   | 66        |

|       |  |    |
|-------|--|----|
| 1.2   | Objetivos específicos.....   | 66 |
| 2     | MATERIAIS E MÉTODOS .....  | 67 |
| 2.2   | Condições de crescimento e desenho experimental .....                            | 67 |
| 2.3   | Extração de RNA total .....  | 67 |
| 2.4   | Ensaio de PCR em tempo real .....  | 67 |
| 2.5   | Desenho dos <i>primers</i> para as reações de qPCR.....                          | 67 |
| 2.6   | Seleção dos genes candidatos a normalizadores .....                              | 69 |
| 2.6.1 | Análise da estabilidade de expressão dos genes candidatos.....                   | 69 |
| 2.6.2 | Validação dos genes normalizadores.....  | 70 |
| 3     | RESULTADOS .....   | 71 |
| 3.5   | Perfil de Expressão dos genes candidatos a normalizadores.....                   | 71 |
| 3.6   | Análises da estabilidade de expressão dos genes candidatos a normalizadores..... | 73 |
| 3.6.1 | Análises pelo GeNorm.....  | 73 |
| 3.6.2 | Análises pelo NormFinder .....   | 80 |
| 3.7   | Validação dos genes de referência para RT-qPCR .....                             | 83 |
| 4     | DISCUSSÃO .....  | 85 |
| 5     | CONCLUSÕES.....  | 89 |
| 6     | REFERÊNCIAS BIBLIOGRÁFICAS .....   | 90 |
|       | ANEXOS.....  | 93 |

## INTRODUÇÃO GERAL

A soja (*Glycine max* (L) MERR) é um dos principais produtos agrícolas de exportação nacional, sendo cultivada praticamente em todo o território brasileiro. No período de 2010/2011 foram produzidas quase 75 milhões de toneladas do grão de soja, o que corresponde em torno de 28% da produção mundial (USDA). O grande interesse no cultivo de soja se deve a sua grande importância a nível nacional e mundial, devido ao seu alto conteúdo proteico (em torno de 40%) que serve como base para a alimentação de animais e humanos. Além de excelente fonte de proteínas e óleo vegetal a soja é rica em outros constituintes funcionais que fornecem benefícios à saúde e prevenção de algumas doenças (BENEVENTI, 2006).

A soja tem sido alvo de muitos estudos moleculares, visando o aumento da resistência da planta contra estresses bióticos e abióticos, melhoramento de características nutricionais (conteúdo de proteínas e óleo) e aumento da produtividade. A maior parte desses estudos envolve análise de expressão gênica, realizada principalmente através da técnica de qRT-PCR (reação em cadeia da DNA polimerase em tempo real). Essa técnica apresenta alta sensibilidade, não requer nenhum processamento pós-PCR e os resultados podem ser visualizados no momento em que a reação está acontecendo. Contudo, a referida técnica deve ser executada cuidadosamente, pois procedimentos, tais como erros de pipetagem, amostras com concentrações iniciais de RNA não homogêneas, ineficiências nas reações de transcrição reversa, podem invalidar os resultados. Desta forma, é necessária a adição de genes com expressão estável, a fim de normalizar essas variações experimentais.

Genes comumente utilizados para normalização são aqueles envolvidos em processos celulares básicos, podendo-se citar, os genes codificadores do RNA 18S, Actina (ACT), Gliceraldeído -3- fosfato desidrogenase (GAPDH), Fator de alongação 1 alfa (EF1 $\alpha$ ), Ubiquitina (UBQ),  $\alpha$ - Tubulina (TUA),  $\beta$ -Tubulina (TUB) dentre outros (CZECHOWSKI et al., 2005; HU et al., 2011).

O fator de alongação 1 $\alpha$  (EF1 $\alpha$ ) está envolvido em um processo celular básico, pois participa do processo de síntese proteica, promovendo a ligação do aminoacil – tRNA ao sítio A do ribossomo durante a fase de alongamento da síntese de proteínas, processo este dependente de GTP. Os genes *EF1 $\alpha$*  são codificados por famílias multigênicas em várias espécies de plantas, como exemplos podem ser citadas *Arabidopsis* (RANSON-HODGKINS

et al., 2009) e arroz (KIDOU e EJIRI, 1998), cujos genes pertencentes à família multigênica do *EF1 $\alpha$*  foram caracterizados, tendo seus padrões de expressão elucidados.

Na soja pouco se sabe sobre os genes que codificam o *EF1 $\alpha$*  em soja, tendo sido dois genes reportados por Aguilar et al (1991). Igualmente, a normalização de genes constitutivos de famílias multigênicas é escassa e em buscas feitas em bancos de dados do Pubmed ([www.pubmed.gov.com.br](http://www.pubmed.gov.com.br)) foram encontrados apenas os trabalhos de Huis et al (2010) e Lee et al (2010) Portanto, como a maioria dos genes constitutivos são codificados por famílias multigênicas, torna-se importante realizar estudos de normalização visando identificar quais membros poderiam ser utilizados em diferentes condições.

Sendo assim, objetivamos caracterizar a família multigênica do fator de alongação 1  $\alpha$  de soja, identificando o número de genes pertencentes a esta família, bem como analisar, filogeneticamente estes genes, comparando-os com genes homólogos de espécies pertencentes a ordem Fabales, espécies estas, com genoma sequenciado e disponível nos bancos de dados. Além desta caracterização, ensaios de qRT-PCR em situações de estresse e durante o desenvolvimento da soja serão executados a fim de identificar os genes mais estáveis pertencentes a família multigênica do *EF1 $\alpha$* , genes estes, que poderão ser utilizados como normalizadores em experimentos posteriores de qRT-PCR em soja e em espécies da ordem Fabales.

# 1 REVISÃO DE LITERATURA

## 1.1 *Glycine max* (L.) MERR

### 1.1.1 Origem e classificação botânica

O centro de origem da soja é a região leste da Ásia (LI et al., 2010). No Brasil, a soja foi introduzida inicialmente na Bahia, em 1882 e no século XIX foi levada para São Paulo. No Rio Grande do Sul (região produtora mais importante nos dias atuais) foi cultivada pela primeira vez em 1900 havendo uma expansão acentuada a partir de 1936 (COSTA, 1996). Na década de 80 foi introduzida na região dos Cerrados, tornando-se a cultura anual de maior área plantada no Brasil.

A soja é uma planta pertencente ao reino Plantae, divisão Magnoliophyta, classe Magnoliopsida, ordem Fabales, família Fabaceae (Leguminosae), subfamília Faboideae (Papilionoideae), gênero *Glycine*, espécie *Glycine max* e forma cultivada *Glycine max* (L.) Merr (MISSÃO, 2006) (Tabela 1). A soja tem como provável progenitor a espécie *Glycine ussuriensis* (LI et al., 2010).

**Tabela 1:** Classificação botânica da soja (USDA).

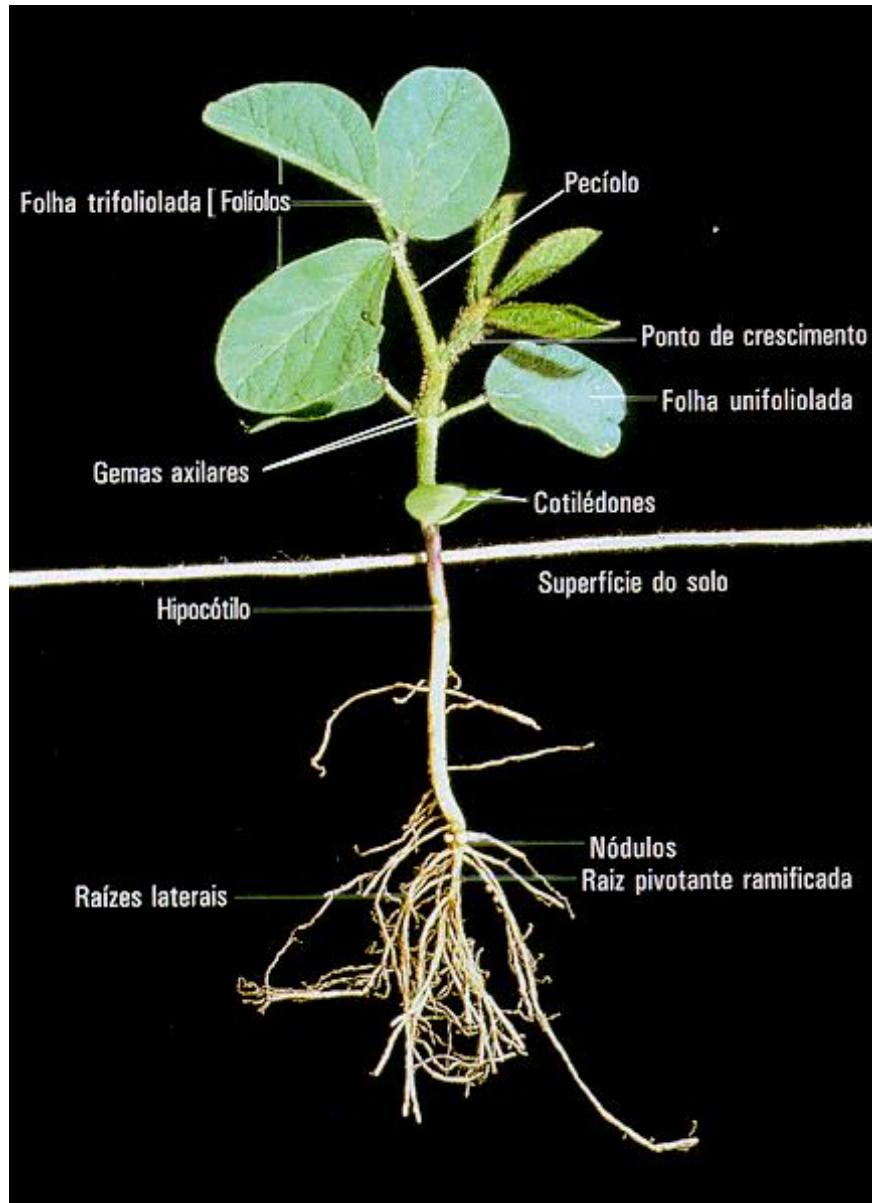
| Classificação botânica |                    |
|------------------------|--------------------|
| Reino                  | Plantae            |
| Divisão                | Magnoliophyta      |
| Classe                 | Magnoliopsida      |
| Ordem                  | Fabales            |
| Família                | Fabaceae           |
| Subfamília             | Faboideae          |
| Tribo                  | Phaseolae          |
| Subtribo               | Phaseolinae        |
| Gênero                 | <i>Glycine</i>     |
| Espécie                | <i>Glycine max</i> |

**Fonte:** Departamento de Agricultura dos Estados Unidos – USDA

A soja é uma cultura anual, herbácea, ereta, autógama, apresentando variabilidade quanto às características morfológicas, que ainda podem ser influenciadas pelo ambiente, como a altura que pode variar de 30 a 200 cm, apresentando mais ou menos ramificações. O sistema radicular da soja é constituído de raiz axial principal e de raízes secundárias (FEHR e CAVINESS, 1977; MÜLLER, 1981). Quanto ao ciclo de vida, este, pode variar de 75 dias para as culturas mais precoces e 200 dias para as mais tardias (EMBRAPA, 2012).

Ao longo do ciclo de vida, a soja possui três tipos de folhas: as cotiledonares, as simples ou unifolioladas e as compostas ou trifoliadas; todas com tamanho, formato e posicionamentos diferentes (Figura 1). Possui flores completas que ocorrem em racemos terminais ou axilares, variando de duas a trinta e cinco flores por racemo (ROCHA, 2009). A soja é considerada uma planta de dias curtos, ou seja, precisa de um mínimo de horas de noite ou escuro para indução floral.





**Figura 1:** Planta de soja, mostrando as principais partes que a compõem.

**Fonte:** Embrapa Soja

O fruto é do tipo vagem e pode chegar a 400 vagens por planta, com número de grãos variando de um a cinco por vagem. Contudo, a maioria das cultivares apresenta vagens com dois ou três grãos. Suas sementes possuem variações quanto à forma, tamanho e cor (EMBRAPA, 2012).

A soja é uma planta com grande variabilidade genética, tanto no ciclo vegetativo (período compreendido da emergência da plântula até a abertura das primeiras flores), como no reprodutivo (período do início da floração até o fim do ciclo da cultura).

Segundo a Embrapa (2012), o cultivo da soja vem se destacando cada vez mais como uma opção rentável e de alta qualidade para os agricultores da Região Nordeste do Brasil. Devido à importância supracitada, a soja, apresenta desenvolvimento fenológico bem caracterizado, compreendido por estádios de desenvolvimento vegetativos (V) e reprodutivos (R) – denominações propostas por (FEHR e CAVINESS, 1977) (Tabela 2).

**Tabela 2:** Estádios de desenvolvimento da soja proposto por (FEHR e CAVINESS 1977).

| ESTÁDIO | DENOMINAÇÃO                        | DESCRIÇÃO   |
|---------|------------------------------------|---|
| VE      | Emergência                         | Cotilédones acima da superfície do solo   |
| VC      | Cotilédone                         | Cotilédones completamente abertos   |
| V1      | Primeiro nó                        | Folhas unifolioladas completamente desenvolvidas  |
| V2      | Segundo nó                         | Primeira folha trifoliolada completamente desenvolvida  |
| V3      | Terceiro nó                        | Segunda folha trifoliolada completamente desenvolvida   |
| V4      | Quarto nó                          | Terceira folha trifoliolada completamente desenvolvida  |
| V5      | Quinto nó                          | Quarta folha trifoliolada completamente desenvolvida  |
| V6      | Sexto nó                           | Quinta folha trifoliolada completamente desenvolvida  |
| V...    | .....                              | .....   |
| Vn      | Enésimo nó                         | Anti-enésima folha trifoliolada completamente desenvolvida  |
| R1      | Início do florescimento            | Uma flor aberta em qualquer nó do caule (haste principal)   |
| R2      | Florescimento pleno                | Uma flor aberta num dos dois últimos nós do caule com folha completamente desenvolvida  |
| R3      | Início da formação da vagem        | Vagem com 5 mm de comprimento num dos quarto últimos nós do caule com folha completamente desenvolvida                                |
| R4      | Vagens completamente desenvolvidas | Vagem com 2 cm de comprimento num dos quarto últimos nós do caule com folha completamente desenvolvida                                |
| R5      | Início do enchimento do grão       | Grão com 3 mm de comprimento em vagem num dos quarto últimos nós do caule com folha completamente desenvolvida                        |
| R6      | Grão cheio ou completo             | Vagem contendo grãos verdes preenchendo as cavidades da vagem um dos quarto últimos nós do caule com folha completamente desenvolvida |
| R7      | Início da maturação                | Uma vagem normal no caule com coloração madura  |
| R8      | Maturação plena                    | 95% das vagens com coloração madura   |

**Obs:**

- 1. Nó cotiledonar não é considerado.**
- 2. Uma folha é considerada completamente desenvolvida quando as bordas do trifólio da folha seguinte não mais se tocam.**
- 3. Caule significa a haste principal da planta.**
- 4. A expressão “últimos nós” refere-se aos últimos nós superiores.**

### 1.1.2 A soja no Brasil e no mundo

Os principais países produtores mundiais de soja em ordem decrescente são os Estados Unidos, o Brasil e a Argentina (EMBRAPA SOJA, 2012).

A soja é uma espécie de leguminosa com grande valor econômico destacando-se como uma das mais importantes culturas do mundo. Em termos de produtividade de grãos no mundo a safra de 2011/2012 foi projetada em 867,73 milhões de toneladas (Departamento de Agricultura dos Estados Unidos – USDA). É a segunda principal oleaginosa cultivada no mundo, depois do dendê, apresentando – se, como um dos grãos mais consumidos mundialmente. Este grande consumo se deve as inúmeras aplicações da soja na indústria farmacêutica e alimentícia, podendo-se citar: a fabricação da proteína texturizada de soja, leite de soja, dentre outros (DALL’AGNOL et al., 2010). Os grãos de soja aparecem como a maior fonte de proteínas usadas na alimentação humana e animal, visto que, possui mais proteínas que outras fontes de alimentos comumente utilizadas, tais como carne bovina, ovos e peixes. (KRISHNAN, 2005; COPPER et al., 2008).

No Brasil, a soja ocupa a maior área plantada em relação às outras culturas (BARROS, 2006). No cenário mundial, o Brasil se destaca como o segundo maior produtor de grãos desta leguminosa (EMBRAPA SOJA, 2012). O desenvolvimento de variedades geneticamente adaptadas às diferentes condições de clima, solo e latitude proporcionaram o cultivo desta leguminosa em todas as regiões do país. Como melhor exemplo desta conquista genética, plantas originadas de latitudes elevadas têm sido transformadas tornando-se produtivas em médias e baixas latitudes.

O Mato Grosso, atualmente é o maior Estado produtor e também detentor da melhor produtividade média. O segundo e o terceiro Estados com maiores áreas plantados, Paraná e Rio Grande do Sul, estão na Região Sul onde a soja foi inicialmente cultivada comercialmente. Goiás, Minas Gerais e Bahia são os demais estados produtores. Os dados de produção da cultura são mostrados na (Tabela 3).

**Tabela 3:** Principais Regiões Brasileiras produtoras de Soja e estimativas de produção.

| REGIÃO/UF             | 2008/09<br>(toneladas) | 2009/10 Previsão<br>(toneladas) | 2010/11 Estimativa<br>(toneladas) |
|-----------------------|------------------------|---------------------------------|-----------------------------------|
| <b>NORTE</b>          | <b>497,6</b>           | <b>574,9</b>                    | <b>586,7</b>                      |
| RR                    | 8,0                    | 1,4                             | 2,1                               |
| RO                    | 106,0                  | 122,3                           | 129,4                             |
| AC                    | -                      | -                               | -                                 |
| AM                    | -                      | -                               | -                                 |
| AP                    | -                      | -                               | -                                 |
| PA                    | 72,2                   | 86,9                            | 86,9                              |
| TO                    | 311,4                  | 364,3                           | 368,3                             |
| <b>NORDESTE</b>       | <b>1.608,0</b>         | <b>1.861,7</b>                  | <b>1.966,5</b>                    |
| MA                    | 387,4                  | 502,1                           | 553,8                             |
| PI                    | 273,1                  | 343,1                           | 389,1                             |
| CE                    | -                      | -                               | -                                 |
| RN                    | -                      | -                               | -                                 |
| PB                    | -                      | -                               | -                                 |
| PE                    | -                      | -                               | -                                 |
| AL                    | -                      | -                               | -                                 |
| SE                    | -                      | -                               | -                                 |
| BA                    | 947,5                  | 1.016,5                         | 1.023,6                           |
| <b>CENTRO-OESTE</b>   | <b>9.900,1</b>         | <b>10.539,2</b>                 | <b>10.804,7</b>                   |
| MT                    | <b>5.828,20</b>        | <b>6.224,50</b>                 | <b>6.331,60</b>                   |
| MS                    | 1.715,8                | 1.712,2                         | 1.864,6                           |
| GO                    | 2.307,2                | 2.549,5                         | 2.549,5                           |
| DF                    | 48,90                  | 53,00                           | 59,00                             |
| <b>SUDESTE</b>        | <b>1.460,4</b>         | <b>1.591,2</b>                  | <b>1.637,9</b>                    |
| MG                    | 929,1                  | 1.019,0                         | 1.030,8                           |
| ES                    | -                      | -                               | -                                 |
| RJ                    | -                      | -                               | -                                 |
| SP                    | 531,3                  | 572,2                           | 607,1                             |
| <b>SUL</b>            | <b>8.277,0</b>         | <b>8.900,9</b>                  | <b>9.124,6</b>                    |
| PR                    | 4.069,2                | 4.485,1                         | 4.610,7                           |
| SC                    | 385,3                  | 439,6                           | 458,2                             |
| RS                    | 3.822,5                | 3.976,2                         | 4.055,7                           |
| <b>NORTE/NORDESTE</b> | <b>2.105,6</b>         | <b>2.436,6</b>                  | <b>2.553,2</b>                    |
| <b>CENTRO-SUL</b>     | <b>19.637,5</b>        | <b>21.031,3</b>                 | <b>21.567,2</b>                   |
| <b>BRASIL</b>         | <b>21.743,1</b>        | <b>23.467,9</b>                 | <b>24.120,4</b>                   |

Fonte: Empresa brasileira de agropecuária - EMBRAPA

No ano de 2010, as exportações brasileiras totalizaram 17,1 bilhões de reais (EMBRAPA, 2012) e em 2011/2012 a exportação brasileira do grão compreendeu 34% da exportação mundial DEPARTAMENTO DE AGRICULTURA DOS ESTADOS UNIDOS (USDA).

O relatório do USDA, publicado em 2011, estimou a produção brasileira em 75,00 milhões de toneladas para o ano agrícola de 2011/2012 e estoque final de 18,18 milhões de toneladas. O Brasil participou com 28% da produção mundial.

A importância dessa cultura é enfatizada pelo sequenciamento do genoma, cultivar Williams, publicado no ano de 2010 (SCHMUTZ et al., 2010). A anotação do genoma da soja revelou um elevado número de genes (em torno de 46.000 genes) apresentando quase o dobro dos genes encontrados em *Arabidopsis* (26.000 genes), primeira planta cujo genoma foi sequenciado no ano 2000 (THE ARABIDOPSIS GENOME INITIATIVE, 2000). Portanto, estudos serão necessários a fim de compreender a função da maioria desses genes em soja. Nesse sentido, a PCR quantitativa em tempo real (RT-qPCR), aparece como uma técnica robusta e precisa no estudo da expressão gênica.

## 1.2 PCR quantitativa em tempo real (RT-qPCR)

Estudos de expressão gênica são atualmente muito utilizados, uma vez que possibilitam o entendimento de muitos processos biológicos no âmbito molecular. Comparado aos métodos tradicionais de análise da expressão gênica - incluindo Northern blot, hibridização *in situ* e PCR semi – quantitativa, a qRT-PCR foi recentemente considerada como o método mais confiável para a detecção dos níveis de RNAm, devido à sua alta sensibilidade e nenhum processamento pós-PCR. O que permite uma comparação direta entre RNAs que diferem amplamente em suas concentrações (JIAN et al., 2008).

Metodologias baseadas na PCR quantitativa em tempo real (RT-qPCR) melhoraram a sensibilidade e a objetividade da interpretação dos resultados. O uso da RT-qPCR em análises das mais diversas amostras permitiu o processamento das informações em um único passo, diminuindo o tempo do processo (HU et al., 2009).

A RT-qPCR é baseada no processo de transcrição reversa, seguida da reação em cadeia da DNA polimerase (PCR). Nas reações de qPCR, moléculas fluorescentes se ligam covalentemente a nucleotídeos dupla fita, tornando-se marcadores durante a cinética da reação. Os produtos amplificados são monitorados a cada ciclo, permitindo uma interpretação rápida e precisa do processo de amplificação. Atualmente é o método de maior sensibilidade e especificidade na análise dos níveis de transcritos (BUSTIN, 2002).

A reação de PCR consiste na síntese bidirecional e repetitiva de DNA através da extensão de uma região específica (*amplicon*) delimitada pelos “primers” ou iniciadores. A reação de PCR requer um par de iniciadores (senso e anti-senso), quatro desoxirribonucleotídeos trifosfato (dNTPs), DNA polimerase termoestável e cofatores, tais como o magnésio. A PCR ocorre em três etapas: o primeiro passo é a desnaturação da dupla fita de DNA que acontece a uma temperatura em torno de 95 °C; o segundo passo é o anelamento (temperatura variável entre 50 e 65 °C) dos *primers* na região do DNA que será amplificada; a última parte do ciclo corresponde a extensão da fita de DNA a partir dos *primers* pela DNA polimerase termoestável, acontecendo a uma temperatura de aproximadamente 72 °C.

Quando a reação de qPCR é iniciada, os sinais de fluorescência são gerados por fluoróforos que são específicos para a fita dupla de DNA (dsDNA). (KLEIN et al., 2002). O

sinal é proporcional à quantidade dos produtos de PCR formados. O equipamento de PCR em tempo real detecta o acúmulo do produto amplificado durante cada ciclo da reação (BUSTIN, 2002). Os dados gerados são, então, medidos na fase exponencial da reação de PCR (HEID et al., 1996).

Uma importante descoberta para a aplicação dessa técnica foi o sistema “Taq Man”, que utiliza a atividade exonuclease 5' → 3' da enzima DNA polimerase I bacteriana na construção de sondas de oligonucleotídeos marcados, que emitem um sinal de fluorescência somente quando clivada. Esta técnica é baseada no princípio de transferência de energia ressonante por fluorescência (SCHMITTGEN et al., 2008). Atualmente, outros sistemas de detecção de níveis de transcritos têm sido desenvolvidos. No entanto nestes novos sistemas não há necessidade de hidrólise pela atividade nucleásica da enzima DNA polimerase (BUSTIN, 2000). Sendo assim, alternativas de menor custo como o uso de corantes intercalantes, como “SYBR Green”, são também usados com frequência na detecção do produto de PCR (GIULIETTI et al., 2001) e têm permitido a comparação dos diferentes níveis de expressão de um mesmo gene em diferentes cultivares, tecidos ou estádios de desenvolvimento. Entretanto, vale salientar que a técnica de qPCR necessita de normalização para a correta interpretação dos dados obtidos. Esta normalização é feita pelo uso de genes constitutivos, que apresentem expressão estável nas diferentes condições testadas.



### 1.3 Genes constitutivos ou “*Housekeeping*”

Um dos principais fatores em estudos de expressão gênica por RT – qPCR é a normalização dos dados obtidos por meio do uso de genes constitutivos de expressão estável. Esse é um método simples e amplamente utilizado, pois possibilita controlar variações na quantidade e qualidade do RNA, bem como nas diferenças de síntese de cDNA permitindo, desta forma, avaliar se as variações de expressão encontradas entre as diferentes amostras estão relacionadas aos tratamentos estabelecidos ou a artefatos experimentais.

A normalização é feita com base em um ou mais genes que possuam expressão uniforme (estável) nas condições experimentais em que se deseja avaliar a expressão gênica, podendo englobar a maioria das células de um organismo, diferentes fases ou etapas de desenvolvimento ou determinadas condições ambientais. Genes que possuem esse padrão de expressão estável ou uniforme são chamados genes constitutivos ou “housekeeping” – HKGs (EXPÓSITO - RODRÍGUEZ et al., 2008). Estes genes durante muito tempo foram escolhidos com base no conhecimento ou provável papel em processos celulares básicos, tais como, a manutenção da estrutura celular e do metabolismo primário (WANG et al., 2011).

Genes comumente utilizados para normalizar experimentos em plantas compreendem: codificadores do RNA 18S, Actina (ACT), Gliceraldeído - 3 - fosfato desidrogenase (GAPDH), Fator de alongação 1 alfa (EF1 $\alpha$ ), Ubiquitina (UBQ),  $\alpha$ - Tubulina (TUA),  $\beta$ -Tubulina (TUB) dentre outros (CZECHOWSKI et al., 2005; HU et al., 2011).

Durante anos os genes citados anteriormente foram utilizados como normalizadores em experimentos de NORTHERN *blots* e RT-PCR convencionais. No entanto, embora esses genes estejam envolvidos em processos celulares básicos, vários estudos (EXPÓSITO - RODRÍGUEZ et al., 2008; JIAN et al., 2008; DHEDA et al., 2004; DING et al., 2004; CZECHOWSKI et al., 2005; GUTIERREZ et al., 2008; HU et al., 2011; LONG et al., 2011; MOREIRA et al., 2011; PAOLACCI et al., 2009; REMANS et al., 2011; WANG et al., 2011) têm demonstrado que genes constitutivos nem sempre apresentam um padrão de expressão estável dependendo das condições ambientais, experimentais, estádios de desenvolvimento e tipo celular estudado. Portanto, estes autores argumentam a necessidade de determinar especificamente, bons genes normalizadores para a espécie em estudo, para as condições experimentais pelas quais o vegetal está sendo submetido, para o tipo de tecido estudado ou estádio de desenvolvimento.

Existem muitas publicações com animais e humanos, descrevendo a identificação de múltiplos genes de referência para normalização em experimentos por RT-qPCR, mas ainda são poucos os trabalhos com plantas, visto a diversidade existente (REMANS et al., 2011).

Devido à importância dos genes de referência para normalização em ensaios usando RT-qPCR, a estabilidade de vários genes “housekeeping” tem sido avaliada em vários organismos e condições específicas. Em plantas como arroz (JAIN et al., 2006; DING et al., 2004; KIM et al., 2003), poplar (BRUNNER et al., 2004), batata (NICOT et al., 2005), soja (JIAN et al., 2008; HU et al., 2009; MOREIRA et al., 2011; WANG et al., 2011; LE et al., 2012), café (CAVALLARI et al., 2009; GOULÃO et al., 2011), trigo (PAOLACCI et al., 2009; LONG et al., 2011) e *Arabidopsis thaliana* (CZECHOWSKI et al., 2005; REMANS et al., 2011) os genes constitutivos foram avaliados em diferentes condições.

Os genes constitutivos em sua grande maioria são codificados por famílias multigênicas, no entanto quando estes genes são testados em relação a sua estabilidade de expressão geralmente apenas um gene membro da família é analisado. Neste contexto não se menciona qual membro pertencente à família multigênica de uma dada espécie está sendo utilizado, dificultando, por exemplo, usar esse conhecimento para estabelecer genes de referência usando genes ortólogos em espécies próximas filogeneticamente.

## 1.4 Fator de alongação

O fator de alongação da síntese proteica em eucariotos - eEF1 é membro da família das proteínas G e é composto de quatro subunidades: A, B $\alpha$ , B $\beta$  e B $\gamma$  (BROWNING, 1996 apud CARNEIRO et al, 1999) sendo a subunidade EF1A também conhecida como EF1 $\alpha$ . Esta subunidade desempenha um importante papel na síntese proteica, promovendo a ligação do aminoacil – tRNA ao sítio A do ribossomo em um mecanismo dependente de GTP (XU et al, 2007). Durante a síntese de proteínas o EF1 $\alpha$  interage com o EF1 $\beta$  formando um complexo binário favorecendo, em seguida, a ligação do GTP a esse complexo formando um complexo ternário. Posteriormente, o EF1 $\beta$  se dissocia e o EF1 $\alpha$  ativado (ligado a GTP) interage com o aminoacil – tRNA promovendo sua ligação ao sítio A do ribossomo. Em seguida o GTP é clivado, liberando o GDP, o EF1 $\alpha$  inativo volta a se ligar com o EF1 $\beta$ . (Figura 2).

A subunidade EF1 $\alpha$  possui três domínios: O domínio I é responsável pela ligação ao GTP, o II pelo reconhecimento da extremidade 3' do tRNA e o III faz a ligação do aminoacil – tRNA ao ribossomo (ALVES, 2010). EF1 $\alpha$  foi caracterizado a partir de uma variedade de fontes e tem peso molecular de 45,000-55,000 kD, pI de 8,5 a 9,5 que podem ser modificados, quer por metilação e / ou fosforilação (POKALSKY et al, 1989).

A principal função do EF1 $\alpha$  é sem dúvida relacionada a síntese de proteínas, porém outras funções são atribuídas como desempenhar um papel na transdução de sinais no interior das células (YANG et al., 1990 ). Acredita-se também que o EF1 $\alpha$  possa se ligar a proteínas do citoesqueleto, como actina e tubulina. O EF1 $\alpha$  é um monitor sensível do pH intracelular e níveis de Ca<sup>2+</sup> (RANSOM-HODGKINS et al., 2000). Ativa a phosphoinositol 4 quinase, uma enzima pertencente à cascata de transdução de sinais e é ativada por fatores de crescimento (YANG et al., 1993; INFANTE et al., 2008). Além disso, têm-se encontrado uma expressão alterada do EF1 $\alpha$  em diversos tipos de tumores (ADAMS et al., 1992; TOMLINSON et al., 2005).

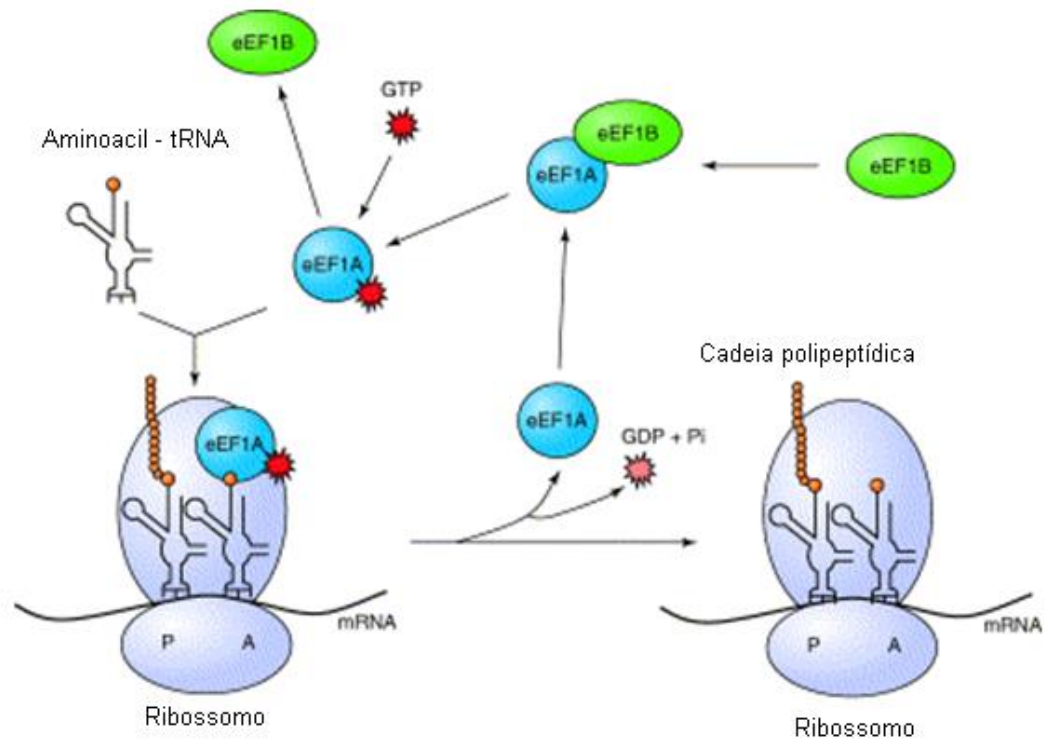
O EF1 $\alpha$  é a segunda proteína mais abundante na célula depois da actina, compreendendo 1-5% do teor total de proteína em células normais em crescimento (INFANTE et al, 2008).

O EF1 $\alpha$  em eucariotos é codificado por famílias multigênicas. Em plantas modelo como *Arabidopsis thaliana* e Arroz quatro genes são encontrados (AXELOS et al., 1989; RANDSON- HODGKINS, 2009; KIDOU e EJIRI, 1998). Já em milho, sete genes foram

relatados, mas há previsão de 10-15 genes (CARNEIRO et al., 1999). Seis genes são encontrados em tomate (POKALSKY et al., 1989), nove genes em algodoeiro (XU et al., 2007), seis genes em mandioca (SUHANDONO et al., 2001) e vinte genes em cana de açúcar (VIJAYKUMAR et al., 2002). O número elevado de cópias deste gene pode estar relacionado com a sua função na síntese de proteínas, a fim de assegurar a fidelidade da tradução.

Os genes *EF1 $\alpha$*  são altamente expressos em todos os tecidos de plantas, sendo considerados genes constitutivos “Housekeeping” e, portanto, são usados para normalizar experimentos de análise da expressão gênica por PCR em tempo real. No entanto a expressão destes genes varia em diferentes tecidos, estádios de desenvolvimento, condições experimentais ou estímulos ambientais (XU et al., 2007).

Em soja esses genes ainda não foram caracterizados e dessa forma, não há trabalhos avaliando o perfil de expressão de todos os membros da família gênica, portanto, fazendo-se necessário um estudo de expressão gênica, avaliando o comportamento dos diferentes genes membros em condições ambientais diversas. Assim, no presente trabalho avaliou-se a expressão dos genes *EF1 $\alpha$*  em diferentes tecidos durante o desenvolvimento da soja.



**Figura 2:** Mecanismo de ação da proteína EF1 $\alpha$  na síntese proteica em eucariotos (MITTAL et al., 2012).

## 2 REFERÊNCIAS BIBLIOGRÁFICAS

ADAMS SM, SHARP MG, WALKER RA, et al. Differential expression of translation-associated genes in benign and malignant human breast tumors. *Br J Cancer* 1992; 65: 65-71.

AGUILAR F, MONTANDON PE, STUTZ E. Two genes encoding the soybean translation elongation factor eEF-1 $\alpha$  are transcribed in seedling leaves. *Plant Mol Bio* 1991;17: 351-360.

ALVES RA. Caracterização funcional de complexos mRNA – proteínas (mRNPs) em *Trypanosoma cruzi*. 2010. 64p. Tese (Pós-Graduação em Biologia Celular e Molecular). Instituto Oswaldo Cruz, Curitiba.

AXELOS M, LIBOZ T, BARDET C, et al. The gene family encoding the Arabidopsis thaliana translation elongation factor-1A: molecular cloning, characterization and expression. *Molecular and General Genetics*. 1989; 219:106–112.

BARROS, B. A. Construção de cassetes de expressão para silenciamento gênico de fatores antinutricionais da Soja, via interferência de RNA. 2006. 102p. Dissertação. Universidade Federal de Viçosa. Minas Gerais.

BENEVENTI, M. A. Transformação genética em soja pela inserção da construção genética contendo a região promotora do gene rd29a e a região codante do gene dreb1a de Arabidopsis thaliana, visando tolerância à seca. 2006. 126p. Dissertação. Universidade Federal de Viçosa. Minas Gerais.

BRUNNER AM, YAKOVLEV IA, STRAUSS SH. Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biology* 2004; 4: 14.

BUSTIN SA. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology* 2000; 25: 169-193.

CARNEIRO NP, PETER AH, BRIAN AL. The eEF1A gene family is differentially expressed in maize endosperm. *Plant Molecular Biology* 1999; 41: 801-804.

CAVALLARI - Barsalobres CF, SEVERINO FE, MALUF MP, et al. Identification of suitable internal control genes for expression studies in *Coffea arabica* under

different experimental conditions.  
BMC Mol Biol 2009; 6:10-1.

COOPER JL, BRADLEY JT, ROBERT G, et al. Tilling to detect induced mutations in soybean. BMC Plant Biology 2008; 8: 9.

COSTA, J.A. Cultura da Soja. Porto Alegre. Evangraf. 1996. 233p.

CZECHOWSKI T, STIIT M, ALTMANN T, et al. Genome – wild identification and testing of superior reference –genes for transcript normalization in Arabidopsis. Plant Physiology 2005; 139: 15-17.

DALL'AGNOL A, LAZAROTTO JJ, HIRAKURI MH. Desenvolvimento, mercado e rentabilidade da soja brasileira. Londrina: Embrapa soja, 2010. 20 p.

DEPARTAMENTO DE AGRICULTURA DOS ESTADOS UNIDOS – USDA. Disponível em <<http://www.sistemafaep.org.br/noticia.aspx?id=562>> acesso em 01/03/2012.

DHEDA K, HUGGETT JF, BUSTIN SA, et al. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. Bio Techniques 2004; 37: 112–119.

DING J, JUNWEI J, LITAO Y, et al. Validation of a Rice Specific Gene, Sucrose Phosphate Synthase, Used as the Endogenous Reference Gene for Qualitative and Real-Time Quantitative PCR Detection of Transgenes. J Agric Food Chem 2004; 52: 3372-3377.

EMBRAPA. Cultivares de soja, regiões norte e nordeste do brasil. Embrapa soja, Londrina, PR, 2012. Disponível em <[http://www.cnpso.embrapa.br/index.php?op\\_page=294&cod\\_pai=17](http://www.cnpso.embrapa.br/index.php?op_page=294&cod_pai=17)>. Acesso em 02/08/2012.

EMBRAPA SOJA: <http://www.cnpso.embrapa.br/> acesso em 01/03/2012.

EXPOSITO-RODRIGUEZ M, BORGES AA, BORGESPEREZ A, et al. Selection of internal control genes for quantitative real-time RTPCR studies during tomato development process. BMC Plant Biology 2008; 8: 131.

FHER WR, CAVINESS CE. Stages on Soybean development. Ames: Iowa State University/Cooperative Extention Service. Special Report 1977; 80: 11.

GIULIETTI A, OVERBERGH L, VALCKX D, et al. An Overview of Real-Time Quantitative PCR: Applications to Quantify Cytokine. *Gene Expression* 2001; 25: 386-401.

GOULAO LF, FORTUNATO AS, RAMALHO JC. Selection of reference genes for normalizing quantitative real-time PCR gene expression data with multiple variables in *Coffea* spp. *Plant Mol Biol Rep* 2012; 30: 741-759.

GUTIERREZ L, MAURIAT M, GUÉNIN S et al. The lack of a systematic validation of reference genes: A serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. *Plant Biotechnol J* 2008; 6: 609-618.

HEID CA, STEVENS J, LIVAK KJ, et al. Real time quantitative PCR. *Genome Research* 1996; 6: 986–994.

HOAGLAND D, ARNON DI. The water culture method for growing plants without soil. *California Agriculture Experimental Station Circular* 1950 347p.

HU R, FAN C, LI H, et al. Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. *BMC Molecular Biology* 2009; 10: 93.

INFANTE C, ESTHER A, JOSÉ P, et al. Molecular characterization and expression analysis of five different elongation factor 1 alpha genes in the flatfish Senegalese sole (*Solea senegalensis* Kaup): Differential gene expression and thyroid hormones dependence during metamorphosis. *BMC Molecular Biology* 2008; 30: 9-19.

JAIN M, NIJHAWAN A, TYAGI AK, et al. Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. *Biochem. Biophys Res Commun* 2006 285: 645-651.

JIAN B, LIU B, BI Y, et al. Validation of internal control for gene expression study in soybean by quantitative real-time PCR. *BMC Molecular Biology* 2008; 9:59.

KIDOU S, EJIRI S. Isolation, characterization and mRNA expression of four cDNAs encoding translation elongation factor 1A from rice (*Oryza sativa* L.). *Plant Mol Biol* 1998; 36:137–148.

KIM BR, NAM HY, KIM SU, et al. Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice. *Biotechnol Lett* 2003; 25: 1869-1872.



- KLEIN CA, SEIDL S, PETAT-DUTTER K, et al. Combined transcriptome and genome analysis of single micrometastatic cells. *Nature Biotechnology* 2002; 20: 387–392.
- KRISHNAN H. Engineering soybean for enhanced sulfur amino acid content. *Crop Sci* 2005; 43: 454-461.
- LE DT, ALDRICH DL, VALLIYODAN B, et al. Evaluation of candidate reference genes for normalization of quantitative RT-PCR in soybean tissues under various abiotic stress conditions. *PLoS ONE* 2012; 7: 464-474.
- LI YH, LI W, ZHANG C, et al. Genetic diversity in domesticated soybean (*Glycine max*) and its wild progenitor (*Glycine soja*) for simple sequence repeat and single-nucleotide polymorphism loci. *New Phytologist* 2010; 188: 242–253.
- LIVAK KJ, SCHMITTEN TD. Analysis of gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods*. 2001; 25:402–408.
- LONG XY, LIU YX, ROCHELEAU H, et al. Identification and validation of internal control genes for gene expression in wheat leaves infected by Strip Rust. *International journal of plant breeding and genetic* 2011; 5: 255-267.
- LUND A, KNUDSEN SM, VISSING H. Assignment of human elongation factor 1 genes: EEF1A maps to chromosome 6q14 and EF1A2 to 20q13.3. *Genomics* 1996; 36: 359-361.
- MENESES, CHSG. Qualidade fisiológica de sementes de algodão submetidas a estresse hídrico induzido por polietilenoglicol-6000. 2007. 126p. Dissertação (Programa de Pós-Graduação em Agronomia) Universidade Federal da Paraíba, Paraíba, Brasil.
- MISSÃO MR. Maringá Management: *Revista de Ciências Empresárias* 2006; 3: 7-15.
- MITTAL N, SUBRAMANIAN G, BTIKOFER P, et al. Unique Posttranslational Modifications in Eukaryotic Translation Factors and their Roles in Protozoan Parasite Viability and Pathogenesis. *Molecular and Biochemical Parasitology* 2013; 127: 21-31.
- MIYASAKA S. *A Soja no Brasil*. 2000, 1062 p.
- MOREIRA RS, LEMOS EGM, ABDELNOOR RV, et al. Identification of reference genes for expression analysis by real time quantitative PCR in drought stressed soybean. *Pesq Agropec Bras* 2011; 46: 58-65.

- MÜLLER L. Morfologia, anatomia e desenvolvimento. In: MIYASAKA, S.; MEDINA, J.C., (Eds). A soja no Brasil. 1981; 6: 65-104.
- NICOT N, HAUSMAN J-F, HOFFMANN L, et al. Housekeeping geneselection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *J Exp Bot* 2005; 56: 2907-2914.
- PAOLACCI AR, ORONZO AT, ENRICO P, et al. Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *BMC Molecular Biology* 2009;10:11.
- POKALSKY AR, HIATT WR, RIDGE N, et al. Structure and expression of an elongation factor 1<sub>α</sub> in tomato. *Nucl Acids Res* 1989; 17: 4661–4673.
- RANSOM-HODGKINS WD. The application of expression analysis in elucidating the eukaryotic elongation factor one alpha gene family in *Arabidopsis thaliana*. *Mol Genet Genomics* 2009; 281: 391–405.
- REMANS T, KAREN S, KELLY O, et al. Normalization of real-time RT-PCR gene expression measurements in *Arabidopsis thaliana* exposed to increased metal concentrations. *Environmental Biology* 2011; 227: 1343-1349.
- ROCHA, R S. Avaliação de variedades e linhagens de soja em Condições de baixa latitude. 2009. 120p. Dissertação (Centro de Ciências Agrárias) Programa de Pós-Graduação em Agronomia, Universidade Federal do Piauí – UFPI. TERESINA, Piauí.
- SCHMITTGEN TD, LIVAK KJ. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* 2008; 3: 1101–1108.
- SEDIYAMA T. Melhoramento de soja (*Glycine max* (L.) Merrill) na Universidade Federal de Viçosa. In: SIMPÓSIO SOBRE CULTURA E PRODUTIVIDADE DA SOJA, 1., Piracicaba, Anais Piracicaba: FEALQ 1992; 1: 82-88.
- SUHANDONO S, HUGHES J, BROWN K, et al. Expression and structure of an elongation factor-1 $\alpha$  gene (MeEF1A) from cassava (*Manihot esculenta* Crantz). *Euphytica* 2001; 120: 49-58.
- THE ARABIDOPSIS GENOME INITIATIVE. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 2000; 4008: 796-815

TOMLINSON VA, NEWBERY HJ, WRAY NR, et al. Translation elongation factor eEF1A2 is a potential oncoprotein that is overexpressed in two-thirds of breast tumours. *BMC Cancer* 2005; 5:113.

VANDESOMPELE J, DE PRETER K, PATTYN F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002; 3: 18.

VIJAYKUMAR D. Molecular cloning, characterization and tissue specific expression of an elongation factor 1A gene in *Saccharum officinarum* L. *Plant Sci* 2002; 162: 315–321.

WANG YI, YU K, POYSA V, et al. Selection of reference genes for normalization of qRT-PCR analysis of differentially expressed genes in soybean exposed to cadmium. *MolBiol Rep* 2011; 39: 1585-1594.

WANG Y, BAO Z, ZHU Y, et al. Analysis of temperature modulation of plant defense against biotrophic microbes. *Mol Plant Microbe Interact* 2009; 22: 498–506.

XU W-L, WANG X-L, WANG H, et al. Molecular characterization and expression analysis of nine cotton GhEF1A genes encoding translation elongation factor 1A. *Gene* 2007; 389: 27-35.

YANG F, DEMMA M, WARREN V, et al. Identification of an actin-binding protein from *Dictyostelium* as elongation factor 1A. *Nature* 1990; 347: 494-496.

YANG W, BURKHART W, CAVALLIUS J, et al. Purification and characterization of a phosphatidylinositol 4-kinase activator in carrot cells. *J Biol Chem* 1993, 268: 392-398.

## **CAPÍTULO I**

# **CARACTERIZAÇÃO, ANÁLISE FILOGENÉTICA E PERFIL DE EXPRESSÃO DA FAMÍLIA MULTIGÊNICA DOS GENES EF1 $\alpha$ DURANTE O DESENVOLVIMENTO DA SOJA.**

# 1 OBJETIVOS

## 1.1 Objetivo geral

Caracterizar, analisar filogeneticamente e avaliar o perfil de expressão da família multigênica do *fator de alongação 1 alfa (EF1 $\alpha$ )* em diferentes tecidos durante o desenvolvimento da soja (*Glycine max*).

## 1.2 Objetivos específicos

- Caracterizar a família multigênica do *EF1 $\alpha$*  de soja (*Glycine max*), bem como de outras leguminosas (*Cajanus cajan*, *Medicago truncatula*, *Vigna unguiculata* e *Phaseolus vulgaris*) através de buscas em bancos de dados e anotação gênica;
- Analisar filogeneticamente a família multigênica do *EF1 $\alpha$*  de soja (*Glycine max*) e de outras leguminosas (*Cajanus cajan*, *Medicago truncatula*, *Vigna unguiculata* e *Phaseolus vulgaris*);
- Avaliar a expressão gênica da família multigênica do *EF1 $\alpha$*  em diferentes tecidos durante o desenvolvimento da soja (*Glycine max*).

## 2 MATERIAIS E MÉTODOS

### 2.1 Material vegetal

As sementes de soja (*Glycine max*) cultivar (BRS Pala – Safra de 2009), utilizadas para obtenção do material vegetal usado nos estudos de expressão gênica, foram doadas pela Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA Trigo, Passo Fundo, Rio Grande do Sul) e encaminhadas ao Laboratório de Bioenergética da Universidade Federal do Ceará (UFC) onde foram transferidas para frascos hermeticamente fechados e armazenadas em geladeira a 10<sup>0</sup>C.

### 2.2 Condições de crescimento

As sementes de soja (*Glycine max*), livres de danos mecânicos aparentes foram selecionadas e esterilizadas em solução de Hipoclorito de Sódio (NaClO) 1% durante cinco minutos, visando eliminar possíveis patógenos presentes em seu tegumento. Posteriormente as sementes foram deixadas em água destilada durante 2 horas. Transcorrido esse tempo as sementes foram lavadas em água corrente e em água destilada (ddH<sub>2</sub>O).

Estas sementes esterilizadas anteriormente foram germinadas em papel de filtro embebido com solução de Hoagland (25%) em casa de vegetação. Esse procedimento de germinação em solução de Hoagland foi adotado para evitar sintomas de deficiência nutricional que foram observadas quando as sementes eram germinadas apenas em água destilada. Seis dias após a semeadura, as mudas foram transferidas para hidroponia contendo a solução nutritiva de Hoagland e Arnon (1950) concentrada a 100%.

As coletas dos órgãos para extração de RNA total foram feitas com 10, 22, 45 dias após a semeadura (DAS) no início da floração, 54 dias (nove dias após o início da floração) e 63 dias (dezoito dias após o início da floração) para estudos nos estádios de desenvolvimento. Foram coletados raízes, folhas unifolioladas e trifolioladas, cotilédones, hipocótilos, epicótilos, flores e vagens. Também foram coletadas sementes no início da germinação (semente seca e a semente embebida 24 horas na solução de Hoagland) (Tabela 4).

**Tabela 4:** Órgãos coletados e idade das plantas usadas para os ensaios durante o desenvolvimento.

| Órgão extraído   | Idade da planta |
|--|-----------------|
| Semente seca.  | -               |
| Semente embebida na solução nutritiva de Hoagland e Arnon (1950).                  | 24h             |
| Folha unifoliolada, folha trifoliolada, cotilédones, raiz, epicótilo e hipocótilo. | 10 DAS          |
| Folha unifoliolada, folha trifoliolada, cotilédones, raiz, epicótilo e hipocótilo. | 22 DAS          |
| Folha trifoliolada, raiz, epicótilo e hipocótilo e flores.                         | 45 DAS          |
| Vagens   | 09 e 18 DAF     |

DAS- Dias após a sementeira e DAF- Dias após a floração

Após as coletas dos órgãos, as amostras foram congeladas imediatamente em nitrogênio líquido e armazenadas no freezer  $-80^{\circ}\text{C}$  até o momento da extração de RNA total.

### 2.3 Análises *in silico*

#### 2.3.1 Identificação *in silico* dos genes do fator de alongação 1 alfa em *Glycine max* e em espécies da ordem Fabales.

A caracterização dos genes pertencentes à família multigênica do *EF1 $\alpha$* , foi feita através de buscas no NCBI (<http://www.ncbi.nlm.nih.gov>), usando como banco de dados o WGS (Whole genome shotgun), onde as sequências obtidas do sequenciamento do genoma da soja estão depositadas. Desta forma, os genes identificados foram anotados. A anotação consistiu em identificar nos genes *EF1 $\alpha$*  a região dos promotores e a região transcrita {subdividida em regiões não traduzidas 5' e 3' [5'UTR (interrompida por 1 intron) e 3' UTR] e região codante (constituída de 2 exons e interrompida por intron)}. Após a anotação do gene, deduziu-se a sequência de cDNA que, foi então, traduzida na sequência de aminoácidos através da ferramenta de tradução do servidor da web Expasy (<http://www.expasy.org/>). De maneira semelhante, os genes *EF1 $\alpha$*  de outras leguminosas da ordem Fabales, tais como: *Cajanus cajan* (NCBI: <http://www.ncbi.nlm.nih.gov/>), *Medicago truncatula* (Phytozome: <http://www.phytozome.net/>), *Vigna unguiculata* (<http://www.ncbi.nlm.nih.gov/> Banco de dados GSS) e *Phaseolus vulgaris* (Phytozome: <http://www.phytozome.net/>) foram anotados e as sequências de cDNAs e de aminoácidos foram deduzidas

### 2.3.2 Alinhamento das sequências e análise filogenética

Após a anotação dos genes das espécies citadas anteriormente, procedeu-se a análise filogenética a partir dos cDNAs anotados. Estas sequências, juntamente com as de *Gossypium hirsutum* da ordem Malvales (usada como grupo externo) foram alinhadas através do programa CLUSTAL W e a relação evolucionária determinada pelo programa MEGA 5.05 empregando o método neighbor-joining (SAITOU e NEI, 1987) com valores de bootstrap (1.000 replicatas) (TAMURA et al., 2011).

### 2.3.3 Desenho dos oligonucleotídeos iniciadores

Os oligonucleotídeos iniciadores (*primers*) usados para as reações de PCR em tempo real foram desenhados usando o programa Perl primer V1.1.20 (MARSHALL, 2011). Os *primers* foram desenhados a partir da região 3' não traduzida (3'UTR), devido essas regiões serem menos conservadas em comparação com a região codante que apresenta alta identidade entre suas sequências de nucleotídeos.

Os *primers* desenhados variaram em comprimento de 19 a 25 pb, com temperatura de anelamento entre 50,7 a 63,6 °C e tamanho do fragmento a ser amplificado entre 60 a 199 bp. Os *primers* desenhados são mostrados na Tabela 5. Nesta mesma tabela também são mostrados os genes *MTP*, *EF1β*, *SKIP 16* e *UKNI* (HU et al, 2009) que foram usados como genes de referência para verificação da expressão relativa dos genes *EF1α* durante o desenvolvimento e em diferentes tecidos da soja. Os alinhamentos usados para o desenho dos *primers* podem ser encontrados nos anexos.



**Tabela 5:** Genes utilizados nas análises de expressão por RT – qPCR de *Glycine max* (L) Merr, sequência de iniciadores, tamanho do produto amplificado e temperatura de anelamento.

| Gene                         | Cromossomo | Produto gênico                                   | Sequência do primer (5'-3')          | Temperatura de anelamento °C | Eficiência de amplificação (%) | Tamanho do amplicon (pb) |                 |
|------------------------------|------------|--|--------------------------------------|------------------------------|--------------------------------|--------------------------|-----------------|
| <i>EF1a1a</i>                | 05         | Fator de alongação 1 $\alpha$                    | Fwd 5' GATTCATGTAGCCGTAGCC 3'        | 59.8                         | 113                            | 182                      |                 |
|                              |            |  | Rev 5' ATTAAGACATCCCTCCTCAG 3'       |                              |                                |                          |                 |
| <i>EF1a2a</i>                | 05         | Fator de alongação 1 $\alpha$                    | Fwd 5' GGATGTCGTTTCTTATGGT 3'        | 57.5                         | 82                             | 189                      |                 |
|                              |            |  | Rev 5' CAAACACAACACATTAACACAG 3'     |                              |                                |                          |                 |
| <i>EF1a1c</i>                | 16         | Fator de alongação 1 $\alpha$                    | Fwd 5' ATGCGATCATAGTTACATTTATAG 3'   | 52.9                         | 90                             | 198                      |                 |
|                              |            |  | Rev 5' AGTTCTCATAACAGCTTATAAAATAG 3' |                              |                                |                          |                 |
| <i>EF1a1b</i>                | 19         | Fator de alongação 1 $\alpha$                    | Fwd 5' ATCATCGTGGTTACTCCTTTAT 3'     | 59.8                         | 95                             | 199                      |                 |
|                              |            |  | Rev 5' TCAGACTCTTCTTACCATCA 3'       |                              |                                |                          |                 |
| <i>EF1a2b</i>                | 17         | Fator de alongação 1 $\alpha$                    | Fwd 5' TTTCTGACTCTTGTGCTTCT 3'       | 57.5                         | 95                             | 159                      |                 |
|                              |            |  | Rev 5' ACAACACAACACATTAACACAC 3'     |                              |                                |                          |                 |
| <i>EF1a3</i>                 | 10         | Fator de alongação 1 $\alpha$                    | Fwd 5' GTGCTGGATGGTTTATCAAGG 3'      | 63.6                         | 129                            | 140                      |                 |
|                              |            |  | Rev 5' AAAGTTGCGAGATGACGGAG 3'       |                              |                                |                          |                 |
| <i>EF1<math>\beta</math></i> |            | Fator de alongação 1 $\beta$                     | Fwd 5' CCACTGCTGAAGAAGATGATGATG 3'   | 63.6                         | 94                             | 134                      | Hu et al., 2009 |
|                              |            | Rev 5' AAGGACAGAAGACTTGCCACTC 3'                 |                                      |                              |                                |                          |                 |
| <i>SKIP 16</i>               |            | SKIP16/Proteína de interação 16                  | Fwd 5' CCACTGCTGAAGAAGATGATGATG 3'   | 50.7                         | 102                            | 60                       | Hu et al., 2009 |
|                              |            | Rev 5' AAGGACAGAAGACTTGCCACTC 3'                 |                                      |                              |                                |                          |                 |
| <i>MTP</i>                   |            | Metaloprotease, enzima de degradação da insulina | Fwd 5' CGCTCCAAGTGCTCCTCATTAG 3'     | 63.6                         | 93                             | 71                       | Hu et al., 2009 |
|                              |            | Rev 5' TGAAGTAACCGACGCCAACG 3'                   |                                      |                              |                                |                          |                 |
| <i>UKN1</i>                  |            | Proteína hipotética                              | Fwd 5' TGGTGCTGCCGCTATTTACTG 3'      | 63.6                         | 96                             | 74                       | Hu et al., 2009 |
|                              |            |  | Rev 5' GGTGGAAGGAAGTCTAACAATC 3'     |                              |                                |                          |                 |

## 2.4 Análise da expressão gênica pela Reação em Cadeia da DNA Polimerase quantitativa em tempo real (RT- qPCR)

### 2.4.1 Extração de RNA total

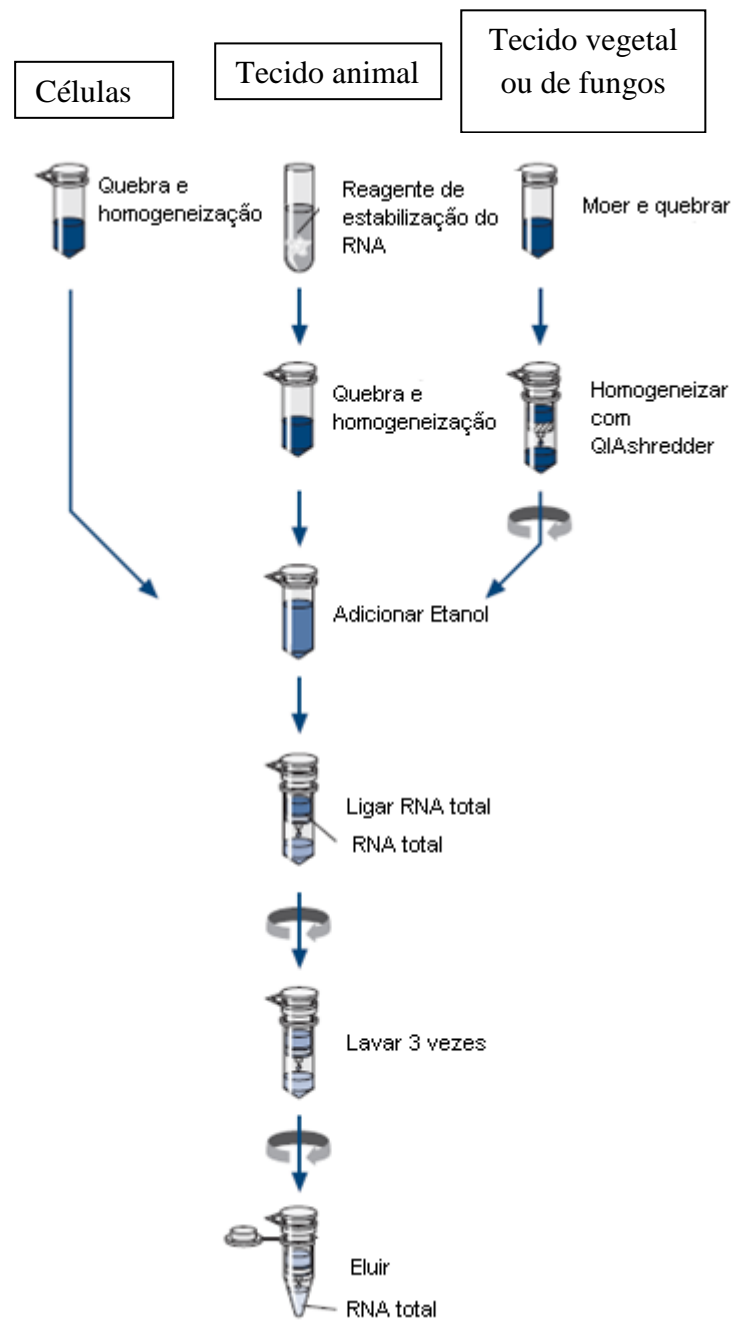
Para extração de RNA total dos órgãos coletados, 200 mg do material vegetal foi macerado em nitrogênio líquido, utilizando-se almofariz e pistilo. Do material vegetal macerado obtido, fez-se a extração de RNA total, através do *RNeasy plant mini Kit* (Qiagen) de acordo com as instruções do fabricante.

Às amostras maceradas foram adicionados 700  $\mu\text{L}$  do tampão RLT e 10  $\mu\text{L}$  de  $\beta$  – mercaptoetanol, homogeneizadas usando-se o vórtex, e a solução obtida foi transferida através de pipetagem para uma mini coluna do kit (QIAshredder spin column - lilás, acoplada a um tubo coletor de 2 mL) e centrifugada a 25  $^{\circ}\text{C}$  por 2 minutos a 14.000 RPM. O sobrenadante foi recuperado e transferido para um novo tubo eppendorf, onde foram adicionadas 0,5 vezes do volume inicial de etanol 95% (400l) e a solução obtida foi misturada por inversão.

A solução obtida anteriormente (em torno de 700  $\mu\text{L}$ ) foi transferida para uma nova coluna do kit (RNeasy mini spin column – rósea, acoplada a um tubo coletor de 2 ml) e centrifugada a temperatura ambiente (25  $^{\circ}\text{C}$ ) por 15 segundos a 10.000 RPM. O eluído foi descartado sendo adicionado na coluna rósea 350  $\mu\text{L}$  de tampão RW1, centrifugando-se a temperatura ambiente por 15 segundos a 10.000 RPM. Foram aplicados 80  $\mu\text{L}$  de desoxirribonucleases (DNase) na coluna, sendo esta deixada a temperatura ambiente durante 15 minutos. Transcorrido este tempo, adicionou-se novamente 350  $\mu\text{L}$  de RW1 e fez-se nova centrifugação a temperatura ambiente por 15 segundos a 10.000 RPM. Após a última lavagem com RW1 a coluna foi transferida para um novo tubo coletor e foram adicionados 500  $\mu\text{L}$  de RPE na coluna, que foi submetida a uma centrifugação a temperatura ambiente por 15 segundos a 10.000 RPM. O eluído foi descartado e foram adicionados mais 500  $\mu\text{L}$  de RPE na coluna, submetendo-a a uma nova centrifugação a temperatura ambiente por 2 minutos a 10 mil RPM, a fim de eliminar qualquer resíduo de etanol que possa interferir em reações posteriores.

O eluído e o tubo coletor foram descartados e a coluna foi transferida para um novo tubo coletor de 1,5 mL. Em sequência foram adicionados 40  $\mu\text{L}$  de água livre de RNases diretamente na coluna, submetendo a nova centrifugação por 1 minuto a 10.000 RPM para eluição do RNA. O RNA eluído é armazenado a -20  $^{\circ}\text{C}$  para posterior quantificação, análises

eletroforéticas, reações de transcrição reversa (TR) e ensaios de RT- qPCR. Os passos de extração do RNA total usando-se o *RNeasy plant mini Kit* estão resumidos na Figura 3.



**Figura 3:** Etapas da extração de RNA total.

### **2.4.2 Quantificação e pureza do RNA total**

Após extração do RNA total a concentração e pureza das amostras foram determinadas através do espectrofotômetro Nanodrop 2000 (Thermo Scientific) em comprimentos de onda de 260nm. Para estimar a pureza do RNA extraído a razão de 260/280 nm de absorbância foi utilizada para identificar contaminações por proteínas (razão deve estar entre 1,8 a 2) e a razão de 260/230 nm de absorbância utilizada para contaminações por polissacarídeos (razão deve estar acima de 2).

### **2.4.3 Integridade do RNA total**

A integridade do RNA total e contaminação com DNA genômico foram verificadas usando-se 0,5 µg de RNA total. Para isso foi realizada uma eletroforese em gel de agarose a 1%. A agarose (ainda líquida) foi colocada em uma bandeja apropriada para inserção de pentes, o que possibilita a formação de canaletas (poços). Nas canaletas foram aplicados um volume X de amostra, volume este determinado de acordo com a concentração do RNA total extraído, 1,5 µL do tampão da amostra (Azul de bromo fenol), sendo um volume final de 5 µL completado com água DEPC (água livre de RNases) .

Após solidificação, o gel de agarose foi colocado em uma cuba eletroforética, e esta foi preenchida com tampão MOPs 10X e ligada a um gerador que polariza a cuba com amperagem constante de 35 mA. A corrida eletroforética durou aproximadamente 30 minutos a temperatura ambiente (25 °C). Devido ao RNA possuir carga elétrica negativa em pH neutro, esta molécula migra entre os poros do gel em direção ao polo positivo da cuba, o que permite a formação de bandas separadas de acordo com o tamanho molecular do fragmento de RNA. Após a corrida eletroforética o gel foi corado com brometo de etídeo (0,5 µg/mL) durante 10 minutos. O brometo de etídeo é um agente intercalante de RNA/DNA, que permite a visualização das bandas de DNA/RNA, quando exposto à luz ultravioleta.

Para visualização do RNA ribossomal (18S e 28S) no gel de agarose foi utilizado o transiluminador de luz UV, sendo o RNA fotodocumentado por meio do sistema Mini BIS Pro (Bio-Imaging Systems) com auxílio do software GEICapture™.

### **2.4.4 Síntese de DNA complementar (cDNA)**

Após as análises de pureza e integridade descritas anteriormente, o RNA total foi utilizado para a síntese do DNA complementar (cDNA), utilizando o Kit da IMpromII™

Transcriptase Reverse (Promega, Estados Unidos) de acordo com as instruções do fabricante. Na reação foram adicionados aproximadamente 1 µg de RNA total, oligodT<sub>24</sub>, MgCl<sub>2</sub> 25 mM, dNTPs 10 mM, tampão de reação 5X e água livre de RNases (Quiagen). A Tabela 6 mostra os reagentes utilizados para esta reação, bem como os seus respectivos volumes.

A solução obtida após a adição dos reagentes supracitados (19 µl) foi submetida a 65 °C por 5 minutos para desnaturação e, posteriormente, transferidos para o gelo. Logo após, 1 µL da enzima transcriptase reversa foi adicionada em cada tubo de reação, completando o volume para 20 µL.

A reação de síntese da fita complementar ocorreu a 42 °C por 1 hora, seguida de 75 °C por 15 minutos. O cDNA obtido foi armazenado a -20 °C para posteriores reações de qPCR.

**Tabela 6.** Componentes e volumes usados nas reações de transcrição reversa.

| Reagentes                        | Volumes   |
|----------------------------------|---|
| RNA total                        | Variável em função da leitura espectrofotométrica para 1 µg de RNA total. |
| Tampão 5X                        | 4 µL  |
| MgCl <sub>2</sub> (25 mM)        | 2,4 µL  |
| dNTPs (10 mM)                    | 1,0 µL  |
| OligodT24 (20 pmols/µL)          | 3,0 µL  |
| H <sub>2</sub> O livre de RNAses | Variável em função do volume de RNA total                                 |
| Transcriptase reversa (160 µ/µL) | 1,0 µL  |

## 2.4.5 Padronização para as reações de qPCR

### 2.4.5.1 Quantificação das amostras de RNA total e dos iniciadores específicos (PRIMERS).

Inicialmente o RNA obtido foi quantificado em espectrofotômetro Nanodrop 2000 (Thermo Scientific), como descrito anteriormente. A quantificação foi realizada em triplicata, sendo feita uma média destas leituras. Este procedimento foi adotado a fim de evitar possíveis erros de quantificação que poderiam levar a resultados não confiáveis. Posteriormente foi realizada uma corrida eletroforética visando identificar se as amostras de RNAs estavam realmente homogêneas.

Após a quantificação e determinação da homogeneidade das amostras de RNA, foram realizadas as reações de transcrição reversa. Os cDNAs obtidos após as reações de RT foram diluídos para uma concentração final de 25 ng/uL considerando-se que a concentração inicial era em torno de 1000 ng/µL para todas as amostras.

Inicialmente os *primers* foram diluídos para uma concentração de 500 µM, sendo posteriormente, quantificados em triplicata usando-se o espectrofotômetro Nanodrop 2000 (Thermo Scientific). Da média das concentrações obtidas das leituras foi calculada a concentração real de cada *primer*, sendo estes, posteriormente diluídos para uma concentração

de 6  $\mu\text{M}$ , a fim de se obter uma concentração final de 300 nM/uL, concentração esta que foi utilizada nas reações de qPCR.

#### **2.4.5.2 Curva de eficiência dos primers**

Para mensurar a eficiência dos *primers* foram utilizadas diluições seriadas do *pool* de cDNA (1:1; 1:10; 1:100; 1:1000; 1:10000). A curva foi feita através da qPCR no termociclador Mastercycler ep realplex (Eppendorf). A eficiência foi calculada usando a equação  $Ef = 10^{(-1/\text{slope})} - 1$ .

#### **2.4.5.3 Gradiente de temperatura**

A fim de determinar a temperatura ótima para cada par de primer foi feito um gradiente de temperatura usando um *pool* das amostras de cDNA. O procedimento foi realizado em termociclador Mastercycler ep realplex 4S (Eppendorf®). As temperaturas usadas foram (47,2; 50,5; 52,7; 55; 57,5; 59,8; 63,6 e 65 °C). Foram selecionadas as temperaturas de menor Ct e com as curvas mais típicas (sigmóides).

#### **2.4.5.4 PCR semi-quantitativa**

A reação de PCR semi-quantitativa foi realizada para testar a qualidade do cDNA, produzido como descrito anteriormente. Utilizou-se para tal procedimento o par de *primers* referente ao gene *EF1 $\alpha$  1a*. Para cada reação foi utilizado 1  $\mu\text{L}$  de cDNA, acrescido de tampão de reação 5x, 1  $\mu\text{L}$  de dNTPs (5 mM), 1  $\mu\text{L}$  de cada iniciador específico (*primers*), 0,1  $\mu\text{L}$  da enzima Go Taq DNA polimerase. O volume total foi completado com água livre de RNases, totalizando 25  $\mu\text{L}$ . O resultado da amplificação foi visualizado em gel de agarose 2%.

#### **2.4.5.5 RT – qPCR**

As reações de amplificação por PCR quantitativa foram realizadas a partir do cDNA, obtido das reações de transcrição reversa das amostras de RNA e diluídos para uma concentração de 25  $\eta\text{g}/\mu\text{L}$ . A diluição das amostras de cDNA foi feita considerando-se a concentração de 1000  $\eta\text{g}/\mu\text{L}$ . Nas reações de qPCR foram utilizados 4  $\mu\text{L}$  de cDNA para cada reação, totalizando 100  $\eta\text{g}/\mu\text{L}$  por poço. Alíquotas da mesma amostra de cDNA foram utilizadas para todas as reações gene-específicas.



A qPCR foi realizada de acordo com o seguinte protocolo: para cada reação (poço da placa) foram adicionados 10  $\mu\text{L}$  do Power SYBR Green PCR Master Mix 2X (*Applied Biosystems*), 4  $\mu\text{L}$  de cDNA (100  $\eta\text{g}$ ), 1  $\mu\text{L}$  de cada oligonucleotídeo iniciador (senso e anti-senso), em uma concentração de 300  $\eta\text{g}/\mu\text{L}$ , e água ultrapura (Milli-Q) autoclavada e livre de nucleases, totalizando 20  $\mu\text{L}$  de reação. O monitoramento em tempo real da PCR foi realizado em um termociclador Mastercycler ep realplex 4S (Eppendorf®) através da detecção dos níveis de fluorescência do SYBR Green. As análises dos dados de fluorescência obtidos foram realizadas pelo *Realplex Software*. As reações ocorreram em placas com 96 poços e foram realizadas em triplicata e conduzidas de acordo com os seguintes parâmetros: 1) 95<sup>0</sup> C por 10 minutos; 2) 40 ciclos: 15 segundos de desnaturação a 95<sup>0</sup> C, 15 segundos de anelamento, variando a temperatura de acordo com cada par de primer (Tabela 2.6), e 20 segundos de extensão a 60<sup>0</sup> C. Também foi realizada a curva de dissociação (Melting), que consistiu de 15 segundos de incubação a 95<sup>0</sup> C, 15 segundos de incubação a 60<sup>0</sup> C e uma subida de 95<sup>0</sup> C.

#### 2.4.6 Análise dos dados de expressão por RT-qPCR

Após a realização das reações de PCR em tempo real, os níveis de expressão gênica foram determinados pelo número de ciclos de amplificação necessários para a fluorescência (emitida pelo SYBR Green) ultrapassar um limiar durante a fase exponencial da reação de PCR, determinado pelo  $C_t$  (*Cycle threshold*). Os  $C_t$ s utilizados para as análises foram obtidos através da média aritmética entre as triplicatas de reações para cada condição e gene. Para normalização dos resultados, foi utilizada a equação  $\Delta C_t = C_t$  (gene alvo) –  $C_t$  (controle endógeno). Os valores brutos de  $C_t$  foram convertidos em dados de expressão relativa utilizando-se o método delta  $C_t$  e delta-delta  $C_t$  (LIVAK & SCHMITTGEN, 2001) com o auxílio do programa GeNorm (VAMDESOMPELE et al., 2002).

## 3 RESULTADOS

### 3.1 Qualidade do RNA

O RNA total foi extraído de vários tecidos (flores, sementes, vagens, cotilédones, folhas unifolioladas, folhas trifolioladas, raízes, hipocótilos e epicótilos) em diferentes estádios de desenvolvimento. O RNA total também foi isolado de folhas e raízes de plantas com treze dias após a semeadura em duas condições de estresse (AS e PEG).

A quantificação em espectrofotômetro mostrou que as relações 260/280 e 260/230 ficaram entre 2 e 2,03 e 1,0 e 2,29, respectivamente. Com base nas relações 260/280 e 260/230, a qualidade do RNA extraído foi considerada boa. A média de concentração do RNA total foi de 136,9 e 1149  $\eta\text{g}/\mu\text{L}$ . (Tabela 7).

Em relação à integridade, os perfis observados no gel de agarose demonstram a integridade dos RNAs analisados, já que foram visualizadas as bandas do RNA ribossomal referentes às subunidades 18S e 28S (Figura 4).

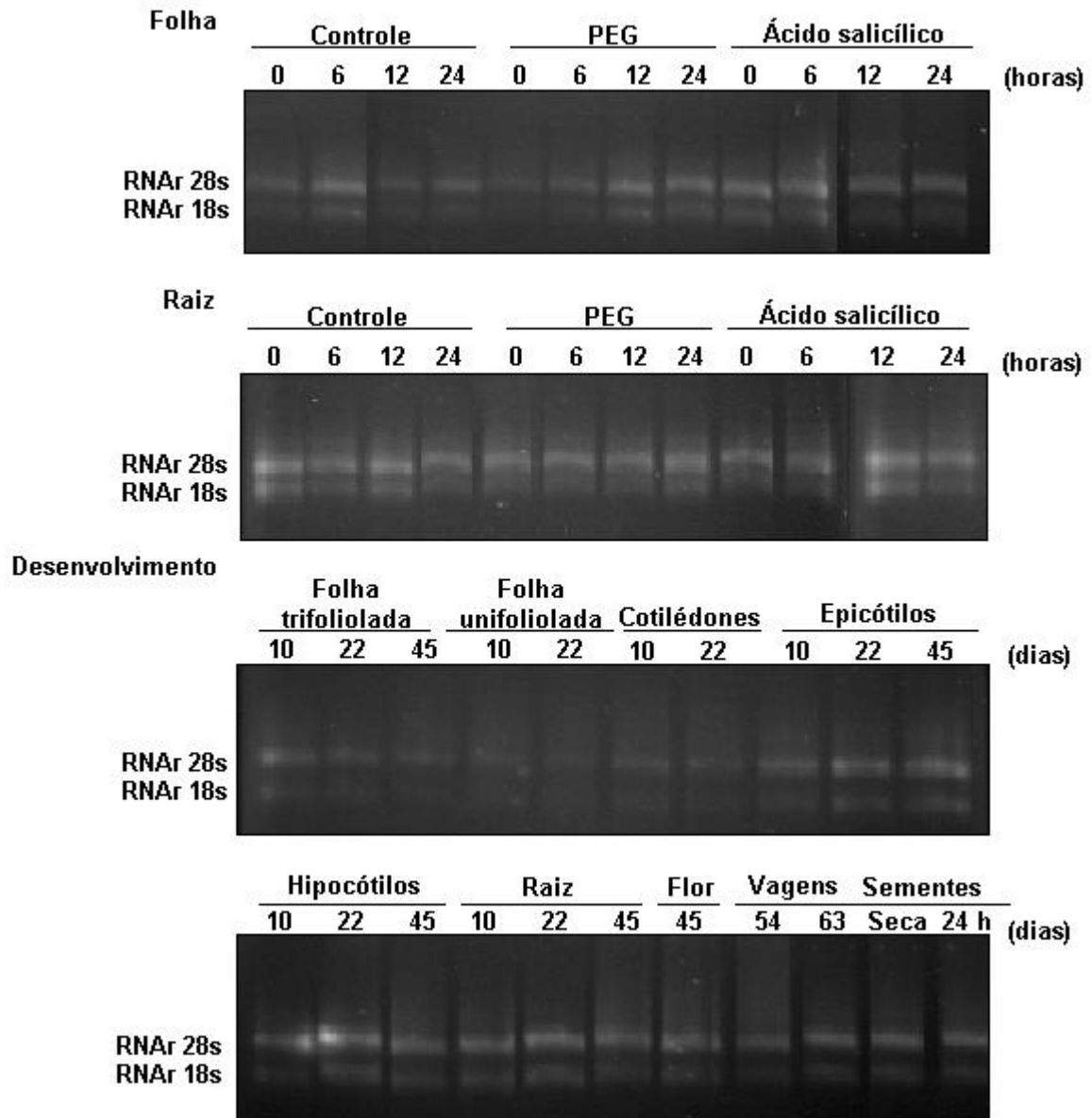
### 3.2 Qualidade do cDNA obtido pela reação de transcrição reversa

A qualidade do cDNA produzido via transcrição reversa foi verificada através da PCR semi - quantitativa, usando todas as amostras de cDNA produzido. A banda referente ao *amplicon* do gene *EF1 $\alpha$  1a*, usado para as reações, foi evidenciada nas diferentes amostras analisadas, demonstrando boa qualidade do cDNA produzido (Figura 5).

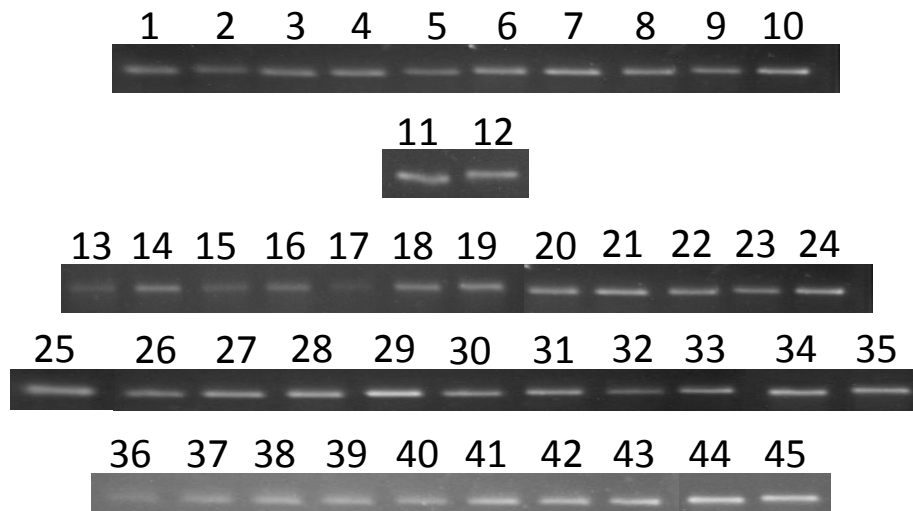
**Tabela 7:** Valores médios das concentrações e relações de absorvâncias dos RNAs extraídos em vários tecidos de soja (cultivar BRS- Pala), durante o desenvolvimento e em condições de estresses.

| <b>Tratamentos</b>           | <b>Média das concentrações (ng/μL)</b> | <b>Relação 260/280</b> | <b>Relação 260/230</b> |
|------------------------------|--|------------------------|------------------------|
| <b>Controle 0h ft</b>        | 313,8                                  | 2,13                   | 1,49                   |
| <b>Controle 6h ft</b>        | 523                                    | 2,11                   | 2,0                    |
| <b>Controle 12h ft</b>       | 361                                    | 2,0                    | 1,5                    |
| <b>Controle 24h ft</b>       | 462,3                                  | 2,15                   | 2,1                    |
| <b>Controle 0h raiz</b>      | 346,7                                  | 2,0                    | 1,62                   |
| <b>Controle 6h raiz</b>      | 587                                    | 2,09                   | 2,21                   |
| <b>Controle 12h raiz</b>     | 407,2                                  | 2,13                   | 1,65                   |
| <b>Controle 24h raiz</b>     | 592,7                                  | 2,04                   | 1,42                   |
| <b>PEG 0h ft</b>             | 426,2                                  | 2,1                    | 1,51                   |
| <b>PEG 6h ft</b>             | 569,1                                  | 2,12                   | 2,29                   |
| <b>PEG 12h ft</b>            | 589,1                                  | 2,09                   | 2,0                    |
| <b>PEG 24h ft</b>            | 551,9                                  | 2,13                   | 1,80                   |
| <b>PEG 0h raiz</b>           | 559                                    | 2,11                   | 1,23                   |
| <b>PEG 6h raiz</b>           | 369,6                                  | 2,12                   | 1,75                   |
| <b>PEG 12h raiz</b>          | 363,1                                  | 2,1                    | 2,07                   |
| <b>PEG 24h raiz</b>          | 248,2                                  | 2,11                   | 2,05                   |
| <b>AS 0h ft</b>              | 620,2                                  | 2,10                   | 2                      |
| <b>AS 6h ft</b>              | 371,9                                  | 2,10                   | 1,7                    |
| <b>AS 12h ft</b>             | 381,5                                  | 2,13                   | 2,24                   |
| <b>AS 24h ft</b>             | 341,26                                 | 2,12                   | 1,39                   |
| <b>AS 0h raiz</b>            | 449,7                                  | 2,12                   | 1,5                    |
| <b>AS 6h raiz</b>            | 610,2                                  | 2,08                   | 2,17                   |
| <b>AS 12h raiz</b>           | 682,26                                 | 2,11                   | 2,16                   |
| <b>AS 24h raiz</b>           | 417,2                                  | 2,12                   | 1,7                    |
| <b>Flores</b>                | 420,4                                  | 2,11                   | 1,53                   |
| <b>Vagens 9 dias</b>         | 1149,7                                 | 2,15                   | 2,04                   |
| <b>Vagens 18 dias</b>        | 758,0                                  | 2,12                   | 2,21                   |
| <b>Semente seca</b>          | 592,4                                  | 2,11                   | 2,3                    |
| <b>Semente embebida 24 h</b> | 238,0                                  | 2,15                   | 2,07                   |

|                                   |        |      |      |
|-----------------------------------|--------|------|------|
| <b>Cotilédones 10 dias</b>        | 841,1  | 2,14 | 2,31 |
| <b>Cotilédones 22 dias</b>        | 136,9  | 2,11 | 1,52 |
| <b>Folha trifoliolada 10 dias</b> | 783,65 | 2,1  | 2,3  |
| <b>Folha trifoliolada 22 dias</b> | 545    | 2,09 | 2,26 |
| <b>Folha trifoliolada 45 dias</b> | 246,5  | 2,13 | 2,13 |
| <b>Folha unifoliolada 10 dias</b> | 591,95 | 2,10 | 1,8  |
| <b>Folha unifoliolada 22 dias</b> | 370,2  | 2,14 | 2,06 |
| <b>Raiz 10 dias</b>               | 788,4  | 2,12 | 1,88 |
| <b>Raiz 22 dias</b>               | 529,6  | 2,11 | 1,76 |
| <b>Raiz 45 dias</b>               | 330,7  | 2,12 | 1,45 |
| <b>Hipocótilo 10 dias</b>         | 260,5  | 2,11 | 1,9  |
| <b>Hipocótilo 22 dias</b>         | 159,9  | 2,14 | 1,54 |
| <b>Hipocótilo 45 dias</b>         | 409,4  | 2,13 | 1,96 |
| <b>Epicótilo 10 dias</b>          | 817,7  | 2,13 | 1,43 |
| <b>Epicótilo 22 dias</b>          | 382,3  | 2,13 | 1    |
| <b>Epicótilo 45 dias</b>          | 578,3  | 2,11 | 1,78 |



**Figura 4:** Análise da integridade do RNA extraído em gel de agarose (1%), evidenciando as bandas do RNA ribossomal 18S e 28S. A definição das amostras está mostrada no gel.



**Figura 5:** Eletroforese em gel de agarose (2%) dos produtos de amplificação do gene do *fator de alongação 1 alfa 1a (EF1α 1a)* para avaliação da qualidade do cDNA produzido na PCR semi - quantitativa. Raias 1-12: folhas de soja de plantas controle e tratadas com polietileno glicol (PEG 100g/L) e ácido salicílico (AS 0,5mM); Raias 13-24: raízes de plantas controle e tratadas com polietileno glicol (PEG 100g/L) e ácido salicílico (AS 0,5mM); Raias 25-45: amostras de vários tecidos (Semente seca, semente embebida na solução nutritiva de Hoagland por 24 horas, raiz, folha unifoliolada, folha trifoliolada, cotilédones, hipocótilo, epicótilo, flores e vagens) da planta em vários estádios de desenvolvimento, compreendendo estádios vegetativos (10 e 22 dias após a germinação – DAG) e reprodutivos 45 DAG (início da floração), 54 DAG (nove dias após o início da floração) e 63 DAG (dezoito dias após o início da floração).

### 3.3 Escolha da melhor temperatura de anelamento

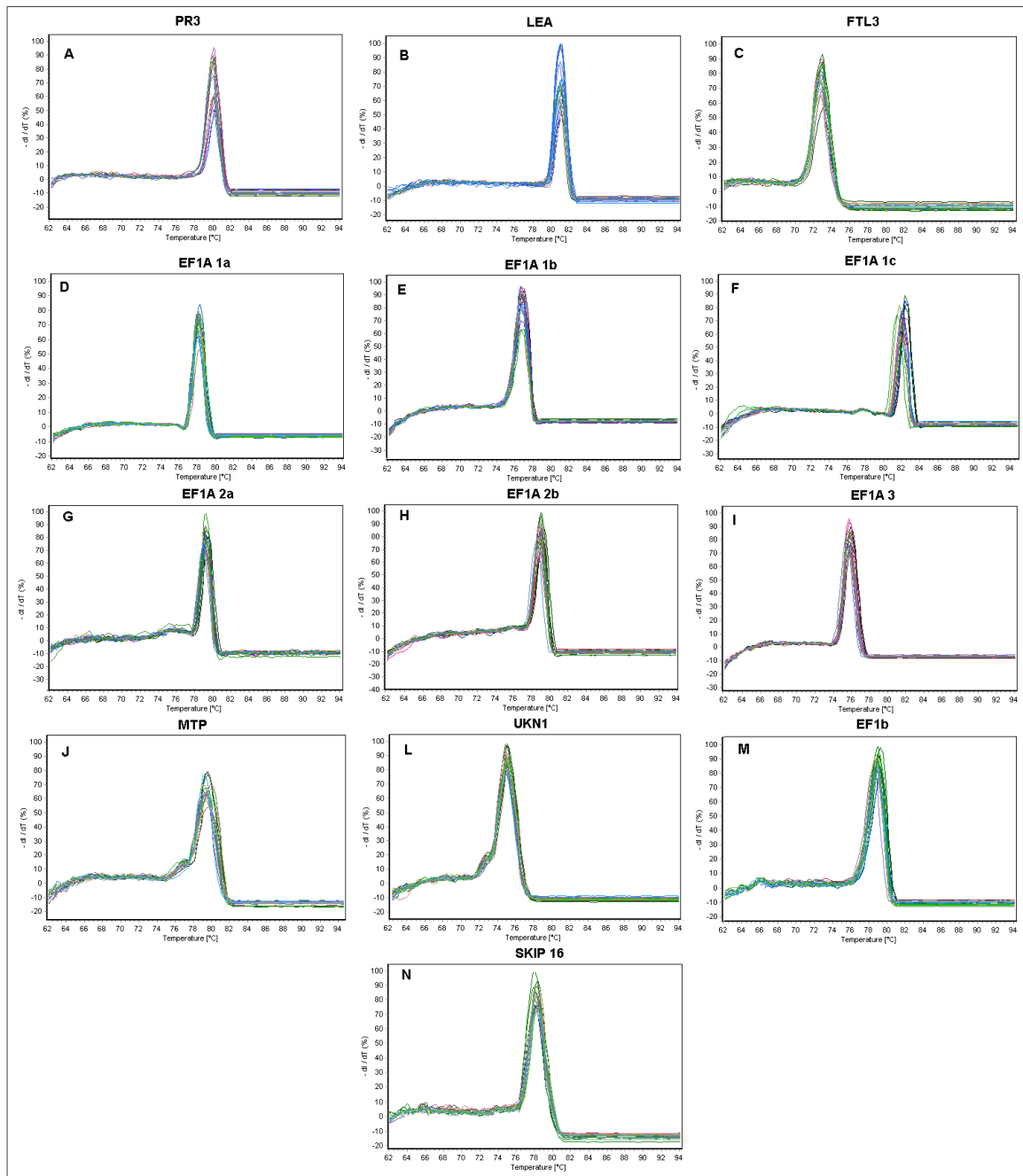
A melhor temperatura de anelamento foi escolhida com base na realização de um gradiente de temperatura. As temperaturas selecionadas para cada par de *primers* variaram. A Tabela 8 mostra as temperaturas de anelamento dos *primers* selecionadas para realização dos ensaios de qRT-PCR.

### 3.4 Especificidade e eficiência dos *primers*

Neste estudo, seis genes *EF1 $\alpha$*  nomeados anteriormente (*EF1 $\alpha$  1a, 1b, 1c, 2a, 2b e 3*) e outros quatro genes (*MTP, SKIP 16, UKNI e EF1 $\beta$* ) tiveram sua expressão avaliada por RT-qPCR em vários tecidos durante o desenvolvimento. Em condições de estresse (PEG e AS) somente os genes pertencentes à família multigênica do *EF1 $\alpha$*  foram avaliados. A especificidade dos dez pares de *primers* dos genes candidatos a normalizadores, juntamente com os *primers* dos genes *LEA, PR3 e FTL3* foi avaliada por RT-qPCR.

As análises das curvas de dissociação mostraram a presença de um único pico, evidenciando, desta forma, que um único produto foi amplificado e que não houve formação de dímeros de *primers* (Figura 6). A especificidade também foi confirmada através de eletroforese em gel de agarose 2%, onde se observou a presença de um único fragmento com o tamanho do amplicon esperado para todos os genes testados.

A eficiência de amplificação dos *primers* variou de 82 (*EF1 $\alpha$  2a*) a 129% (*EF1 $\alpha$  3*), mostrando que todos os *primers* eram adequados para ensaios por RT – qPCR (Tabela 9). Os valores de eficiência de todos os genes são mostrados na (Tabela 9).



**Figura 6:** Curvas de dissociação obtidas por RT-qPCR usando *primers* específicos. A- *PR3*: quitinase tipo 1; B- *LEA*: proteína abundante na embriogênese tardia; C- *FTL3*: proteína da floração do locus T; D- *EF1 $\alpha$  1a*: fator de alongação 1 alfa 1a; E- *EF1 $\alpha$  1b*: fator de alongação 1 alfa 1b; F- *EF1 $\alpha$  1c*: fator de alongação 1 alfa 1c; G- *EF1 $\alpha$  2a*: fator de alongação 1 alfa 2a; H- *EF1 $\alpha$  2b*: fator de alongação 1 alfa 2b; I- *EF1 $\alpha$  3*: fator de alongação 1 alfa 3; J- *MTP*: metaloprotease, enzima de degradação da insulina; L- *UKN1*: proteína hipotética; M- *EF1 $\beta$* : fator de alongação 1 beta e N- *SKIP 16*: SKP1/Ask- proteína de interação16.



**Tabela 8:** Temperaturas de anelamento selecionadas para os genes da família multigênica do *EF1 $\alpha$* , *PR3*, *LEA*, *FTL3*, *MTP*, *UKN1*, *SKIP 16* e *EF1 $\beta$* .

| Gene                                  | <i>EF1A1a</i>       | <i>EF1A1b</i>       | <i>EF1A1c</i>       | <i>EF1A2a</i>       | <i>EF1A2b</i>       | <i>EF1A3</i>        | <i>PR3</i>          | <i>LEA</i>        | <i>FTL3</i>         | <i>MTP</i>          | <i>UKN1</i>         | <i>SKIP 16</i>      | <i>EF1<math>\beta</math></i> |
|---------------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|-------------------|---------------------|---------------------|---------------------|---------------------|------------------------------|
| Temperatura de anelamento selecionada | 59,8 <sup>o</sup> C | 59,8 <sup>o</sup> C | 63,6 <sup>o</sup> C | 57,5 <sup>o</sup> C | 57,5 <sup>o</sup> C | 63,6 <sup>o</sup> C | 50,5 <sup>o</sup> C | 65 <sup>o</sup> C | 52,8 <sup>o</sup> C | 63,3 <sup>o</sup> C | 63,6 <sup>o</sup> C | 50,7 <sup>o</sup> C | 63,6 <sup>o</sup> C          |

**Tabela 9:** Cálculos da inclinação (*slope*) da curva e da eficiência de amplificação dos diferentes pares de *primers* obtidos na RT-qPCR.

| Iniciador                        | Inclinação | Eficiência (%) |
|----------------------------------|------------|----------------|
| <i>EF1<math>\alpha</math> 1a</i> | -3,01      | 1,13           |
| <i>EF1<math>\alpha</math> 2a</i> | -3,79      | 0,82           |
| <i>EF1<math>\alpha</math> 2b</i> | -3,50      | 0,95           |
| <i>EF1<math>\alpha</math> 1b</i> | -3,55      | 0,90           |
| <i>EF1<math>\alpha</math> 3</i>  | -3,79      | 1,29           |
| <i>EF1<math>\alpha</math> 1c</i> | -3,90      | 0,90           |
| <i>FTL3</i>                      | -3,70      | 0,86           |
| <i>LEA</i>                       | -3,30      | 0,99           |
| <i>PR3</i>                       | -3,34      | 0,99           |
| <i>EF1<math>\beta</math></i>     | -          | 0,94           |
| <i>SKIP 16</i>                   | -          | 1,02           |
| <i>MTP</i>                       | -          | 0,93           |
| <i>UKN1</i>                      | -          | 0,96           |

### 3.5 Determinação da estrutura e número de genes *EF1α* em espécies da ordem Fabales

Os genes *EF1α* são codificados por famílias multigênicas nas espécies da ordem Fabales (*Glycine max*, *Cajanus cajan*, *Phaseolus vulgaris*, *Vigna unguiculata* e *Medicago truncatula*) aqui analisadas. Todas essas espécies com exceção de *Vigna unguiculata* tiveram os seus genomas sequenciados e se encontram atualmente disponíveis nos bancos de dados, o que possibilitou as análises aqui realizadas. Inicialmente identificou-se nos bancos de dados às sequências relativas aos genes *EF1α* e determinou-se o número de genes em todas essas espécies analisadas.

O número de genes encontrados variou de 3 a 6 entre as diferentes espécies de leguminosas. Em soja foram identificados seis genes codificando para o *EF1α*, em *Cajanus cajan* e *Medicago truncatula* encontraram-se cinco genes e em *Phaseolus vulgaris* e *Vigna unguiculata* três genes. Em *Vigna unguiculata* tem-se apenas uma estimativa do número de genes, tendo em vista que o seu genoma não está sequenciado. A (Tabela 10) mostra o número de genes em cada espécie. Os genes *EF1α* em soja estão distribuídos em 5 diferentes cromossomos 5, 10, 16, 17 e 19 (Tabela 11).

**Tabela 10:** Caracterização e número de genes EF1 $\alpha$  em espécies da ordem Fabales.

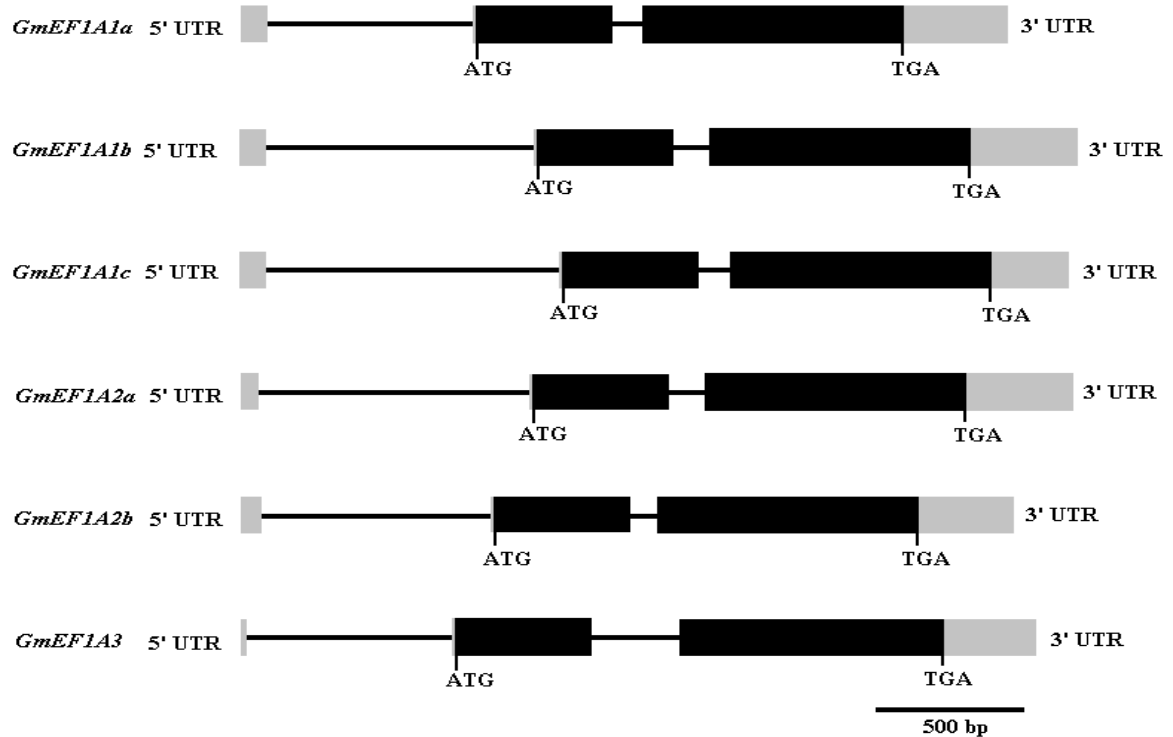
| <b>Espécies</b>            | <i>Medicago<br/>truncatula</i> | <i>Phaseolus<br/>vulgaris</i> | <i>Cajanus<br/>cajan</i> | <i>Glycine<br/>max</i> | <i>Vigna<br/>unguiculata</i> |
|----------------------------|--------------------------------|-------------------------------|--------------------------|------------------------|------------------------------|
| Número de genes            | Cinco                          | Três                          | Cinco                    | Seis                   | Três                         |
| Número de éxons e introns. | Dois/dois                      | Dois/dois                     | Dois/dois                | Dois/dois              | -                            |

**Tabela 11:** Localização genômica dos genes EF1 $\alpha$  de soja.

| <b>Gene</b>                      | <b>Localização (Cromossomo)</b> |
|----------------------------------|---------------------------------|
| <i>EF1<math>\alpha</math> 1a</i> | 5                               |
| <i>EF1<math>\alpha</math> 1b</i> | 19                              |
| <i>EF1<math>\alpha</math> 1c</i> | 16                              |
| <i>EF1<math>\alpha</math> 2a</i> | 5                               |
| <i>EF1<math>\alpha</math> 2b</i> | 17                              |
| <i>EF1<math>\alpha</math> 3</i>  | 10                              |

Quando se analisou a estrutura dos genes *EFl $\alpha$* , pode-se concluir que todas as espécies de leguminosas (excetuando *Vigna unguiculata*: sem dados genômicos) apresentavam uma estrutura gênica conservada, composta de dois éxons e dois íntrons (Tabela 10 e Figura 7). Um destes íntrons está localizado na região 5' UTR (Figura 7).

A (Figura 7) mostra a representação esquemática dos genes *EFl $\alpha$*  em Soja. Pode-se observar que os tamanhos dos éxons e íntrons são proporcionais nos seis genes.



**Figura 7:** Representação esquemática da estrutura dos seis genes *EF1α* membros da família multigênica em *Glycine max* (soja). Os genes foram anotados através de buscas em bancos de dados do NCBI (<http://www.ncbi.nlm.nih.gov/>) usando sequências homólogas do *EF1α*. Éxons e UTRs são representados pelas caixas pretas e cinza respectivamente. As linhas entre as caixas representam os íntrons.

### 3.6 Análises das sequências de cDNAs, proteínas e promotores

Os cDNAs, proteínas e promotores deduzidos de *Glycine max* foram alinhados usando o programa CLUSTAL W (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). Entre os seis genes analisados, a identidade entre as sequências de cDNAs deduzidas foi alta variando de 80 a 94% sendo a maior identidade observada entre os genes *EF1α 1a/EF1α 1b* e *EF1α 2a/EF1α 2b* (Tabela 12).

As sequências de aminoácidos deduzidas também mostraram alta identidade entre os genes *EF1α* analisados (80 a 99%). A maior identidade observada foi entre os genes *EF1α 1a/EF1α 1b* e *EF1α 2a/EF1α 2b* (Tabela 13), seguindo perfil semelhante ao observado para as sequências de cDNAs deduzidas. Em adição, as sequências de cDNAs e aminoácidos deduzidos do *EF1α* de *Glycine max* foram comparadas com as sequências deduzidas de outras espécies de leguminosas (*Cajanus cajan*, *Phaseolus vulgaris*, *Vigna unguiculata* e *Medicago truncatula*). Nesta análise, observou-se alta identidade entre as sequências de cDNAs (63 a 93%) e de aminoácidos deduzidos (77 a 98%) (dados não mostrados).

As sequências de promotores (1000 pb) dos seis genes *EF1α* de soja também foram estudadas. As análises mostraram alta divergência na região dos promotores, no entanto os genes *EF1α 1a/1b* (53%) e os genes *EF1α 2a/2b* (51%) revelaram as maiores identidades (Tabela 14).

**Tabela 12:** Identidade das sequências de cDNAs deduzidas dos genes *EF1 $\alpha$*  em soja (*Glycine max*) obtidos do alinhamento no Clustal W.

|                                     | <i>Gm EF1<math>\alpha</math> 1a</i> | <i>Gm EF1<math>\alpha</math> 2a</i> | <i>Gm EF1<math>\alpha</math> 3</i> | <i>Gm EF1<math>\alpha</math> 1c</i> | <i>Gm EF1<math>\alpha</math> 2b</i> | <i>Gm EF1<math>\alpha</math> 1b</i> |
|-------------------------------------|-------------------------------------|-------------------------------------|------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| <i>Gm EF1<math>\alpha</math> 1a</i> | -                                   | 80%                                 | 83%                                | 86%                                 | 83%                                 | 94%                                 |
| <i>Gm EF1<math>\alpha</math> 2a</i> | 80%                                 | -                                   | 86%                                | 81%                                 | 93%                                 | 81%                                 |
| <i>Gm EF1<math>\alpha</math> 3</i>  | 83%                                 | 86%                                 | -                                  | 82%                                 | 86%                                 | 83%                                 |
| <i>Gm EF1<math>\alpha</math> 1c</i> | 86%                                 | 81%                                 | 86%                                | -                                   | 81%                                 | 86%                                 |
| <i>Gm EF1<math>\alpha</math> 2b</i> | 83%                                 | 93%                                 | 86%                                | 81%                                 | -                                   | 82%                                 |
| <i>Gm EF1<math>\alpha</math> 1b</i> | 94%                                 | 81%                                 | 82%                                | 86%                                 | 82%                                 | -                                   |

**Tabela 13:** Identidade das sequências de aminoácidos deduzidas dos genes *EF1 $\alpha$*  em soja (*Glycine max*) obtidos do alinhamento no Clustal W.

|                                     | <i>Gm EF1<math>\alpha</math> 1a</i> | <i>Gm EF1<math>\alpha</math> 2a</i> | <i>Gm EF1<math>\alpha</math> 3</i> | <i>Gm EF1<math>\alpha</math> 1c</i> | <i>Gm EF1<math>\alpha</math> 2b</i> | <i>Gm EF1<math>\alpha</math> 1b</i> |
|-------------------------------------|-------------------------------------|-------------------------------------|------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| <i>Gm EF1<math>\alpha</math> 1a</i> | -                                   | 96%                                 | 97%                                | 97%                                 | 96%                                 | 99%                                 |
| <i>Gm EF1<math>\alpha</math> 2a</i> | 96%                                 | -                                   | 96%                                | 97%                                 | 99%                                 | 96%                                 |
| <i>Gm EF1<math>\alpha</math> 3</i>  | 97%                                 | 96%                                 | -                                  | 97%                                 | 97%                                 | 97%                                 |
| <i>Gm EF1<math>\alpha</math> 1c</i> | 97%                                 | 97%                                 | 97%                                | -                                   | 96%                                 | 97%                                 |
| <i>Gm EF1<math>\alpha</math> 2b</i> | 96%                                 | 99%                                 | 97%                                | 96%                                 | -                                   | 96%                                 |
| <i>Gm EF1<math>\alpha</math> 1b</i> | 99%                                 | 96%                                 | 97%                                | 97%                                 | 96%                                 | -                                   |

**Tabela 14:** Identidade dos promotores dos genes *EF1 $\alpha$*  em soja (*Glycine max*) obtidos do alinhamento no Clustal W.

|                                     | <i>Gm EF1<math>\alpha</math> 1a</i> | <i>Gm EF1<math>\alpha</math> 2a</i> | <i>Gm EF1<math>\alpha</math> 3</i> | <i>Gm EF1<math>\alpha</math> 1c</i> | <i>Gm EF1<math>\alpha</math> 2b</i> | <i>Gm EF1<math>\alpha</math> 1b</i> |
|-------------------------------------|-------------------------------------|-------------------------------------|------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| <i>Gm EF1<math>\alpha</math> 1a</i> | -                                   | 41%                                 | 37%                                | 42%                                 | 43%                                 | 53%                                 |
| <i>Gm EF1<math>\alpha</math> 2a</i> | 41%                                 | -                                   | 36%                                | 40%                                 | 51%                                 | 41%                                 |
| <i>Gm EF1<math>\alpha</math> 3</i>  | 37%                                 | 36%                                 | -                                  | 39%                                 | 40%                                 | 39%                                 |
| <i>Gm EF1<math>\alpha</math> 1c</i> | 42%                                 | 40%                                 | 39%                                | -                                   | 43%                                 | 42%                                 |
| <i>Gm EF1<math>\alpha</math> 2b</i> | 43%                                 | 51%                                 | 40%                                | 43%                                 | -                                   | 44%                                 |
| <i>Gm EF1<math>\alpha</math> 1b</i> | 53%                                 | 41%                                 | 39%                                | 42%                                 | 44%                                 | -                                   |

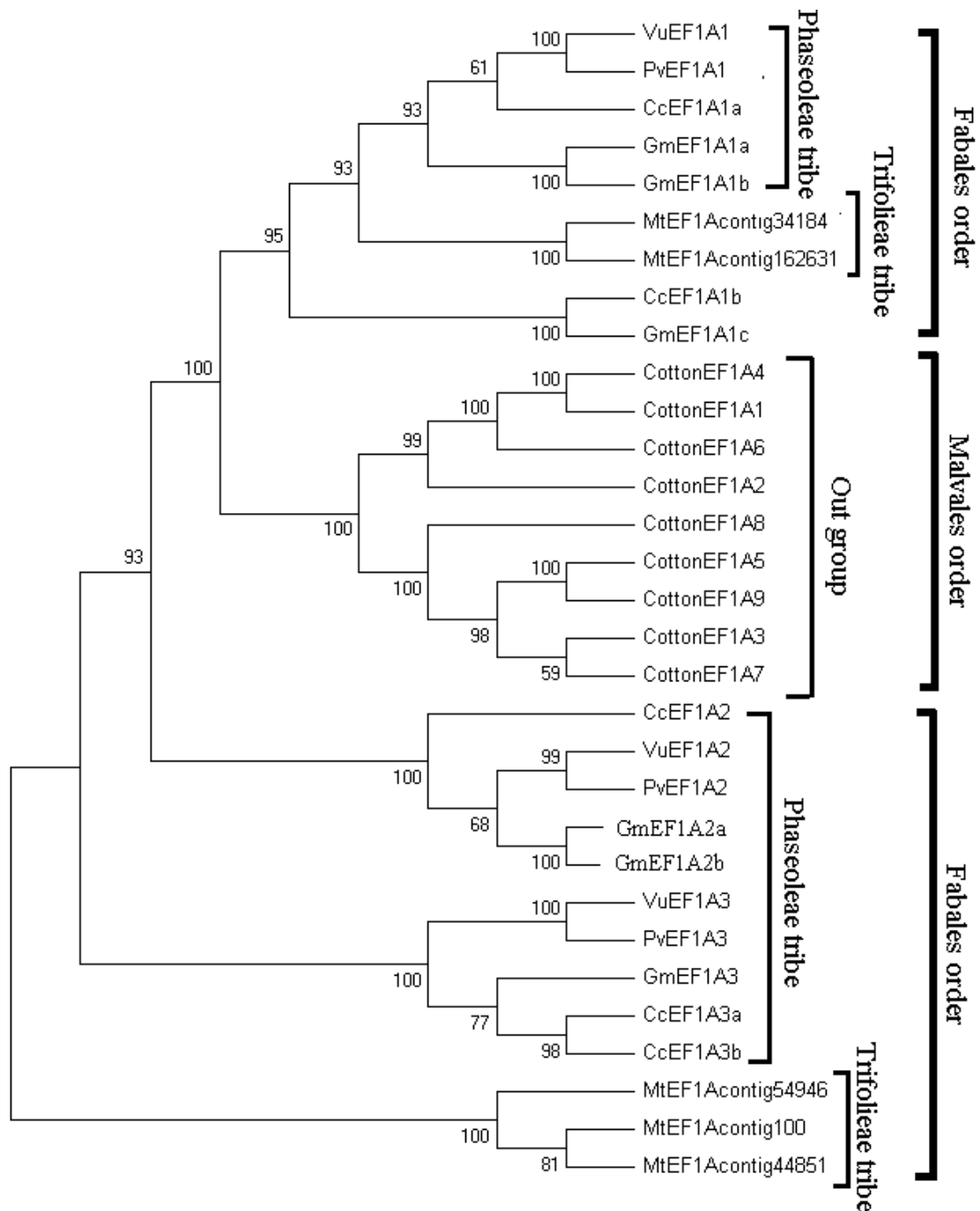


### 3.7 Análises das relações filogenéticas entre os genes *EF1α* de leguminosas

A análise filogenética dos genes *EF1α* das espécies de leguminosas da ordem Fabales (*Medicago truncatula*, *Phaseolus vulgaris*, *Cajanus cajan*, *Vigna unguiculata* e *Glycine max*), foi feita através do programa Mega 5.05. Os genes *EF1α* de *Gossypium hirsutum* (XU et al, 2007), da ordem Malvales foram usados como grupo externo.

A análise demonstrou que os genes *EF1α* das espécies da ordem Fabales, tribo Phaseoleae (*Phaseolus vulgaris*, *Cajanus cajan*, *Vigna unguiculata* e *Glycine max*) são divididos em três clados (Figura 5) de genes ortólogos, e são divergentes dos de *Medicago truncatula* (ordem Fabales, tribo Trifolieae) e *Gossypium hirsutum* (ordem malvales).

Após as análises filogenéticas propusemos uma classificação para os genes *EF1α* da ordem Fabales, tribo phaseoleae (Figura 8). Não foi possível estabelecer a nomenclatura para os genes *EF1α* de *Medicago truncatula*, visto que os genes ortólogos entre esta espécie e as demais analisadas neste trabalho não foram identificados.



**Figura 8:** Análise filogenética dos genes *EF1α* (cDNA deduzidos) das fabales (*Mr: Medicago truncatula*, *Pv: Phaseolus vulgaris*, *Cc: Cajanus cajan*, *Vu: Vigna unguiculata* e *Gm: Glycine max*) feita através do uso da ferramenta MEGA 5.05 empregando o método neighbor joining (NJ). Os genes *EF1α* de *Gossypium hirsutum* (ordem malvales) foram utilizados como grupo externo.

### 3.8 Perfil de expressão dos genes *EF1α* em *Glycine max*

Os genes membros da família multigênica do *EF1α* foram investigados quanto ao perfil de expressão nos diferentes tecidos (flores, vagens, sementes, cotilédones, folhas unifolioladas e trifolioladas, raízes, hipocótilos e epicótilos) durante o desenvolvimento da soja. Para as análises de expressão gênica através de qPCR os genes *MTP*, *UKNI*, *EF1β* e *SKIP 16*, validados em estudos prévios realizados por Hu et al., (2009) foram utilizados como controles internos para normalizar as possíveis diferenças nas concentrações iniciais de transcritos.

A análise de expressão por qPCR revelou que todos os genes *EF1α* de soja são funcionais, sendo bem expressos em todos os tecidos, contudo apresentam expressão diferencial dependente do tecido e do estágio de desenvolvimento (Figura 6).

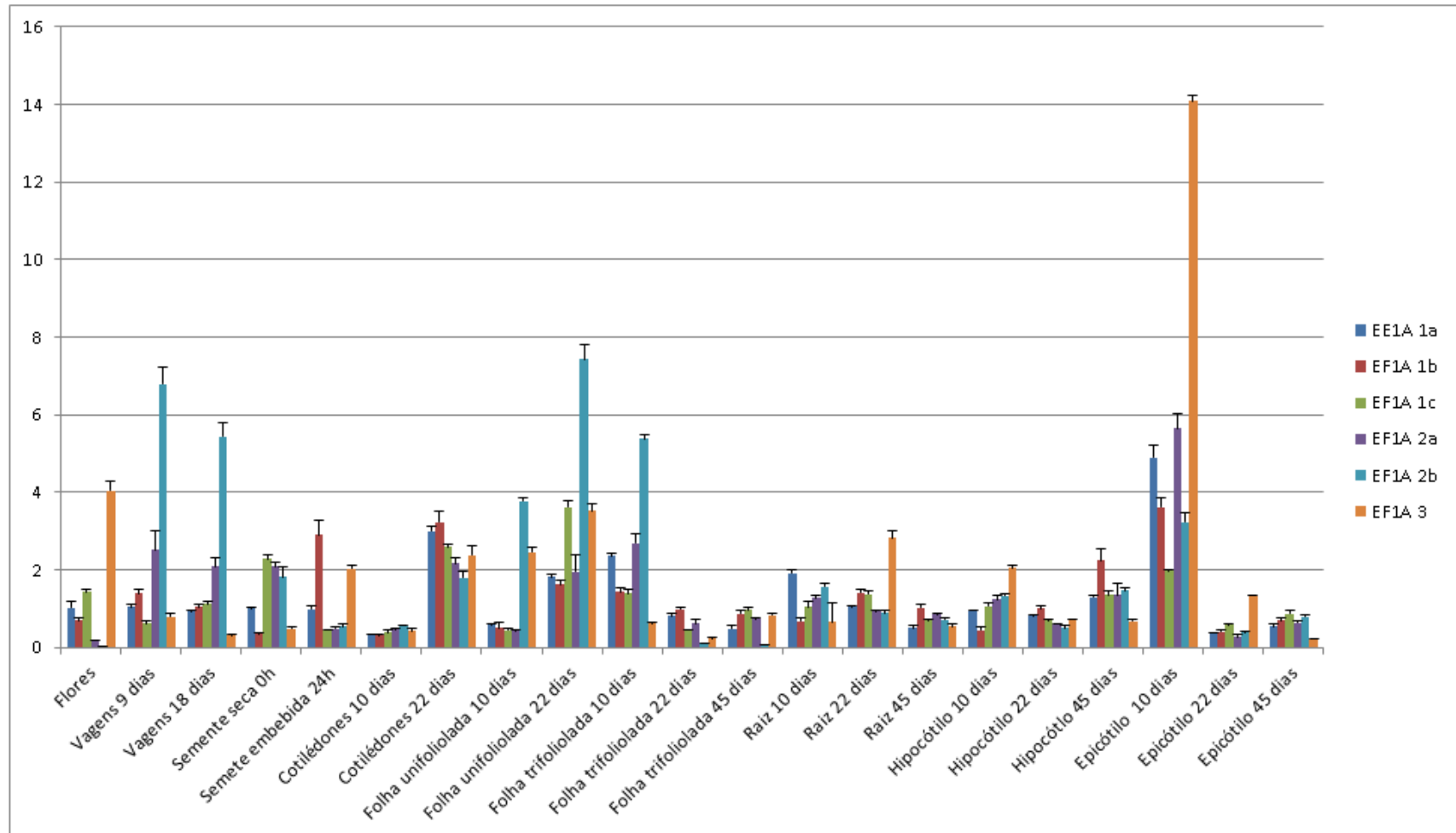
A Figura 9 mostra que o gene *EF1α 1a* foi o gene membro que apresentou a menor variação de expressão nos diferentes tecidos testados. Já o gene *EF1α 1b* foi mais expresso em semente embebida (germinação) e durante o desenvolvimento do hipocótilo, visto que, os níveis de transcritos, neste tecido, aumentam gradativamente no decorrer do desenvolvimento da planta.

A análise do perfil de expressão do gene *EF1α 1c* mostrou que há um predomínio de expressão deste gene durante o desenvolvimento da semente (vagem), tendo em vista que há um aumento de expressão gradativo desde vagens com nove dias até semente seca, no entanto há uma queda dos níveis de transcritos em sementes embebidas. Esse mesmo perfil é observado para os genes *EF1α 2a* e *2b*. Para os genes *EF1α 2a* e *2b* observa-se também um aumento dos níveis de transcritos em folha unifoliolada com vinte e dois dias. Menores níveis de transcritos foram detectados para os genes *EF1α 2a* e *2b* em folhas trifolioladas no decorrer do desenvolvimento.

O gene *EF1α 3* foi o que apresentou o perfil de expressão mais variável. Os maiores níveis de transcritos foram observados em tecidos com dez dias, tais como hipocótilos e epicótilos. Observa-se também um aumento de expressão em folhas unifoliolada com vinte e dois dias, raízes com vinte e dois dias e cotilédones com vinte e dois dias. Altos níveis de transcritos também foram visualizados em amostras de flores e sementes em germinação.

Em amostras de folhas trifolioladas e epicótilos (dez dias) todos os genes *EFl $\alpha$*  apresentam-se mais expressos. Já em amostras de cotilédones e folhas unifolioladas com vinte e dois dias, observa-se que todos os genes tiveram suas expressões aumentadas em relação a cotilédones e folhas unifolioladas com dez dias, respectivamente. Vale salientar que ambos os tecidos coletados com vinte e dois dias estavam em processo de senescência.

Expressão relativa



**Figura 9:** Expressão relativa dos diferentes membros da família gênica do *EF1a* em soja.

## 4 DISCUSSÃO

O fator de alongação ( $EF1\alpha$ ) auxilia a ligação do aminoacil – tRNA ao sítio A do ribossomo durante a fase de alongação da síntese proteica em um mecanismo dependente de GTP, desta forma é considerado um gene constitutivo “Housekeeping gene”, pois participa do metabolismo celular básico e é requerido para a sobrevivência da célula (XU et al, 2007). Genes constitutivos são expressos em todos os tecidos/órgãos da planta (RANSON-HODGKINS, et al., 2009).

No presente estudo, os genes  $EF1\alpha$  de soja foram identificados, caracterizados e analisados filogeneticamente em associação com os  $EF1\alpha$  de outras leguminosas (*Vigna unguiculata*, *Phaseolus vulgaris*, *Medicago truncatula* e *Cajanus cajan*). Além disso, os genes  $EF1\alpha$  tiveram seus padrões de expressão elucidados em diferentes tecidos durante o desenvolvimento da soja.

Muitos estudos têm mostrado que o  $EF1\alpha$  é codificado por famílias multigênicas em plantas (RANSON-HODGKINS et al., 2009), variando de poucos a um grande número de cópias. O grande número de membros nas famílias gênicas pode estar relacionado com a função primária do  $EF1\alpha$  na síntese proteica, a fim de assegurar a fidelidade do processo. Além disso, pode também estar relacionado com a necessidade de adaptação das plantas a um determinado estímulo ambiental.

Em soja, os genes  $EF1\alpha$ , como mostrado na (Figura 7) pertencem a uma família multigênica de seis membros, distribuídos em cinco diferentes cromossomos (5, 10, 16, 17 e 19) e que codificam para proteínas variando em identidade de 80 a 99%. Alta identidade entre as proteínas deduzidas também foi observada em *Arabidopsis* (RANSON-HODGKINS, 2009 e AXELOS et al, 1989;) (99% a 100%) e arroz (100%) (KIDOU e EJIRI et al., 1998). Os genes  $EF1\alpha$  das espécies de leguminosas aqui analisadas, também mostraram alta identidade entre as sequências de aminoácidos e nucleotídeos deduzidas. Isto pode ser explicado devido à proximidade filogenética entre essas espécies.

As análises filogenéticas mostraram que os genes  $EF1\alpha$  de (*Vigna unguiculata*, *Phaseolus vulgaris*, *Medicago truncatula* e *Cajanus cajan*) (ordem fabales, tribo Phaseoleae) são divididos em três clados (Figura 5) de genes ortólogos, e são divergentes dos de *Medicago truncatula* (ordem fabales, tribo trifolieae). Após essa análise, os genes  $EF1\alpha$  de soja foram nomeados como  $EF1\alpha$  1a, 1b, 1c, 2a, 2b e 3, de acordo com a distribuição filogenética,

estabelecendo-se também a nomenclatura dos genes ortólogos aos da soja em algumas espécies da ordem fabales (*Cajanus cajan*, *Vigna unguiculata*, *Phaseolus vulgaris*) (Figura 8). No caso de *Medicago truncatula*, não foi possível estabelecer essa nomenclatura. Isso foi devido ao fato dos genes *EF1 $\alpha$*  serem caracterizados por apresentarem alta identidade entre os parálogos (mesma espécie), fator este que, aparentemente, interfere na caracterização de genes ortólogos (espécies mais distantes filogeneticamente) possibilitando uma identificação apenas em espécies mais próximas. Nesse trabalho, identificamos que o limite para a identificação de genes ortólogos *EF1 $\alpha$*  na ordem Fabales não deve se estender além do nível taxonômico “tribo”.

Em soja, observou-se o maior número de genes *EF1 $\alpha$*  entre as espécies de leguminosas analisadas (que variou de 3 a 6). Este fato pode estar relacionado com eventos de duplicação do genoma da soja (SCHMUTZ et al., 2010). Em *Vigna unguiculata* 3 genes foram identificados. Para esta espécie temos apenas uma estimativa, visto que o seu genoma ainda não está disponível. No entanto, há fortes indícios em afirmar o número de genes *EF1 $\alpha$*  nesta espécie, pois *Phaseolus vulgaris* e *Vigna unguiculata* são espécies muito próximas filogeneticamente, e *Phaseolus vulgaris*, tem seu genoma sequenciado, onde são encontrados os três genes ortólogos aos de *Vigna unguiculata*.

O número de genes *EF1 $\alpha$*  membros da família multigênica nas diferentes espécies é variável. Em espécies monocotiledôneas como milho, são encontrados entre dez a quinze genes (CARNEIRO et al, 1999) e em cana-de-açúcar vinte genes (VIJAYKUMAR et al, 2002) são reportados. Em plantas dicotiledôneas, como tomate, são encontrados entre quatro a oito genes (POKALSKY et al., 1989) em algodão e batata, nove genes (XU et al., 2007 e COSTA et al., 2010) e em mandioca (SUHANDONO et al., 2001) encontram-se seis genes. Em *Arabidopsis* e arroz, duas plantas modelos, são encontrados quatro genes *EF1 $\alpha$* . Em todas estas espécies, os autores mostram expressão diferencial dos genes *EF1 $\alpha$*  nas diferentes condições experimentais.

Neste trabalho, foi observado que os genes *EF1 $\alpha$*  membros da família multigênica em soja, são diferencialmente expressos em diferentes tecidos/órgãos durante o desenvolvimento da planta. Analisando a (Figura 9) pode-se distinguir dois perfis de expressão: Um perfil geral (tendência de expressão de todos os genes) e um perfil específico (expressão de cada gene isoladamente).

Fazendo uma análise do perfil geral de expressão dos genes *EF1 $\alpha$*  observa-se que eles são genes detectados em todos os tecidos e condições estudadas de soja. Além disso, como era de se esperar, tais genes são muito expressos em tecidos jovens onde há uma alta taxa de divisão celular (RANSON-HODGKINS., et al 2009) como: epicótilo (dez dias em relação a vinte e dois e quarenta e cinco dias) e folha trifoliolada (dez dias em relação a vinte e dois e quarenta e cinco dias) (Figura 9). Entretanto, curiosamente, foi também observado que os genes *EF1 $\alpha$*  são bem mais expressos em cotilédones e folhas unifolioladas com vinte e dois dias em relação aos mesmos tecidos coletados com dez dias. Esse achado sugere uma função dos genes *EF1 $\alpha$*  na senescência desses tecidos, provavelmente na síntese de proteínas relacionadas com a morte celular programada. Sabe-se que em eventos de PCD há a superexpressão de proteínas pré apoptóticas (PETROS et al., 2004).

Analisando-se o perfil específico de cada gene observa-se predomínio de expressão em alguns tecidos. Pela análise da (Figura 9), pode-se observar que o gene *EF1 $\alpha$  3* é o que apresenta maiores variações e predomínio de expressão em vários tecidos, tais como: epicótilos e hipocótilos de plantas jovens, sugerindo que este gene tem um papel crucial na síntese proteica em epicótilos e hipocótilos no início do desenvolvimento da planta. Além disso, predomínio de expressão do *EF1 $\alpha$  3* foi encontrado em sementes em germinação, flores e raízes (vinte e dois dias). O gene *EF1 $\alpha$  1a* foi o que apresentou a menor variação entre os seis genes *EF1 $\alpha$*  de soja. Os genes *EF1 $\alpha$  1c*, *2a* e *2b* de soja apresentaram predomínio de expressão durante o desenvolvimento da vagem. O *EF1 $\alpha$  2a* também apresentou predomínio de expressão em folhas unifolioladas e trifolioladas e o *EF1 $\alpha$  2b* em folhas unifolioladas. Essa expressão diferencial de alguns membros observada em soja também foi observada em outras espécies de plantas. Em *Arabidopsis thaliana*, a família gênica do *EF1 $\alpha$*  é composta por quatro genes, designados *EF1 $\alpha$  A1*, *A2*, *A3* e *A4*. Os genes nomeados *EF1 $\alpha$  A3* e *A4* são bem expressos em flores, já o gene *EF1 $\alpha$  A2* é altamente expresso em raízes e em amostras de folhas jovens o gene *EF1 $\alpha$  A4* é o mais expresso (RANSON-HODGKINS., et al 2009).

Observando o perfil geral de expressão dos quatro genes *EF1 $\alpha$*  de *Arabidopsis* (RANSON-HODGKINS., et al 2009), verifica-se que são expressos diferencialmente nos diferentes tecidos, sendo os maiores níveis de transcritos detectados em amostras de sementes durante a germinação e meristemas de raízes. A menor expressão destes genes foi observada em mudas em desenvolvimento com quatorze dias após a semeadura. A alta expressão dos genes *EF1 $\alpha$*  nos tecidos jovens, esta relacionada, segundo os autores, com a alta taxa de



divisão celular destes tecidos, sendo os genes *EF1 $\alpha$*  requeridos para a síntese proteica das células em divisão (RANSON-HODGKINS, et al 2009).

Em Arroz a família multigênica do *EF1 $\alpha$*  é também composta por quatro genes, que codificam para proteínas altamente idênticas (100% de identidade), nomeados como *refa1*, *refa2*, *refa3* and *refa4* (KIDOU e EJIRI, 1998). Estes genes são também diferencialmente expressos durante o desenvolvimento do arroz. O maior o nível de transcritos foi observado em mudas de arroz com cinco dias após a semeadura e endospermas de sementes imaturas. Menores níveis de RNAm foram detectados em sementes secas e embebidas com água por vinte e quatro horas para induzir a germinação. Tais resultados corroboram com os encontrados nesse trabalho, pois os genes *EF1 $\alpha$  1c*, *2a* e *2b* também diminuem seus níveis de transcritos em sementes durante a germinação.

Em tomate, altos níveis de transcritos foram também encontrados em tecidos durante o desenvolvimento como folhas jovens e frutos verdes (POKALSKY et al., 1989). Em mandioca, altos níveis de mRNA também são detectados em tecidos jovens (SUHANDONO et al., 2001). Estes resultados corroboram com os encontrados para soja, onde se observa maior nível de transcritos em folhas trifolioladas e epicótilos jovens.

Quando se observou o perfil de expressão dos genes *EF1 $\alpha$*  em soja, os genes *EF1 $\alpha$  2a* e *2b* apresentaram um perfil de expressão semelhante, excetuando-se as amostras de folhas unifolioladas e trifolioladas (Figura 9). Embora ambos os genes estejam em cromossomos diferentes (*EF1 $\alpha$  2a* – Cromossomo 5; *EF1 $\alpha$  2b* – Cromossomo 17), os referidos promotores apresentam extensas regiões com alta identidade sugerindo que esses genes compartilham fatores de transcrição semelhantes corroborando com a co-expressão encontrada.

Os resultados encontrados aqui reforçam a necessidade de estudos de expressão com os diferentes membros das famílias gênicas, tendo em vista que, esse padrão de expressão diferencial pode estar relacionado com necessidades específicas de plantas em diferentes tecidos durante o desenvolvimento. Além disso, estes estudos podem servir de base para a seleção de genes alvo que desempenhem um papel fundamental durante o ciclo de vida da planta.

## 5 CONCLUSÕES

1. Os genes *EF1α* em soja são membros de uma família multigênica composta de seis genes classificados como *EF1α 1a; 1b; 1c; 2a; 2b* e *3*. Este número de genes é variável entre as diferentes espécies de leguminosas aqui analisadas. O maior número de genes em soja pode estar relacionado com eventos de duplicações recentes no genoma.
2. Todos os genes *EF1α* em soja são funcionais. O perfil geral de expressão revela alta expressão de todos os genes em tecidos de plantas jovens (dez dias) como epicótilos e folha trifoliolada. Ademais, todos os genes também são bem expressos em tecidos senescentes como cotilédones e folha unifoliolada (vinte e dois dias) sugerindo, pela primeira vez, um possível papel do *EF1α* na morte celular programada.
3. Geralmente há predomínio de expressão tempo/tecido-específica de genes membros:

*EF1α 1a*: Não há predomínio;

*EF1α 1b*: Germinação e desenvolvimento do hipocótilo;

*EF1α 1c*: Desenvolvimento da vagem;

*EF1α 2a*: Desenvolvimento da vagem, folha unifoliolada e trifoliolada;

*EF1α 2b*: Desenvolvimento da vagem, folha unifoliolada;

*EF1α 3*: Germinação, flores, raízes, fases iniciais do desenvolvimento do hipocótilo e epicótilo.

## 6 REFERÊNCIAS BIBLIOGRÁFICAS

AXELOS M, LIBOZ T, BARDET C, et al. The gene family encoding the Arabidopsis thaliana translation elongation factor-1A: molecular cloning, characterization and expression. *Molecular and General Genetics*. 1989 219:106–112.

CARNEIRO NP, HUGHES PA, LARKINS BA. The eEF1A gene family is differentially expressed in maize endosperm. *Plant Industry* 1999; 41: 801-813.

COSTA AM, DI GIACOMO I, MASSARELLI, et al. Isolation, characterization and expression of an elongation factor 1 $\alpha$  gene in potato (*Solanum tuberosum*) cell cultures. *Plant Biosystems* 2010; 144: 618-625.

EMBRAPA. Cultivares de soja, regiões norte e nordeste do brasil. Embrapa soja, Londrina, PR, 2012. Disponível em <[http://www.cnpso.embrapa.br/index.php?op\\_page=294&cod\\_pai=17](http://www.cnpso.embrapa.br/index.php?op_page=294&cod_pai=17)>. Acesso em 02/08/2012.

EMBRAPA SOJA: <http://www.cnpso.embrapa.br/> acesso em 01/03/2012

HOAGLAND D, ARNON DI. The water culture method for growing plants without soil. *California Agriculture Experimental Station Circular* 1950. 347 p.

HU R, FAN C, LI H, et al. Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. *BMC Molecular Biology* 2009; 10: 93.

KIDOU S, EJIRI S. Isolation, characterization and mRNA expression of four cDNAs encoding translation elongation factor 1A from rice (*Oryza sativa* L.). *Plant Mol. Biol* 1998; 36: 137-148.

LIVAK KJ, SCHMITTEN TD. Analysis of gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 2001; 25: 402-408.

MARSHALL O. Per primer v 1. 1. 20. 2003-2011.

PETRUS MJ, WILHELM DE, MURAKAMI M, et al. Altered expression of Chk1 disrupts cell cycle remodeling at the midblastula transition in *Xenopus laevis* embryos. *Cell Cycle* 2004; 3: 212-216.

POKALSKY AR, HIATT WR, RIDGE N, et al. Structure and expression of an elongation factor 1A in tomato. *Nucl Acids Res* 1989 17: 4661-4673.

RANSOM-HODGKINS WD. The application of expression analysis in elucidating the eukaryotic elongation factor one alpha gene family in *Arabidopsis thaliana*. *Mol Genet Genomics* 2009; 281:391-405.

RNeasy® Mini Handbook. Quarta edição. Setembro 2010.

SAITOU N, NEI M. The Neighbor-Joining method-a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987; 4: 406-25.

SCHMUTZ J, CANNON SB, SCHLUETER J, et al. Genome sequence of the palaeopolyploid soybean. *Nature* 2010; 463: 178-83.

SUHANDONO S, HUGHES J, BROWN K, et al. Expression and structure of an elongation factor-1 $\alpha$  gene (MeEF1) from cassava (*Manihot esculenta* Crantz). *Euphytica* 2001; 120: 49-58.

TAMURA K, PETERSON D, PETERSON N, et al. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* 28: 2731-2739.

VANDESOMPELE J, DE PRETER K, PATTYN F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002; 3: 18.

VIJAYKUMAR, D. Molecular cloning, characterization and tissue specific expression of an elongation factor 1A gene in *Saccharum officinarum* L. *Plant Sci* 2002; 162: 315–321.

XU W-L, WANG X-L, WANG H, et al. Molecular characterization and expression analysis of nine cotton GhEF1A genes encoding translation elongation factor 1A. *Gene* 2007; 389: 27-35.

## **CAPÍTULO II**

**ANÁLISE DOS GENES DA FAMÍLIA MULTIGÊNICA DO EF1 $\alpha$  DE SOJA COMO NORMALIZADORES PARA qPCR NOS DIFERENTES TECIDOS EM DESENVOLVIMENTO E CONDIÇÕES DE ESTRESSE**

# 1 OBJETIVOS

## 1.1 Objetivo geral

Selecionar o(s) gene(s) de referência mais estáveis, pertencentes à família multigênica do fator de alongação 1 alfa (*EFl $\alpha$* ), para análises de expressão gênica por RT- qPCR em soja (*Glycine max*) em diferentes tecidos durante o desenvolvimento e sob condições de estresse.

## 1.2 Objetivos específicos

- Determinar os melhores genes de referência da família multigênica (*EFl $\alpha$* ) em diferentes tecidos durante o desenvolvimento da soja usando os programas GeNorm e NormFinder;
- Selecionar os genes *EFl $\alpha$*  de expressão mais estável, em diferentes condições de estresse (Polietileno glicol - PEG, Ácido salicílico – AS) usando os programas GeNorm e NormFinder;;
- Validar a utilização dos genes de referência selecionados.

## 2 MATERIAIS E MÉTODOS

### 2.1 Dados de expressão dos genes *EF1α*

A seleção de genes normalizadores *EF1α* de soja foi feita a partir de dados de expressão dos referidos genes em diferentes tecidos durante o desenvolvimento e sob condições de estresse. Os dados de expressão de tecidos em desenvolvimento foram obtidos conforme descrito no capítulo I, enquanto que os dados de expressão sob condições de estresse foram obtidos como segue.

### 2.2 Condições de crescimento e desenho experimental

Para os ensaios em condições de estresse, plantas de soja (BRS PALA) com treze dias após a semeadura (DAS) foram submetidas a estresse osmótico, promovido pelo polietileno glicol (PEG - 100 g/L), e estresse através da aplicação exógena de ácido salicílico (AS - 0,5 mM). Folhas e raízes foram coletadas nos tempos 0, 6, 12 e 24 h de submissão aos estresses, as amostras foram congeladas imediatamente em nitrogênio líquido e armazenadas no freezer -80 °C para posterior extração de RNA total.

### 2.3 Extração de RNA total

A extração do RNA, determinação de sua pureza e integridade, bem como a síntese de cDNA foi feita como descrito no capítulo anterior.

### 2.4 Ensaios de PCR em tempo real

Os ensaios de PCR em tempo real, determinação da eficiência e especificidade dos *primers* foram executados como descrito no capítulo I.

### 2.5 Desenho dos *primers* para as reações de qPCR

Nesse capítulo foram utilizados os *primers* *FTL3*, *LEA* e *PR3*, juntamente com os demais *primers* mencionados anteriormente (*EF1α 1a, 1b, 1c, 2a, 2b e 3/ MTP, EF1β, SKIP 16 e UKN1*). Os *primers* foram desenhados de acordo com os parâmetros descritos no capítulo I e são mostrados na Tabela 15.

**Tabela 15:** Pares de *primers* utilizados nos ensaios de qRT-PCR

| Gene                         | Cromossomo | Produto gênico                                | Sequências dos <i>primers</i> (5'-3') | Temperatura de anelamento <sup>o</sup> C | Eficiência de amplificação (%) | Tamanho do amplicon (pb) |                 |
|------------------------------|------------|---|---------------------------------------|--|--------------------------------|--------------------------|-----------------|
| <i>EF1a1a</i>                | 05         | Fator de alongação 1 $\alpha$                 | Fwd 5' GATTCATGTAGCCGTAGCC 3'         | 59.8                                     | 113                            | 182                      |                 |
|                              |            |   | Rev 5' ATTTAAGACATCCCTCCTCAG 3'       |  |                                |                          |                 |
| <i>EF1a2a</i>                | 05         | Fator de alongação 1 $\alpha$                 | Fwd 5' GGATGTCGTTTCTTATGGT 3'         | 57.5                                     | 82                             | 189                      |                 |
|                              |            |   | Rev 5' CAAACACAACACATTTAAACAG 3'      |  |                                |                          |                 |
| <i>EF1a1c</i>                | 16         | Fator de alongação 1 $\alpha$                 | Fwd 5' ATCGCATCATAGTTACATTATAG 3'     | 52.9                                     | 90                             | 198                      |                 |
|                              |            |   | Rev 5' AGTTCTCATACAGCTTATAAAATAG 3'   |  |                                |                          |                 |
| <i>EF1a1b</i>                | 19         | Fator de alongação 1 $\alpha$                 | Fwd 5' ATCATCGTGGTTACTCCTTTAT 3'      | 59.8                                     | 95                             | 199                      |                 |
|                              |            |   | Rev 5' TCAGACTCTTCTTACCATCA 3'        |  |                                |                          |                 |
| <i>EF1a2b</i>                | 17         | Fator de alongação 1 $\alpha$                 | Fwd 5' TTTCTGACTCTTGTGCTTCT 3'        | 57.5                                     | 95                             | 159                      |                 |
|                              |            |   | Rev 5' ACAAACACAACACATTTAAACAC 3'     |  |                                |                          |                 |
| <i>EF1a3</i>                 | 10         | Fator de alongação 1 $\alpha$                 | Fwd 5' GTGCTGGATGGTTTATCAAGG 3'       | 63.6                                     | 129                            | 140                      |                 |
|                              |            |   | Rev 5' AAAGTTGCGAGATGACGGAG 3'        |  |                                |                          |                 |
| <i>FTL3</i>                  |            | Proteína de floração do locus T               | Fwd 5' GGTTCGGTGAAGGAGG 3'            | 52.8                                     | 86                             | 123                      |                 |
|                              |            |   | Rev 5' CTAATAAAGAGTGTGGGA 3'          |  |                                |                          |                 |
| <i>LEA</i>                   |            | Proteína abundante na embriogenesis           | Fwd 5' GCCGAAGTGAAGAAAAGACGAACC 3'    | 65                                       | 99                             | 124                      |                 |
|                              |            |   | Rev 5' CTTGGGCTGTTTGTGGGTCTTC 3'      |  |                                |                          |                 |
| <i>PR3</i>                   |            | Quitinase tipo 1                              | Fwd 5' ACCCTCCTCTGCTTTCTC 3'          | 55                                       | 99                             | 130                      |                 |
|                              |            |   | Rev 5' GCACTTGGTCTGGATTG 3'           |  |                                |                          |                 |
| <i>EF1<math>\beta</math></i> |            | Fator de alongação 1 $\beta$                  | Fwd 5' CCACTGCTGAAGAAGATGATGATG 3'    | 63.6                                     | 94                             | 134                      | HU et al., 2009 |
|                              |            |   | Rev 5' AAGGACAGAAGACTTGCCACTC 3'      |  |                                |                          |                 |
| <i>SKIP 16</i>               |            | SKP1/Ask-Proteína de interação 16             | Fwd 5' CCACTGCTGAAGAAGATGATGATG 3'    | 50.7                                     | 102                            | 60                       | HU et al., 2009 |
|                              |            |   | Rev 5' AAGGACAGAAGACTTGCCACTC 3'      |  |                                |                          |                 |
| <i>MTP</i>                   |            | Metaloprotease, enzima de degradação insulina | Fwd 5' CGTCCAAGTGCTCCTCATTAG 3'       | 63.3                                     | 93                             | 71                       | HU et al., 2009 |
|                              |            |   | Rev 5' TGAAGTAACCGACGCCAACG 3'        |  |                                |                          |                 |
| <i>UKN1</i>                  |            | Proteína hipotética                           | Fwd 5' TGGTGCTGCCGCTATTTACTG 3'       | 63.6                                     | 96                             | 74                       | HU et al., 2009 |
|                              |            |   | Rev 5' GGTGAAGGAACTGCTAACAAATC 3'     |  |                                |                          |                 |



## 2.6 Seleção dos genes candidatos a normalizadores

Em soja a família multigênica do *EF1 $\alpha$*  é composta por seis genes. Esses genes candidatos a normalizadores tiveram sua estabilidade de expressão testada em duas condições de estresse, descritas anteriormente e em diferentes tecidos durante o desenvolvimento da soja (Tabela 4).

Nas análises durante o desenvolvimento, os genes *MTP*, *EF1 $\beta$* , *SKIP 16* e *UKNI* também foram testados a fim de se estabelecer uma comparação com os genes *EF1 $\alpha$* , membros da família multigênica. Os referidos genes *MTP*, *EF1 $\beta$* , *SKIP 16* e *UKNI* foram analisados por HU et al, (2009) e são indicados para normalizar experimentos com soja durante o desenvolvimento.

### 2.6.1 Análise da estabilidade de expressão dos genes candidatos

Os níveis de expressão gênica foram determinados pelo número de ciclos de amplificação necessários para a fluorescência (emitida pelo SYBR Green) ultrapassar um limiar durante a fase exponencial da reação de PCR. Esse limiar é determinado pelo *cycle threshold* ( $C_t$ ).

Os valores brutos de  $C_t$  foram exportados do programa Realplex para uma planilha do excel e posteriormente importados para o programa GeNorm (Ver 2.4) e NormFinder (Ver. 0.953) de acordo com as especificações de cada programa. Estes programas são utilizados para calcular a estabilidade de expressão dos genes candidatos a normalizadores. O GeNorm calcula a estabilidade dos genes candidatos através de um valor  $M$  e o número de genes necessários para a normalização em diferentes condições através da variação da média em pares (valores  $V$ ) (VANDESOMPELE et al., 2002). O NormFinder considera a variação intra e intergrupar dos candidatos a genes de referência, bem como indica a melhor combinação (pares) de genes que devem ser utilizados em cada condição (ANDERSEN et al., 2004). O valor de estabilidade gerado no programa é utilizado para identificar os melhores genes, sendo indicados aqueles com os menores valores.

### **2.6.2 Validação dos genes normalizadores durante o desenvolvimento e nas condições de estresse**

Para validar os resultados das análises de estabilidade de expressão dos genes constitutivos e determinar como a adoção de um, ou múltiplos genes pode afetar a normalização de dados de expressão relativa, os genes *LEA*, *PR3* e *FTL3* foram utilizados. Os genes *LEA* e *PR3* foram também utilizados como controles positivos, tendo em vista que, estes genes respondem aos estresses com PEG e AS respectivamente. O gene *FTL3* foi utilizado para validar os resultados durante o desenvolvimento.

Os dados de expressão destes genes foram normalizados utilizando os valores de quantificação relativa dos genes mais estáveis indicados pelo GeNorm. Esse procedimento foi realizado no GeNorm qbase<sup>(PLUS)</sup> usando o método  $2^{-\Delta\Delta ct}$  (LIVAK & SCHMITTGEN 2001).

### 3 RESULTADOS

#### 3.5 Perfil de Expressão dos genes candidatos a normalizadores

Ensaio por RT- qPCR, com base na detecção por SYBR Green, foram utilizados para avaliar os perfis de expressão dos seis candidatos a genes de referência (*EF1 $\alpha$  1a*, *EF1 $\alpha$  1b*, *EF1 $\alpha$  1c*, *EF1 $\alpha$  2a*, *EF1 $\alpha$  2b*, *EF1 $\alpha$  3*) em vinte e quatro amostras de *G. max* em condições de estresse. O padrão de expressão também foi avaliado em dez genes candidatos a normalizadores (*EF1 $\alpha$  1a*, *1b*, *1c*, *2a*, *2b*, *3*, *MTP*, *EF1 $\beta$* , *SKIP 16* e *UKNI*) em diferentes tecidos durante o desenvolvimento da soja.

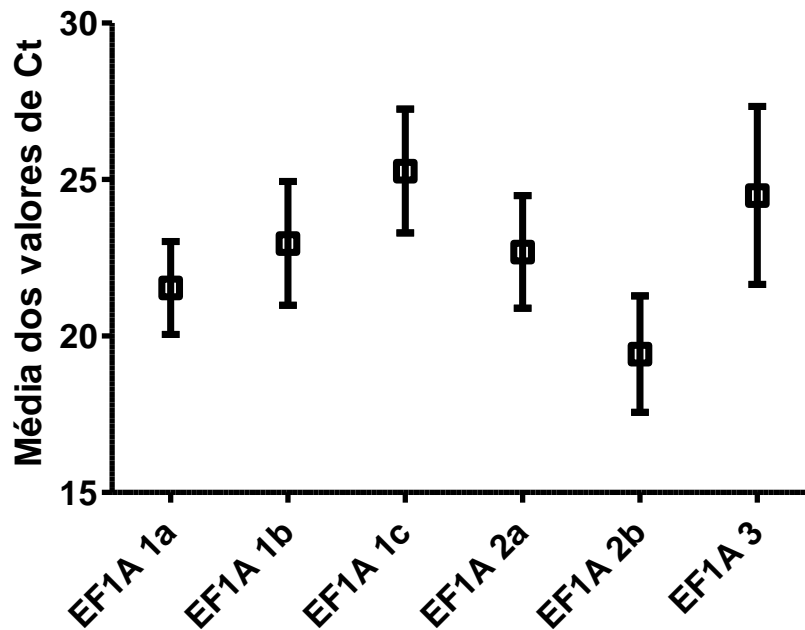
A fim de minimizar variabilidades nas reações de RT - qPCR, todas as amostras de RNA foram ajustadas para uma mesma concentração e posteriormente convertidas em cDNA via transcrição reversa.

O nível de expressão dos genes candidatos obtidos a partir de experimentos de RT- qPCR são apresentados como média dos valores de Ct. A média dos valores de Ct dos genes de referência em condições de estresse variou de 19,42 para 25,27 (Figura 10) e de 20,21 para 25,56 em diferentes tecidos durante o desenvolvimento da soja (Figura 11).

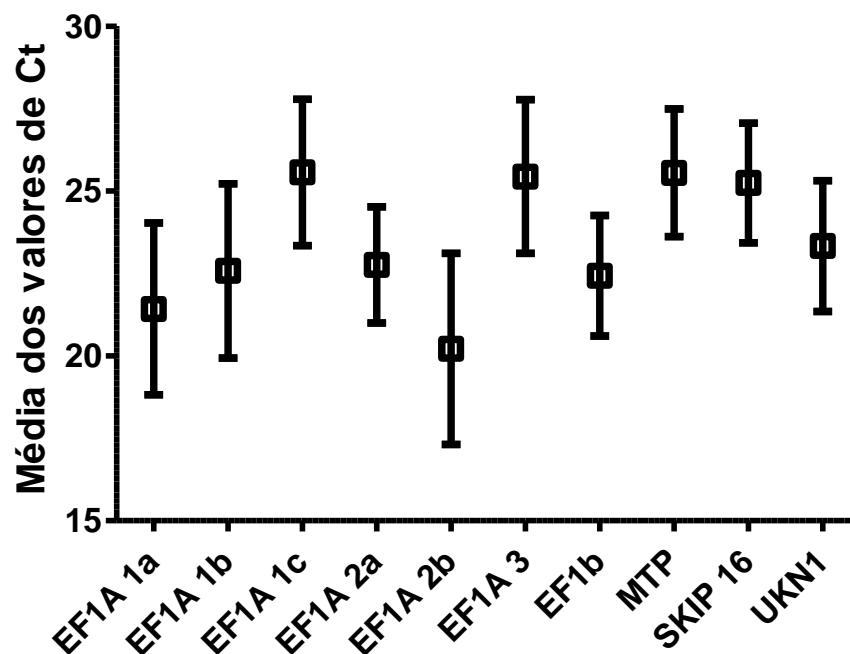
O gene *EF1 $\alpha$  2b* foi o mais expresso (média do Ct = 19.42), enquanto que o *EF1 $\alpha$  1c* foi o menos expresso (média do Ct = 25.27) nas condições de estresse (Figura 10).

Durante o desenvolvimento o gene *EF1 $\alpha$  2b* foi o mais expresso (média do Ct = 20.21) e o gene *EF1 $\alpha$  1c* foi o menos expresso (média do Ct = 25.56) (Figura 11).

Pela análise dos resultados pode-se concluir que o *EF1 $\alpha$  2a* e o *EF1 $\alpha$  2b* foram os genes que mostraram a maior e a menor variação, respectivamente, entre as amostras de desenvolvimento (Figura 11). Já nas amostras em condições de estresses o *EF1 $\alpha$  3* e *EF1 $\alpha$  1a* foram os genes que apresentaram a maior e a menor variação, respectivamente. (Figura 10).



**Figura 10:** Média dos valores de Ct em vinte e quatro amostras de plantas submetidas a condições de estresse com ácido salicílico e polietileno glicol. As caixas representam as médias de Ct e as barras os desvios padrões para cada gene.



**Figura 11:** Média dos valores de Ct em amostras de diferentes tecidos durante o desenvolvimento da soja. As caixas representam as médias de Ct e as barras os desvios padrões para cada gene.

### 3.6 Análises da estabilidade de expressão dos genes candidatos a normalizadores

A fim de verificar o nível de expressão de genes alvo em diferentes tecidos durante o desenvolvimento das plantas e em condições de estresse é crucial normalizar todas as amostras usando um gene de referência de expressão estável.

Para determinar bons genes de referência nestas condições os genes *EF1 $\alpha$*  codificados por uma família multigênica em soja foram avaliados, juntamente com quatro (*MTP*, *UKNI*, *EF1 $\beta$*  e *SKIP 16*) genes já selecionados como bons genes normalizadores (Hu et al., 2009) durante o desenvolvimento da soja. Estes genes (*MTP*, *UKNI*, *EF1 $\beta$*  e *SKIP 16*) foram avaliados por Hu et al., (2009) durante o desenvolvimento da soja em folhas trifolioladas e unifolioladas, mas não foram avaliados em diferentes tecidos durante o desenvolvimento da soja.

A estabilidade de expressão dos genes candidatos foi determinada através de dois programas estatísticos (GeNorm e NormFinder).

#### 3.6.1 Análises pelo GeNorm

A média de estabilidade de expressão gênica (valor M) de todos os genes foi calculada pelo GeNorm. Este programa recomenda a utilização de um valor M abaixo do limite de 1,5 identificando (conjuntos de) genes de referência com expressão estável.

Os resultados mostraram que no conjunto total de amostras de vários tecidos durante o desenvolvimento (Figura 12a) somente o gene *EF1 $\alpha$  2b* ficou acima do limite de corte de 1,5 recomendado pelo GeNorm. A análise individual dos diferentes tecidos mostrou que todos os genes ficaram com um valor M abaixo de 1,5, exceto nas amostras de “vagens/flores” e “sementes”, cujos genes *EF1 $\alpha$  2b* e *EF1 $\alpha$  1b*, respectivamente, ficaram acima do limite de corte (>1,5).

A análise conjunta das amostras “Total” mostrou que os genes *SKIP 16*, *UKNI* e *MTP* foram os genes de referência mais estáveis ao passo que o gene *EF1 $\alpha$  2b* foi o gene de referência menos estável (Figura 12a).

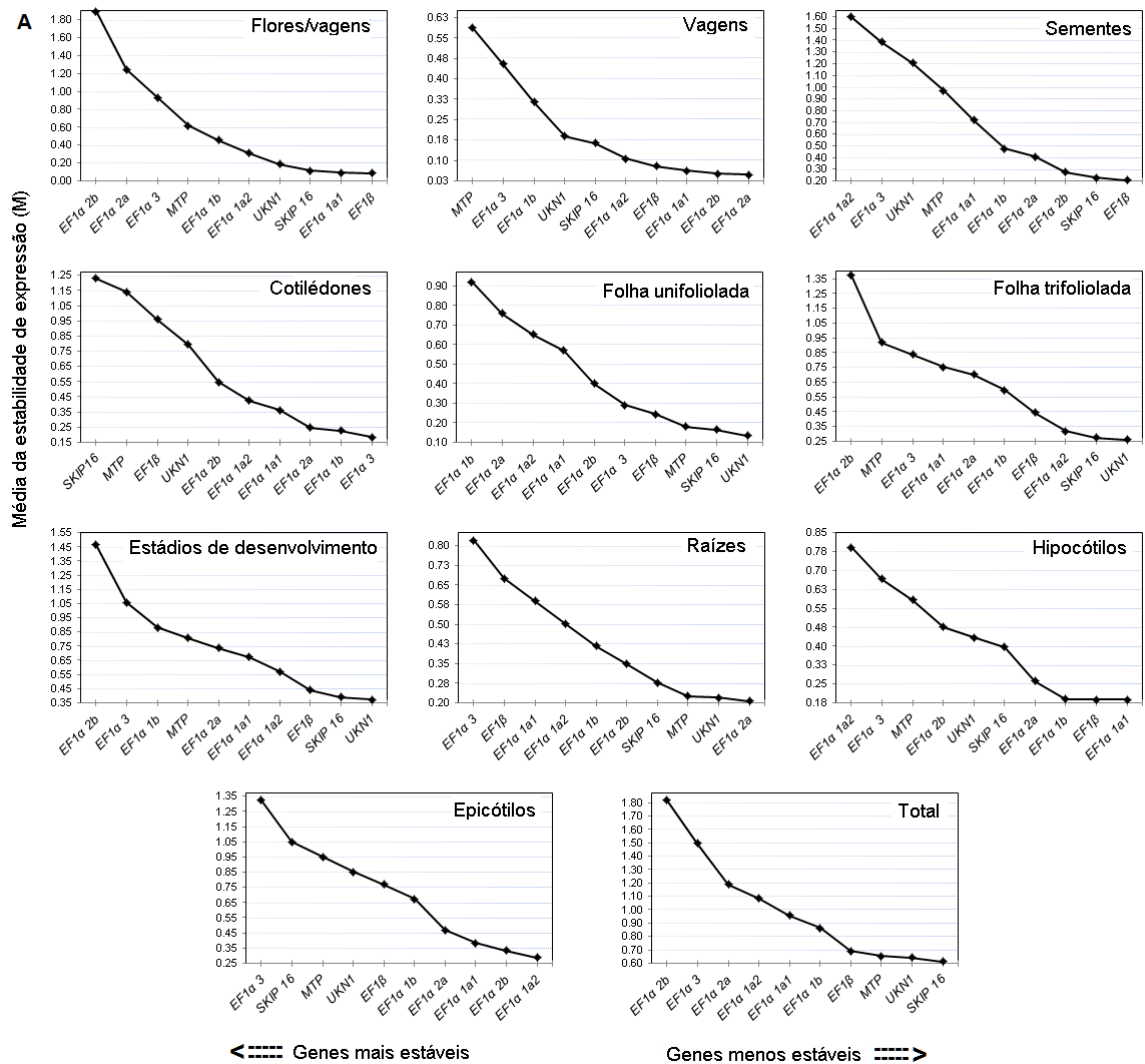
Quando as análises foram executadas em cada tecido de forma individualizada, os genes *EF1α* foram os mais estáveis para a maioria dos tecidos. Em “cotilédones” os três genes mais estáveis foram (*EF1α 3*, *1c* e *2a*) e o menos estável foi o *SKIP 16*. Em “epicótilos” os genes *EF1α 1b*, *2b* e *1a* foram os mais estáveis, ao passo que o *EF1α 3* foi o menos estável (Figura 12a).

Nas análises das amostras de “hipocótilos” os genes *EF1α 1a*, *EF1β* e *EF1α 1b* foram os mais estáveis, ficando o gene *EF1β* como o de expressão mais variável. Em folhas unifolioladas e trifolioladas (“estádios de desenvolvimento”) os genes *UKNI*, *SKIP 16* e *EF1β* foram os que menos variaram, enquanto que o *EF1α 2b* foi o que mais variou. Em “raízes” os genes *EF1α 2a*, *UKNI* e *MTP* foram os de expressão mais estável, ao passo que o gene *EF1α 3* aparece como o menos estável. Em “sementes” os genes *EF1β*, *SKIP 16* e *EF1α 2b* apresentaram o menor valor M, enquanto que o gene *EF1α 1b* foi o que apresentou o maior valor M.

Em “vagens”, os genes com os menores valores M foram *EF1α 2a*, *2b* e *1a*, ficando o gene *MTP* como o menos estável neste órgão. Analisamos também “vagens/flores”, onde os genes com os menores valores M foram *EF1β*, *EF1α 1a* e *SKIP 16* enquanto o gene *EF1α 2b* foi o mais variável.

As análises em amostras de folhas trifolioladas e unifolioladas, mostraram que em folhas trifolioladas os melhores genes foram *UKNI*, *SKIP 16* e *EF1α 1b*, ao passo que o gene *EF1α 2b* foi o menos estável. Já para folhas unifolioladas os genes *UKNI*, *SKIP 16* e *MTP* foram os mais estáveis, ficando o gene *EF1α 1c* com o maior valor M e conseqüentemente com a menor estabilidade de expressão (Figura 12a).

O programa GeNorm também estima o número de genes de referência adequados para normalização através das variações aos pares. Neste estudo verificou-se que para cada tecido/órgão analisado foram recomendadas combinações de genes diferentes (Figura 12b). O maior número de genes requeridos para normalização foi nas amostras de folhas unifolioladas e trifolioladas analisadas em conjunto (“estádios de desenvolvimento”), onde cinco genes foram necessários (*UKNI*, *SKIP 16*, *EF1β*, *EF1α 1b* e *EF1α 1a*) (Figura 12b).



**Figura 12a:** Classificação fornecida pelo GeNorm dos candidatos a genes de referência em diferentes tecidos durante o desenvolvimento da soja (*Glycine max*).

Na maioria dos tecidos a adoção de dois genes foi suficiente para a normalização, tais como: cotilédones (*EF1α 3* e *1c*), hipocótilos (*EF1α 1a* e *EF1β*), folhas trifolioladas e folhas unifolioladas (*UKNI* e *SKIP 16*), raízes (*EF1α 2a* e *UKNI*), vagens (*EF1α 2a* e *2b*), sementes (*EF1β* e *SKIP 16*), vagens/flores (*EF1β* e *EF1α 1a*). Nas amostras de epicótilos (*EF1α 1b*, *2b* e *1a*) e em todas as amostras analisadas em conjunto “Total” (*SKIP 16*, *UKNI* e *MTP*) a adoção da combinação de três genes foi à opção mais adequada (Figura 12b).

Em praticamente todos os grupos experimentais (Figura 12b) encontra-se pelo menos uma combinação de genes onde a variação V fica abaixo do valor de corte (0,15) sugerido pela literatura (VANDESOMPELE et al., 2002). Somente no conjunto “Total” de amostras a variação V ficou um pouco acima do valor recomendado (0,156), no entanto este é um valor que pode ser usado sem maiores problemas, tendo em vista que este valor é muito próximo do sugerido pelo GeNorm (Figura 12b).

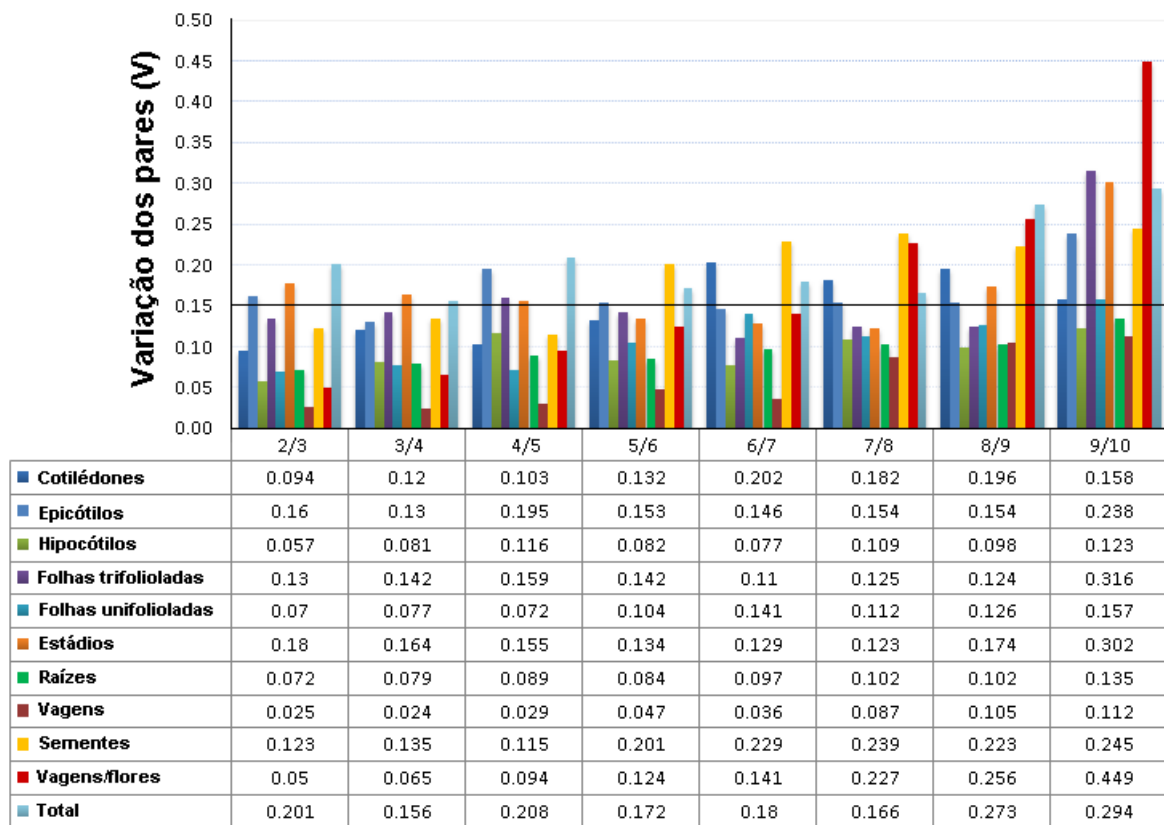
Além dos ensaios realizados durante o desenvolvimento, também analisamos a estabilidade de expressão dos genes *EF1α* em folhas e raízes em condições de estresse (PEG e AS). Os resultados obtidos através de análise no GeNorm mostraram que somente nas amostras de raízes submetidas ao estresse com AS o valor M ficou abaixo do limite de corte de 1,5 para todos os genes *EF1α* testados. Nas demais condições testadas (raízes PEG, folhas AS e PEG) o valor M ficou acima do ponto de corte para alguns dos genes *EF1α* testados, no entanto em todas as condições, pelo menos três dos seis genes *EF1α* testados ficaram abaixo do ponto de corte. Já em raízes submetidas ao estresse com PEG somente o gene *EF1α 3* ficou acima do ponto de corte de 1,5. Os genes *EF1α 2a*, *1c* e *2b* são os de expressão mais estável nestas condições (Figura 13a). Em raízes submetidas ao estresse com AS os genes mais estáveis foram *EF1α 1c*, *3* e *2a*.

Em folhas submetidas ao estresse com AS os genes *EF1α 2a*, *1b* e *2b* são os mais estáveis, o gene *EF1α 1c* foi o que mais variou. Já em folhas submetidas ao estresse com PEG, os genes *EF1α 2b*, *2a* e *1a* foram os três mais estáveis e o gene *EF1α 3* foi o de expressão menos estável (Figura 13a).

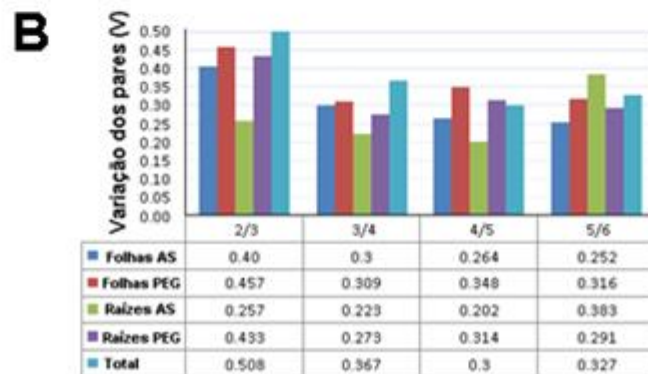
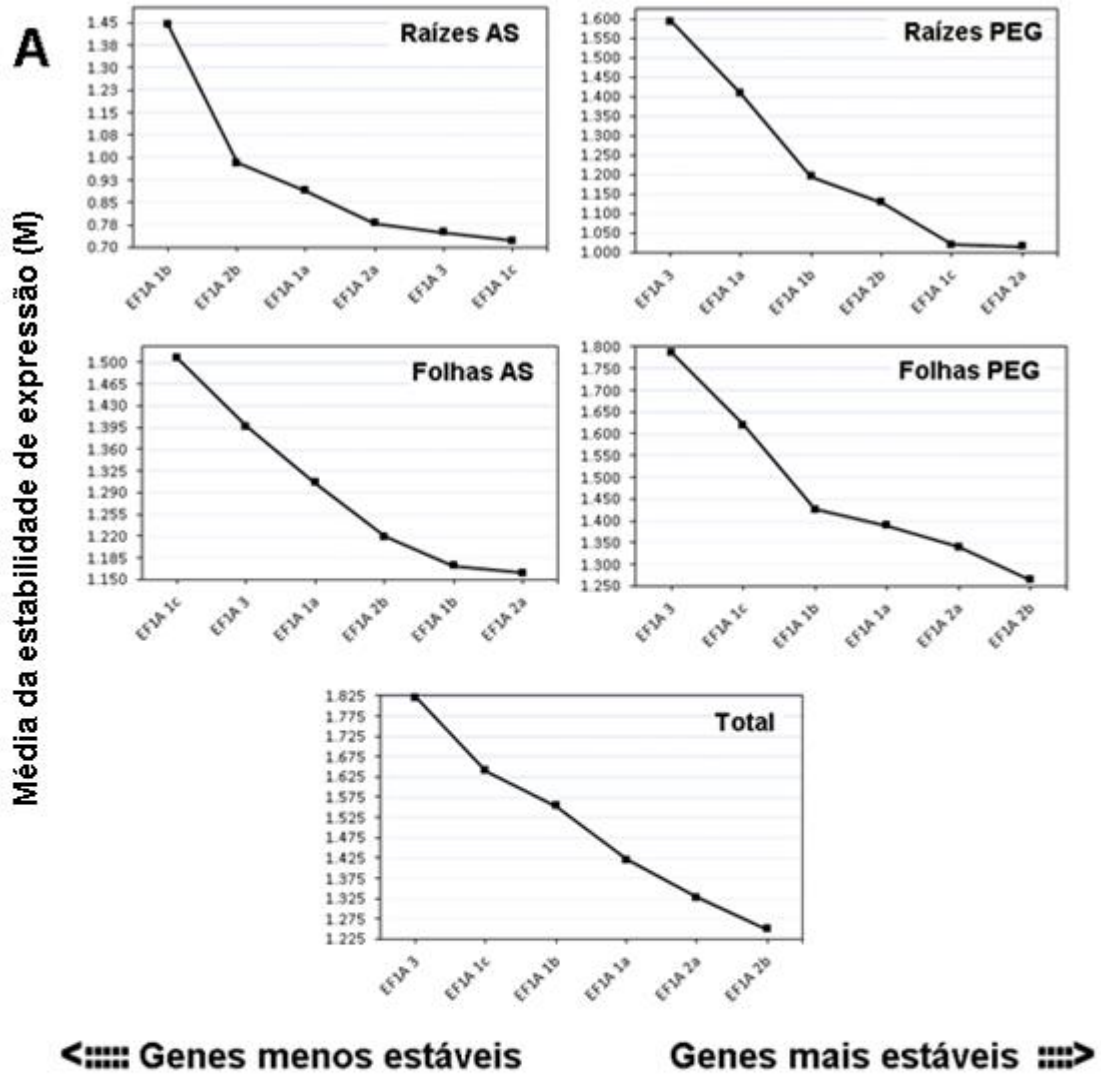
Quando analisamos as amostras submetidas aos estresses em conjunto “Total” os genes *EF1α 2b*, *2a* e *1a* são os mais indicados e o gene *EF1α 3* foi o menos indicado (Figura 13a).



Após a análise da variação dos pares (V) o GeNorm recomenda, para todos os grupos experimentais das amostras submetidas a estresses, a adoção dos três genes de expressão mais estável (com menor valor M – Figura 13a), no entanto, para estas condições, a variação V encontra-se acima do valor de ponto de corte (0,15) sugerido pelo GeNorm (Figura 13b).



**Figura 12b:** Gráfico dos valores V fornecidos pelo GeNorm para os dez candidatos a gene de referência nas amostras de diferentes tecidos/órgãos durante o desenvolvimento de *Glycine max*.



**Figura 13:** **A:** Classificação fornecida pelo geNorm dos candidatos a genes de referência nas condições de estresse. Os genes de menores valores M são os mais estáveis. **B:** Gráfico dos valores V fornecidos pelo geNorm para os seis candidatos a gene de referência nas amostras de *Glycine max* submetidas a condições de estresse.

### 3.6.2 Análises pelo NormFinder

Para verificar e comparar os resultados obtidos pelo GeNorm, os dados de expressão obtidos das amostras de diferentes tecidos durante o desenvolvimento da soja, bem como das amostras coletadas em condições de estresse foram convertidos em dados de expressão relativa, através do método delta Ct e delta-delta Ct (LIVAK & SCHMITTGEN, 2001). Esses dados foram exportados para uma planilha do excel e avaliados pelo programa NormFinder seguindo as recomendações descritas por Andersen et al. (2004).

Este programa se baseia numa abordagem de estimativa de variância e classifica os genes de acordo com a sua estabilidade sob um dado conjunto de condições experimentais. Os genes com os menores valores de estabilidade são os mais estáveis (ANDERSEN et al., 2004).

A classificação gerada pelo NormFinder foi sutilmente diferente da classificação gerada pelo GeNorm. Nas amostras analisadas em conjunto “Total” (tecidos em desenvolvimento) o melhor gene indicado pelo NormFinder foi o *EF1α 1a* com um valor de estabilidade de 0,397 (Tabela 16). Este gene é indicado como o sexto na classificação de estabilidade fornecida pelo GeNorm.

Em vagens/flores a classificação foi igual nos dois programas, com os genes *EF1β* e *EF1α 1a* apresentando expressão mais estável.

Ambos os programas GeNorm e NormFinder identificaram os mesmos quatro genes mais estáveis em vagens, mas com diferenças na ordem de classificação (Tabela 16 e Figura 12a). Além disso, a classificação dos três primeiros genes candidatos em sementes é exatamente igual. Já em hipocótilos dois dos três primeiros genes indicados são os menos nos dois programas. Observou-se também que de acordo com o NormFinder o gene *EF1α 2b* foi o mais estável nas amostras de cotilédones, no entanto na classificação fornecida pelo GeNorm este foi somente o sexto gene mais estável (Figura 12a e Tabela 16).

Em folhas unifolioladas e na combinação de folhas unifolioladas/folhas trifolioladas (estádios de desenvolvimento) o gene *EF1α 3* foi indicado como o gene mais estável, já em folhas trifolioladas analisadas individualmente o melhor gene foi o *EF1α 1b*.

Nas amostras de raízes observa-se que em ambos os programas o gene mais estável foi o *EF1α 2a*. No GeNorm e NormFinder os seis primeiros genes em amostras de epicótilos foram os mesmos, mas há diferenças na ordem de classificação (Tabela 16).

Os resultados revelam que de acordo com o programa NormFinder os genes pertencentes a família multigênica do *EF1α* são os de expressão mais estável, excetuando-se vagens/flores e sementes, onde o gene *EF1β* aparece como o mais estável, esse comportamento também foi observado quando analisamos esses dados pelo GeNorm.

Nas análises das amostras de raízes e folhas submetidas a condições de estresse com PEG e AS, os dois programas mostraram resultados semelhantes, no entanto existem algumas diferenças. No caso da avaliação pelo NormFinder a classificação dos genes para amostras submetidas ao estresse com PEG foi *EF1α 2a > 1a > 1b > 2b > 3 > 1c* em folhas e *EF1α 2b, 1c, 2a, 1a, 3 e 1b* em raízes (Tabela 17).

Nas amostras de raízes e folhas submetidas ao estresse com AS, a classificação de genes mais estáveis para folha foi *EF1α 1b > 2a > 3 > 1a > 1c > 2b*, ao passo que para as amostras de raízes a classificação de estabilidade foi *EF1α 1a > 2b > 1c > 2a > 3 > 1c* (Tabela 17).

**Tabela 16:** Classificação dos genes de referência fornecida pelo programa NormFinder para as amostras de diferentes tecidos durante o desenvolvimento da soja.

| Classificação | Total           |              | Flores/vagens   |              | Vagens          |              | Sementes        |              | Cotilédones     |              | Folha trifoliolada |              | Folha unifoliolada |              | Estádios de desenvolvimento |              | Raiz            |              | Hipocótilos     |              | Epicótilos  |              |
|---------------|-----------------|--------------|-----------------|--------------|-----------------|--------------|-----------------|--------------|-----------------|--------------|--------------------|--------------|--------------------|--------------|-----------------------------|--------------|-----------------|--------------|-----------------|--------------|-------------|--------------|
|               | Gene            | Estabilidade | Gene            | Estabilidade | Gene            | Estabilidade | Gene            | Estabilidade | Gene            | Estabilidade | Gene               | Estabilidade | Gene               | Estabilidade | Gene                        | Estabilidade | Gene            | Estabilidade | Gene            | Estabilidade | Gene        | Estabilidade |
| 1             | EF1 $\alpha$ 1a | 0.397        | EF1 $\beta$     | 0.009        | EF1 $\alpha$ 1a | 0.003        | EF1 $\beta$     | 0.026        | EF1 $\alpha$ 2b | 0.420        | EF1 $\alpha$ 1b    | 0.212        | EF1 $\alpha$ 3     | 0.115        | EF1 $\alpha$ 3              | 0.452        | EF1 $\alpha$ 2a | 0.129        | EF1 $\beta$     | 0.084        | EF1A 1c     | 0.230        |
| 2             | EF1 $\beta$     | 0.375        | EF1 $\alpha$ 1a | 0.009        | EF1 $\beta$     | 0.003        | SKIP 16         | 0.026        | EF1 $\alpha$ 2a | 0.799        | EF1 $\alpha$ 3     | 0.231        | EF1 $\alpha$ 2b    | 0.201        | EF1 $\alpha$ 1b             | 0.463        | EF1 $\alpha$ 1c | 0.278        | EF1A 1a         | 0.097        | EF1A 2b     | 0.759        |
| 3             | EF1 $\alpha$ 1c | 0.415        | SKIP 16         | 0.046        | EF1 $\alpha$ 2a | 0.012        | EF1 $\alpha$ 2b | 0.074        | EF1 $\alpha$ 3  | 0.970        | EF1 $\beta$        | 0.237        | EF1 $\alpha$ 1a    | 0.678        | EF1 $\alpha$ 1a             | 0.467        | EF1 $\alpha$ 1b | 0.341        | EF1A 1c         | 0.230        | EF1A 1b     | 0.833        |
| 4             | EF1 $\alpha$ 2a | 0.459        | UKN1            | 0.073        | EF1 $\alpha$ 2b | 0.012        | EF1 $\alpha$ 1a | 0.397        | EF1 $\alpha$ 1c | 1.118        | EF1 $\alpha$ 1c    | 0.337        | EF1 $\alpha$ 1b    | 0.723        | EF1 $\beta$                 | 0.478        | MTP             | 0.375        | EF1A 2a         | 0.361        | EF1A 1a     | 1.120        |
| 5             | SKIP 16         | 0.559        | EF1 $\alpha$ 1b | 0.121        | EF1 $\alpha$ 1b | 0.028        | EF1 $\alpha$ 2a | 0.482        | UKN1            | 1.303        | EF1 $\alpha$ 2a    | 0.564        | EF1 $\beta$        | 0.924        | EF1 $\alpha$ 2a             | 0.585        | EF1 $\alpha$ 2b | 0.391        | SKIP 16         | 0.436        | EF1 $\beta$ | 1.159        |
| 6             | EF1 $\alpha$ 1b | 0.562        | MTP             | 0.734        | SKIP 16         | 0.192        | EF1 $\alpha$ 1c | 0.642        | EF1 $\alpha$ 1a | 1.387        | EF1 $\alpha$ 1a    | 0.589        | EF1 $\alpha$ 2a    | 1.015        | EF1 $\alpha$ 1c             | 0.644        | UKN1            | 0.433        | UKN1            | 0.461        | EF1A 2a     | 1.334        |
| 7             | UKN1            | 0.582        | EF1 $\alpha$ 1c | 0.736        | UKN1            | 0.206        | MTP             | 1.065        | EF1 $\alpha$ 1b | 1.496        | UKN1               | 0.856        | MTP                | 1.062        | UKN1                        | 0.699        | SKIP 16         | 0.553        | EF1A 2b         | 0.536        | UKN1        | 1.368        |
| 8             | EF1 $\alpha$ 2b | 0.586        | EF1 $\alpha$ 2a | 1.163        | EF1 $\alpha$ 3  | 0.531        | UKN1            | 1.499        | EF1 $\beta$     | 1.500        | SKIP 16            | 0.879        | UKN1               | 1.183        | SKIP 16                     | 0.731        | EF1 $\alpha$ 1a | 0.693        | EF1A 3          | 0.643        | MTP         | 1.642        |
| 9             | EF1 $\alpha$ 3  | 0.640        | EF1 $\alpha$ 3  | 1.708        | EF1 $\alpha$ 1c | 0.640        | EF1 $\alpha$ 3  | 1.646        | MTP             | 1.897        | MTP                | 1.349        | SKIP 16            | 1.279        | MTP                         | 0.793        | EF1 $\beta$     | 0.707        | MTP             | 0.728        | SKIP 16     | 1.896        |
| 10            | MTP             | 0.704        | EF1 $\alpha$ 2b | 3.099        | MTP             | 0.790        | EF1 $\alpha$ 1b | 2.190        | SKIP 16         | 1.905        | EF1 $\alpha$ 2b    | 2.481        | EF1 $\alpha$ 1c    | 1.462        | EF1 $\alpha$ 2b             | 1.147        | EF1 $\alpha$ 3  | 0.952        | EF1 $\alpha$ 1b | 0.884        | EF1A 3      | 1.988        |

**Tabela 17:** Classificação dos genes de referência fornecida pelo programa NormFinder para as amostras de raiz e folha submetidas a condições de estresse com PEG e AS.

| Classificação | Total           |              | Folhas AS       |              | Raízes AS       |              | Folhas PEG      |              | Raízes PEG      |              |
|---------------|-----------------|--------------|-----------------|--------------|-----------------|--------------|-----------------|--------------|-----------------|--------------|
|               | Gene            | Estabilidade | Gene            | Estabilidade | Gene            | Estabilidade | Gene            | Estabilidade | Gene            | Estabilidade |
| 1             | EF1 $\alpha$ 2a | 0.026        | EF1 $\alpha$ 1b | 0.020        | EF1 $\alpha$ 1a | 0.009        | EF1 $\alpha$ 2a | 0.028        | EF1 $\alpha$ 2b | 0.012        |
| 2             | EF1 $\alpha$ 2b | 0.037        | EF1 $\alpha$ 2a | 0.022        | EF1 $\alpha$ 2b | 0.010        | EF1 $\alpha$ 1a | 0.036        | EF1 $\alpha$ 1c | 0.017        |
| 3             | EF1 $\alpha$ 1b | 0.039        | EF1 $\alpha$ 3  | 0.028        | EF1 $\alpha$ 1c | 0.015        | EF1 $\alpha$ 1b | 0.039        | EF1 $\alpha$ 2a | 0.021        |
| 4             | EF1 $\alpha$ 1a | 0.041        | EF1 $\alpha$ 1a | 0.036        | EF1 $\alpha$ 2a | 0.019        | EF1 $\alpha$ 2b | 0.051        | EF1 $\alpha$ 1a | 0.022        |
| 5             | EF1 $\alpha$ 1c | 0.047        | EF1 $\alpha$ 1c | 0.043        | EF1 $\alpha$ 3  | 0.022        | EF1 $\alpha$ 3  | 0.057        | EF1 $\alpha$ 3  | 0.034        |
| 6             | EF1 $\alpha$ 3  | 0.054        | EF1 $\alpha$ 2b | 0.046        | EF1 $\alpha$ 1b | 0.056        | EF1 $\alpha$ 1c | 0.061        | EF1 $\alpha$ 1b | 0.035        |

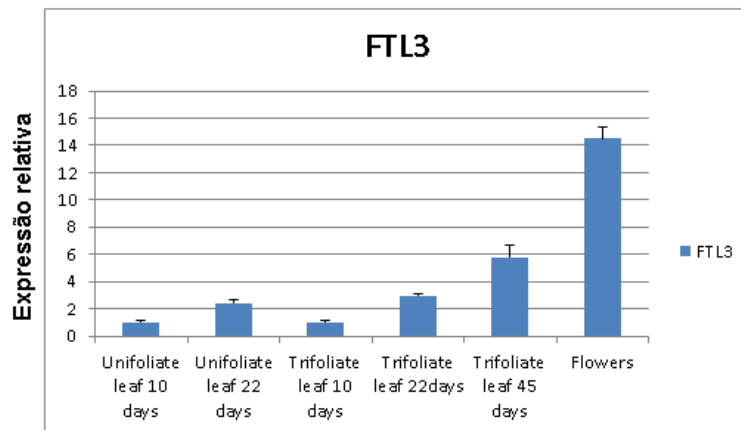
### 3.7 Validação dos genes de referência para RT-qPCR

Com o objetivo de validar o uso dos genes de referência indicados neste estudo, os genes *FTL3*, *LEA* e *PR3* foram analisados em relação aos seus respectivos níveis de expressão. Todos esses genes são reconhecidamente responsivos as condições testadas neste trabalho (Figuras 14, 15 e 16).

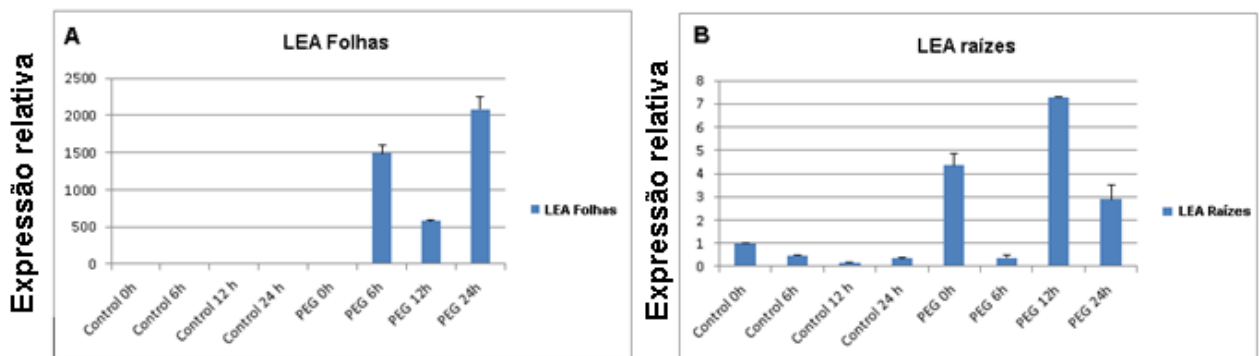
O perfil de expressão do gene *FTL3* foi avaliado em três distintos estágios de crescimento vegetativos da folha [V1 (10 dias), V4 (22 dias) e V9 (45 dias)]. Quando usamos os genes *EF1α 1a*, *EF1α 1b* e *EF1α 2a* como genes de referência, os níveis de transcritos aumentaram gradualmente com o decorrer do tempo, observando-se um pico no início da floração [oitava folha trifoliolada totalmente expandida - V9 (45 dias)] (Figura 14).

O gene *LEA* foi utilizado como controle positivo para os ensaios com estresse promovido por PEG, tendo em vista que o referido gene aumenta o nível de transcritos em resposta a este tipo de estresse (IRSIGLER et al., 2007, PINHEIRO et al., 2009). Nas análises em folhas, quando os genes *EF1α 1a*, *2a* e *2b* foram utilizados como normalizadores o nível de transcritos aumentou nos tempos 6, 12 e 24h. (Figura 15a). Quando usamos os genes *EF1α 1c*, *2a* e *2b* como genes de referência o nível de transcritos em raízes aumentou nos tempos 0, 12 e 24h (Figura 15b).

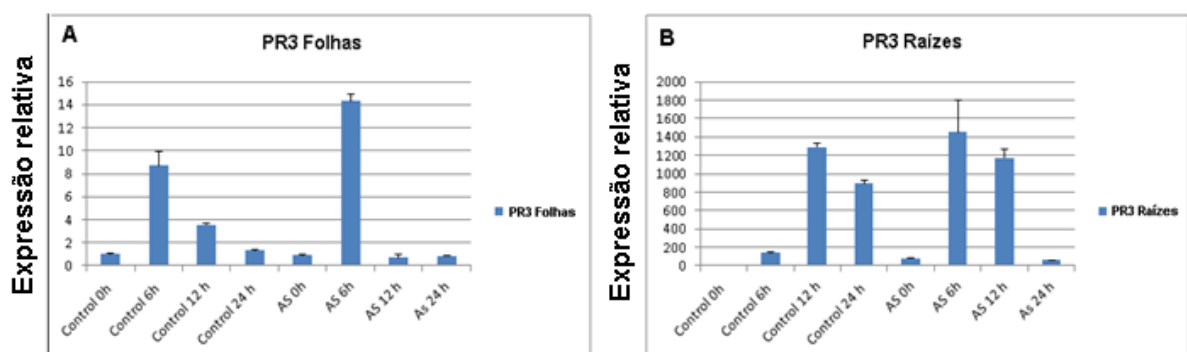
O gene *PR3* foi utilizado como controle positivo para o estresse promovido por ácido salicílico tendo em vista que estudos anteriores (MORAES, 2012) mostraram que este gene responde a este tipo de estresse. Em folhas, quando os genes *EF1α 1a*, *1b*, *2a*, *2b* e *3* foram utilizados como genes de referência, observou-se um aumento de transcritos em seis horas, comportamento este igual ao observado por Moraes (2012). Em raízes, o nível de transcritos também aumentou significativamente em seis horas, quando os genes *EF1α 1a*, *1c*, *2a* e *3* recomendados pelo GeNorm foram utilizados (Figura 16a e 16b).



**Figura 14:** Expressão relativa do gene *FTL3*, usando os genes *EF1 $\alpha$ s* indicados pelo GeNorm como normalizadores.



**Figura 15:** Expressão relativa do gene *LEA* usando os genes *EF1 $\alpha$ s* indicados pelo GeNorm como normalizadores. **A:** folhas **B:** raízes. Gene utilizado como controle positivo para o estresse com polietileno glicol.



**Figura 16:** Expressão relativa do gene *PR3* usando os genes *EF1 $\alpha$ s* indicados pelo GeNorm como normalizadores **A:** folhas **B:** raízes. Gene utilizado como controle positivo para o estresse com ácido salicílico



## 4 DISCUSSÃO

Soja (*Glycine max*) é uma cultura importante mundialmente, apresentando grãos com alto teor proteico essenciais para o consumo humano e animal (LE et al, 2011; SCHMUTZ et al, 2010). A soja é uma espécie de leguminosa da ordem Fabales que teve seu genoma recentemente sequenciado, tornando-se disponível para a comunidade acadêmica, viabilizando a identificação da função de novos genes e possibilitando avanços no estabelecimento de novas estratégias moleculares de melhoramento genético (SCHMUTZ et al, 2010).

A função de um gene pode ser identificada usando diferentes técnicas, sobretudo, através do silenciamento gênico. Contudo, análises de expressão gênica também contribuem para elucidar o papel dos genes na célula. Essas análises, em sua grande maioria são realizadas, atualmente, através de ensaios por PCR em tempo real (qRT-PCR). A técnica de qRT-PCR é altamente sensível, eficiente e confiável, no entanto para o uso da qRT-PCR faz-se necessário a adoção de uma estratégia de normalização adequada. O uso de genes constitutivos ou “Housekeeping” é comumente aceito para normalizar dados de expressão gênica.

Genes “housekeeping” são codificados por famílias multigênicas em muitas espécies de plantas e são altamente idênticos entre si (XU et al., 2007). Embora os genes “housekeeping” sejam amplamente usados na normalização de dados de qPCR, aparentemente não há preocupação entre os pesquisadores em se determinar genes normalizadores usando todos os membros da família gênica. Geralmente são utilizados genes, escolhidos aleatoriamente, que codificam diferentes proteínas.

No presente trabalho, buscou-se identificar genes de referência *EF1 $\alpha$*  com expressão estável em duas condições de estresse e em diferentes tecidos durante o desenvolvimento da soja. Para tal, os genes *EF1 $\alpha$*  membros da família multigênica foram testados nas condições de estresse e nos diferentes tecidos. Os genes *MTP*, *SKIP 16*, *UKNI* e *EF1 $\beta$*  também foram testados nos diferentes tecidos em desenvolvimento. Estudos anteriores, mostraram que os genes *EF1 $\alpha$*  foram selecionados em diferentes condições experimentais como bons genes de referência, apresentando expressão estável (HAN et al., 2012; LUO et al., 2010; MARTIN et al., 2008, REIDE et al., 2006 e WAN et al., 2010), no entanto, nestes estudos, não havia preocupação em se determinar qual gene membro era selecionado. Muitas vezes um único par

de *primers* reconhecia mais de um gene membro devido à alta identidade existente entre tais genes.

No presente trabalho, as análises dos valores das médias de *Ct* mostraram que em diferentes tecidos e condições de estresses os genes *EF1α* apresentaram variações de expressão (Figuras 10 e 11). Estes achados estão de acordo com os estudos de expressão de genes *EF1α* de uva (REIDE et al., 2006) e Tung tree (HAN et al., 2004), onde as médias dos valores de *Ct* dos genes *EF1α* também variaram.

O gene *EF1α 2a*, juntamente com os genes *MTP*, *UKNI*, *SKIP 16* e *EF1β* foram os que mostraram a menor variação de *Ct*, entre as amostras de desenvolvimento (Figura 11). Estes resultados corroboram com a classificação proposta pelo GeNorm, onde os genes *SKIP 16*, *UKNI* e *MTP* aparecem como os mais estáveis quando todas as amostras dos diferentes tecidos “Total” são analisadas em conjunto (Figura 12a). De acordo com o GeNorm o gene *EF1α 2b* apareceu como o mais instável dentre as amostras de desenvolvimento “Total”, este gene também apresentou a maior variação de *Ct* (Figuras 11 e 12a)

Nas amostras submetidas aos estresses com AS e PEG, o gene *EF1α 3* foi o que apresentou a maior variação da média dos valores de *Ct* (Figura 10). Este resultado é semelhante ao observado nos dados fornecidos pelo GeNorm, onde este gene também aparece como o menos estável (Figura 13a). Já o gene com a menor variação dos valores de *Ct* foi o *EF1α 1a*, no entanto no GeNorm o gene mais estável é o *EF1α 2a*. Esses resultados estão de acordo com os reportados na literatura, onde se observa que os diferentes métodos utilizados para análise da estabilidade de expressão de genes candidatos a normalizadores, muitas vezes não produzem resultados idênticos (WANG et al., 2011; JIAN et al., 2008).

Os genes *EF1α 1c* e *2a* segundo o GeNorm foram os genes mais estáveis em raízes submetidas aos estresses com AS e PEG, respectivamente. Em folhas, *EF1α 2a* e *2b* foram os melhores genes em condições de estresse com AS e PEG, respectivamente. Estes achados mostram que em diferentes tecidos e condições de estresse, diferentes genes são requeridos para normalização. Estes resultados estão de acordo com os encontrados na literatura. Em amostras de lírios, o gene *EF1α* é indicado para folhas e raízes em algumas condições de estresse, mas não é indicado em outros tipos de tecidos, nestas mesmas condições (LUO et al., 2010). Em folhas de populus (BASA et al., 2009) e pepino (WAN et al., 2009) sob condições de estresse com Cádmio o gene *EF1α* foi o mais estável.

Em diferentes tecidos durante o desenvolvimento da soja os genes indicados para normalização diferiram, no entanto dentre os dez genes analisados (*EF1α 1a, 1b, 1c, 2a, 2b, 3* e *UKNI, SKIP16, EF1β* e *MTP*), os genes membros da família gênica *EF1α* foram os mais estáveis na maioria dos tecidos testados, tais como cotilédones (*EF1α 3* e *1c*); epicótilos (*EF1α 1b, 2b* e *1a*); hipocótilos (*EF1α 1a* e *EF1β*); vagens (*EF1α 2a* e *2b*) e raízes (*EF1α 2a; UKNI*) (Figura 12a). Nos outros tecidos testados, tais como folhas trifolioladas e folhas unifolioladas os genes *UKNI* e *SKIP16* corresponderam à combinação mais estável. Já em sementes durante a germinação os genes *EF1β* e *SKIP16* são os mais indicados.

Estudos de normalização em *Lolium perenne* mostraram que os genes *EF1α*, pertencentes a uma família multigênica de dois membros, exibem expressão diferencial, assim como observado em soja. No entanto, diferentemente do observado neste trabalho, as análises de 442 amostras de diferentes tecidos revelou que a combinação dos dois genes *EF1α* apresentava a maior estabilidade tornando possível usá-los como genes de referência nessas condições (LEE et al., 2010). Resultados semelhantes são encontrados em *Linum usitatissimum* L. Nesta espécie os genes membros da família multigênica do fator de iniciação da tradução, nomeados *ETIF1, ETIF3E, ETIF3H, ETIF4F* e *ETIF5A* foram analisados em relação à estabilidade de expressão. De acordo com o GeNorm, alguns membros da família gênica são indicados para normalizar amostras de diferentes tipos de tecidos (HUIS et al., 2010).

Nesse trabalho, os programas GeNorm e NormFinder foram utilizados para verificar a estabilidade de expressão dos genes candidatos a normalizadores. De acordo com os resultados obtidos observa-se que ambos os programas indicam os genes *EF1α* como os genes mais estáveis em diferentes tecidos. No entanto, em alguns tecidos os genes *EF1α* membros da família multigênica não são os mais indicados, no entanto outros genes (*MTP, SKIP 16, UKNI* e *EF1β*) são indicados neste estudo. Nas diferentes condições de estresse não foi possível determinar as combinações ideais de genes estáveis para normalização, neste caso, enfatizamos a necessidade da adição de novos genes com expressão estável, a fim de se determinar combinações ideais para normalização.

Estes achados reforçam a necessidade de estudos de normalização utilizando os diferentes membros das famílias multigênicas, já que estes genes são diferencialmente expressos, e os diferentes membros podem ser indicados para normalização em algumas condições, enquanto que em outras não.

Para validar os resultados encontrados o gene *FTL3* foi utilizado. Este gene aumenta o nível de transcritos à medida que a planta se desenvolve e se aproxima do início da floração. A proteína codificada por este gene desempenha uma função primordial na transição do estágio vegetativo para o estágio reprodutivo, favorecendo a floração (ODA et al., 2011). Em *A. thaliana*, *FTL3* atua como um promotor floral e um integrador de várias vias de floração (NOTAGUCHI et al., 2008). Os resultados mostram que este gene exibiu o padrão de expressão esperado, visto que, teve sua expressão aumentada gradativamente à medida que a planta se aproximava da floração (Figura 14). Tais resultados corroboram com os encontrados no trabalho de Hu et al. (2009) mostrando que quando usamos os genes *EF1 $\alpha$*  como normalizadores, os resultados esperados não são alterados.

Os genes *LEA* e *PR3* foram avaliados em relação aos seus perfis de expressão, os resultados mostraram que ambos os genes respondem aos estresses com polietileno glicol e ácido salicílico, respectivamente. Esses resultados demonstram que as plantas foram realmente submetidas ao estresse osmótico promovido pelo PEG e ao estresse com AS. Ademais se observou que os genes *EF1 $\alpha$* , apesar de não serem indicados como os melhores genes de referência para essas análises, não interferiram nos resultados esperados (Figura 15a, 15b, 16a e 16b).

Os dados mostram que existem variações de expressão entre os genes membros da família gênica em diferentes tecidos durante o desenvolvimento e condições de estresse. Entretanto, foi possível determinar genes *EF1 $\alpha$*  com estabilidade de expressão em tecidos/condições específicas que podem ser usados para normalizar ensaios por qRT-PCR em soja. Ademais, esses dados também podem ser usados como base para o estabelecimento de genes de referência em leguminosas da tribo Phaseoleae, tendo em vista que genes *EF1 $\alpha$*  ortólogos foram identificados entre as diferentes espécies.

## 5 CONCLUSÕES

1. Combinações diferentes de genes são requeridas para normalização em diferentes tipos de tecidos durante o desenvolvimento da soja. Em relação aos genes pertencentes à família multigênica do *EF1 $\alpha$* , alguns membros são indicados para a maioria dos tecidos testados, tais como cotilédones (*EF1 $\alpha$  3* e *1c*), raízes (*EF1 $\alpha$  2a*), epicótilos (*EF1 $\alpha$  1b*, *2b* e *1a*), hipocótilos (*EF1 $\alpha$  1a* e *EF1 $\beta$* ) e vagens(*EF1 $\alpha$  2a* e *2b*). Para as amostras de folhas trifolioladas, folhas unifolioladas, vagens/flores, estádios de desenvolvimento e sementes os genes *EF1 $\alpha$*  não são os mais indicados para normalização.
2. Nas situações de estresse com PEG e AS, não foi possível estabelecer combinações ideais de genes para normalização, havendo, portanto necessidade de inclusão de genes candidatos a normalizadores com expressão estável, a fim de se estabelecer estratégias apropriadas de normalização.
3. Os dados mostram que existem variações de expressão entre os genes membros da família gênica nas situações testadas neste trabalho. Entretanto, foi possível determinar genes *EF1 $\alpha$*  com estabilidade de expressão em tecidos/condições específicas. Estes genes de expressão estável podem ser usados para normalizar ensaios por qRT-PCR em soja. Além disso, esses dados podem ser usados como base para o estabelecimento de genes de referência em leguminosas da tribo phaseoleae, tendo em vista que genes *EF1 $\alpha$*  ortólogos foram identificados entre as diferentes espécies.

## 6 REFERÊNCIAS BIBLIOGRÁFICAS

ANDERSEN CL, JENSEN JL, ORNTOFT TF. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 2004; 64: 5245-5250.

AXELOS M, LIBOZ T, BARDET C, et al. The gene family encoding the Arabidopsis thaliana translation elongation factor-1A: molecular cloning, characterization and expression. *Molecular and General Genetics* 1998; 219: 106–112.

BASA B, SOLTI A, SÁRVÁRI É , et al. Housekeeping gene selection in poplar plants under Cd-stress: comparative study for real-time PCR normalisation. *Functional Plant Biology* 2009; 12: 1079–1087.

CZECHOWSKI T, STIIT M, ALTMANN T, et al. Genome – wild identification and testing of superior reference –genes for transcript normalization in Arabidopsis. *Plant Physiology* 2005; 139: 15-17.

DHEDA K, HUGGETT JF, BUSTIN SA, et al. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques* 2004; 37: 112–119.

EMBRAPA. Cultivares de soja, regiões norte e nordeste do brasil. Embrapa soja, Londrina, PR, 2012. Disponível em <[http://www.cnpso.embrapa.br/index.php?op\\_page=294&cod\\_pai=17](http://www.cnpso.embrapa.br/index.php?op_page=294&cod_pai=17)>. Acesso em 02/08/2012.

EXPOSITO-RODRIGUEZ M, BORGES AA, BORGESPEREZ A, et al. Selection of internal control genes for quantitative real-time RTPCR studies during tomato development process. *BMC Plant Biology* 2008; 8: 131.

HAN X, LU M, CHEN Y, et al. Selection of Reliable Reference Genes for Gene Expression Studies Using Real-Time PCR in Tung Tree during Seed Development. *PLoS ONE* 2012; 7: e43084.

HOAGLAND D, ARNON, DI. The water culture method for growing plants without soil. California Agriculture Experimental Station Circular 1950. 347.

HU R, FAN C, LI H, et al. Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. BMC Molecular Biology 2009; 10: 93.

HUIS R, HAWKINS S, NEUTELINGS G. Selection of reference genes for quantitative gene expression normalization in flax (*Linum usitatissimum* L.) BMC Plant Biol 2010; 10:71.

JIAN B, LIU B, BI Y, et al. Validation of internal control for gene expression study in soybean by quantitative real-time PCR. BMC Molecular Biology 2008; 9:59.

## M

LE, DT, NISHIYAMA R, WATANABE Y, et al. Genome-wide survey and expression analysis of the plant-specific NAC transcription factor family in soybean during development and dehydration stress. DNA research 2011; 18: 263-276.

LEE J M, ROCHE J R, DONAGHY D J, et al. Validation of reference genes for quantitative RTPCR studies of gene expression in perennial ryegrass (*Lolium perenne* L.) BMC Molecular Biology 2010, 11:8.

LIVAK KJ, SCHMITTEN TD. Analysis of gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 2001; 25: 402-408.

LUO H, CHEN S, WAN H, et al. Candidate reference genes for gene expression studies in water lily. Analytical Biochem 2010; 404:100–102.

MARTIN R, HOLLENBECK V, DOMBROWSKI J. Evaluation of Reference Genes for Quantitative RT-PCR in *Lolium perenne*. Crop Sci 2008; 48:1881-1887.

MORAIS, V.D. Avaliação da expressão da toxina da soja (SBTX) por indutores da defesa de plantas. 2012. 96p. Dissertação (Departamento de Bioquímica e Biologia Molecular), Universidade Federal do Ceará - UFC. Fortaleza, Ceará.

NOTAGUCHI M, MITSUTOMO ABE, KIMURA T, et al. Long-Distance, Graft-Transmissible Action of Arabidopsis FLOWERING LOCUS T Protein to Promote Flowering. Plant Cell Physiol 2008; 49: 1645-1658.

ODA A, TAKAKO N, TUOPING L, et al. CsFTL3, a chrysanthemum FLOWERING LOCUS T-like gene, is a key regulator of photoperiodic flowering in chrysanthemums. *Journal of Experimental Botany* 2012; 63: 1461-1477.

PINHEIRO GL, MARQUES CS, MAXIMILLER DBL, et al. Complete inventory of soybean NAC transcription factors: Sequence conservation and expression analysis uncover their distinct roles in stress response. *Gene* 2009; 444: 10-23.

REID KE, OLSSON N, SCHLOSSER J, et al. An optimized grapevine RNA isolation procedure and statistical determination of reference genes during berry development. *BMC Mol Biol* 2006; 6:27.

SCHMUTZ J, CANNON SB, SCHLUETER J, et al. Genome sequence of the palaeopolyploid soybean. *Nature* 2010; 463: 178-183.

VANDESOMPELE J, DE PRETER K, PATTYN F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* 2002; 3:18.

WAN H, ZHAO Z, QIAN C, et al. Selection of appropriate reference genes for gene expression studies by quantitative real-time polymerase chain reaction in cucumber. *Anal Biochem* 2010; 399: 257-261.

WANG YI, YU K, POYSA V, et al. Selection of reference genes for normalization of qRT-PCR analysis of differentially expressed genes in soybean exposed to cadmium. *Mol Biol Rep* 2011; 39: 1585-1594.



## ANEXOS

**Anexo 01:** Sequências de cDNAs deduzidas dos diferentes genes membros da família multigênica do *EF1 $\alpha$* . Em rosa são mostradas as sequências dos *primers* senso e antisense.

>*GmEF1 $\alpha$  1a*

```
atcactctgtgctccctcgctctatcttctcagtcctctgtgtttgcggtgaggattccgaacgagtgaccttctt
cgtttctcgcaaaagtttaagtaatcattgggtaaggaaaagggtcacatcaacattgtcgtcattggacatgtcga
ctctgggaagtcaactaccactggtcacttgatctacaagcttggaggtattgacaagcgtgtgattgagagggtt
cgagaaggaggtgccgagatgaacaagaggtcattcaagatgcctgggtgctcgacaagctcaaagctgagcgt
tgaaagaggaattaccattgatattgctttgtggaagtttgaaccaccaagtaactactgcacggtcattgatgc
tcttgacatcgtgactttatcaagaacatgattactggtacctcccaggccgactgtgctgtccttattattga
ctccaccactggtggttttgaagctggtatttctaaggatggacagaccctgagcatgctcttcttggctttcac
cctaggtgtgaagcagatgatctgttgcgtgaacaagatggatgccactaccccccaagtaactctaaggctaggtg
tgatgaaatcgtgaagggaagtctcttcttacttgaagaaggttggttacaacccagacaagattccctttgttcc
catctctggttttgaggggtgacaacatgattgagaggtccaccaaccttgactggtacaagggaccaactctcct
tgaggctcttgaccaaatcaatgagcccaagaggccctccgacaagcctctaaggcttccattgcaggatgtcta
caagattggtggtattggtactgtgccagtgggacgtgtagagactggggttgtgaagcctggtatggtggtgac
ttttggtcccactgggctgacaactgaggttaagctgttgagatgcaccatgaggctctcacagaggctcttcc
aggtgacaatggttgatttaagtgaagaatggtgagcaagatctcaagcgtggttttggttgcatccaactc
caaggacgaccctgccaaaggaagctgccaacttcacatcccaagtcattatcatgaaccatcctggccagattgg
taatggatacgcaccagtccttgactgccacacttctcacattgctgtggaagtttctgaaatcttgaccaagat
tgacagggcagatctggttaaggagcttgagaaggagcccaaatttttgaagaatggtgatgctggtatggttaagat
ggttccaaccaagcccattggtggttgaactttctctgagatcctccccttggctggttttctgtgagggacat
gctgcagaccgtagctgttggagtcacaaagagttgagaagaagacccaccggagccaaggtcacaaggc
tgccgccaagaagaagtgaattgcattttggcaattttgctagcacatgtgatcatcaacgtggtttcaaaaaaac
ttggtcctttacagtagtttatctttgcagagcttaggtgtttgttttaccagttatattttgaagtgctccgc
gatttcatgtagccgtagcccttcaaaaactgggttcttgatcggcggtaacattttcgttgcgttttggttttgat
gagtactgtttttgttttgatggtaaaagtctgagattttcaaatcacaagcagccatagggttttagtccat
ttcctttgctgtctgaggagggatgtcttaaat
```

>*GmEF1 $\alpha$  2a*

```
ctcttcaactccttcggttcaactcttcttcttcttgcggctagggtttttagcgcagcttcttctagatttaaggaaga
tgggtaaggaaaaggtcacatcagatattgtggttattggccatggtgactctgggaagtccactaccactggcc
acctgatctacaagcttggaggcattgacaagcgtgttattgagagatttgagaaggaagctgctgagatgaaca
agaggtctttcaagatgcctgggtgcttgataagcttaaggctgagcgtgaaagaggaatcacaattgatattg
ccttgtggaagtttgaaccaccaagtaactattgcacagtcattgatgccctggacacagggatttcatcaaga
acatgatcactgggacatcccaagctgactgtgctgttcttatacattgattccactactggtggttttgaagctg
gaatttcaaaggatggacagactcgtgaacatgctctgcttcttacccttgggtgaaacagatgatttgtt
gctgtaacaagatggatgctactacaccaagtaactccaaggccaggtatgatgaaattgtgaaggaagtttctt
cctatttgaagaaagtaggatacaaccctgacaagattccttttcttataatctggttttgaggagacaaca
tgattgagaggtccacaaaccttgactggtacaaggtcctactctgctagatgacttgaccagatctctgagc
ccaagaggccttctgacaagcccctcaggctacccctcaggatgtgtacaagattggaggaattggaactgtgc
ctgtgggacgggttgagactggtgtcttgaagcctggaatggtggtgacttttgcaccaactggactgacaaccg
aagtttaagctgtggaatgcaccatgaagctctcacagaggctcttcccggtgataatggttgattcaatgta
agaatgttgcgtgtaaggatctcaagcgtggttatggtgctcgaactcaaaggatgatcctgccaaaggagctg
ctaacttcaactgcccaggttatcatcatgaaccatcctggctcagattggaaatggctatgccctgttcttgact
gccacacttcccacattgctgtcaagtttgcgtgaactcatgaccaagattgacaggcagatctggcaaaagagcttg
agaaggaacccaagtttttgaagaatggtgatgctggttttggtaagatgattccgaccaaacccatggtggttg
aaactttctctgagtaacccccacttggctgctttgctgtcagggatatgctgcaaacgttctgtgtgggagtc
tcaagaacgtggagaagaaggtcctactggagccaaggtcaccaaggtgccagaagaagaagtgaatcgtgc
ggtttggttcatcaggggatgtcgtttcttatggttacaataaatggtggtttcttgccttctgtcttctggtttc
taggtagcttgttttccgacatagtttgaagctccaccatcatctcgcaactttgttcccagaattgggttc
ttgatcgacgggtggcaagactccttttatcattctgttttaaagtggttgggttgagaaccctgattacatt
tttgttaagcgcagcaggttttagggctttgcccgttgcgttgggttggctttttaaagtgcaactttatatt
gtgttcaatttttgcttgggttggcttttaaaaaatcaaatatgtg
```

>GmEF1 $\alpha$  1b

atcactctgctgcctcgtctctgtttctgtctctgtgtttgctgctgaggattccgaacgagcgaccttcttcggt  
 tctcgcagaaagtttaagtcacatggtgtaaggaaaagactcacatcaacattgctcgtcattggacatgtcgcactc  
 ggaagtcaactaccactggtcatttgatctacaagcttgaggattgacaagcgtgtgattgagaggttcgag  
 aaggaggctgctgagatgaacaagaggtcattcaagatgacctgggtgcttgacaagctcaaggctgagcgcgaa  
 agaggaattaccattgatattgctttgtggaagttgaaaccaccaagtaactactgcacagtcacgatgcccc  
 ggacatcgtgactttatcaagaacatgattactggtacctcccaggctgattgtgctgtccttatcattgactcc  
 accactggtggttttgaagctggtatttctaaggatggacagactcgtgagcatgctcttcttgccttaccctt  
 ggtgtgaagcagatgatctgctgctgtaacaagatggatgccactaccccgaagtactctaaggctaggatgat  
 gaaattgtgaaggaagtctcttctacttgaagaaggttggttacaaccagacaagattccatttgtgcccac  
 tctggtttcgagggtgacaacatgattgagaggtccaccaaccttgactggtacaagggaaccaactctccttgag  
 gctcttgaccaaatcaatgagcccaagaggccctcagacaagccactcaggcttccattgcaggatgtctacaaa  
 attggtggtattggtaccgtgccagtgggacgtgttgagactgggtcgtggaagccggatgggtggtgactttt  
 ggtcccactgggtgacaactgaggttaagctgctgtgagatgcaccatgaggctctcacagagctcttccaggt  
 gacaatggtggctttaaagtggaagaatgtgacgtcaagatctcaagcgtggtttggatccaactccaag  
 gatgacctgccaaggaagctgccaacttcacatctcaagtcattatcatgaaccacctggccagattggaat  
 ggatacgcaccgggtgcttgattgccacacttctcacattgctgtcaagtttctgaaatctgaccaagattgac  
 aggcgatctggtaaggagcttgagaaggagcccaaattttgagaatggtgatgcaggatggttaagatgatt  
 ccaaccaagcccattggtggtgaaacttctctgagatcctccccttggctgctttgctgtgagggacatgctg  
 caactgtggctgtcggagtcatcaagagtggtgagaagaagaccccaccggagccaagggtaccaaggccgct  
 gccaagaagaagtgaattgcatttgggcaattttgctagcacatgtgatcattcatcgtggttactcctttatagta  
 gttttatccttgcagagctcttaggtgtttggtttaagttatatttttaagtttctgcccgatttcatgtagccgta  
 actttcaaaaactaggttcttgcagcgggtggtcaattttcattgctgctttgcttttggatgagtagctgctttt  
 ttttgatggtaagaagagctgagatatttcgaatttcacaagcagctatagggttttagtccatttcccttgc  
 gctgagggatgttttaaagttgcatttaattataacgaagttttataaactgtttatggtttaaaggctattatt  
 cttt

>GmEF1 $\alpha$  2b

cttctcttcacacctttgttactcttcttctcttgcggctagggttttagcgcagcttcttctagatttaagga  
 agatgggtaaggaaaaggttcacatcagatattgtggtcattggccatgctgactctgggaaatccactaccactg  
 gtccactgatttacaagcttgaggcattgacaagcgtgttattgagaggtttgagaaggaagctgctgagatga  
 acaagaggtcttcaagatgctggtgctggacaaacttaaggctgagcgtgaaagaggaatcaccattgata  
 ttgctttgtggaagttgaaacaacaaagtattattgcacagttattgatgctgctggacatagggatttcatta  
 agaatatgattactgggacatcccagctgactgtgctgttcttatcattgattcgaccactggtggttttgaag  
 ctggtatttccaaggatggacagactcgtgaacatgctctgcttctcattcacccttgggtgtgaaacagatgatt  
 gctgctgtaacaaaatggatgctactacacccaagattccaaggccaggtatgatgaaattgtgaaggaagtct  
 cttcctacttgaagaaagtaggatacaacctgacaagattccttttgttctatctctggttttgagggagaca  
 acatgattgagaggtccacaaaccttgactggtacaagggtccaactctgcttgatgcacttgaccagattagt  
 agcccaagagccctctgacaagccctcaggcttccccttcaggatgtgtacaagattggaggtattggaactg  
 tgccagtgggacgtgttgagaccggtgcttgaagcctggaatggtggtgacttttgaccaactggactgaca  
 ctgaagtcaagtctgtggagatgcaccatgaatctcttacagaggcacatcctggtgacaatgtgggattcaatg  
 ttaagaatggtgctgttaaggatttgaagcgtggttatggtgctcaaactcaaaggatgacctgcaaaggagg  
 ctgctaacttcacagcccagtcacatcatgaaccacctggtcagattggaaatggctatgcccctgtcctcg  
 actgccacacttctcacattgctgtcaagtttgtgtaactcatgaccaagattgacagggcatccggcaaagagc  
 ttgagaaggagcccagtttttgaagaacgggtgatgctggttttggtaagatgattccaacaaacccatggttg  
 ttgaaacttctccgagatcctccacttggtaggtttgctggttagggacatgctgcaactgttctgtggtggag  
 tcatcaagaacgttgagaagaaggatcctaccggagccaagggtaccaaggctgccagaagaagaagtgaatcg  
 tggggctggttcatcaggggatggttacaataaatggtggtttctttctgactcttgtgtcttcttttct  
 taggtagcttgttttccggacaaagtttgaagctccaccatcatctcgcaactggttccagaactgggttc  
 ttgatcgacgggtggcaaaattgcttttattatctgtgttttaagtgtgtgtgttgcggaaccctgattaca  
 tttttgtaagcgcagcgagtttccaggacttctgctgctgtgtgtgctttggtttatataaagtcaactttctat  
 ttgtagtgctt

>GmEF1 $\alpha$  1c

aaactctgctcttgccttagccatcactaacactcagctctctgaattggggctgagggtttcgagtaacgttta  
 ttttccacgagtttaagatatacattgggtaagagaaggttcacattaacattgtggtcattggccatggtgactc  
 tgggaagtcaaccaccactggccacttgatttacaagcttgaggattgacaagcgtgttattgagaggtttga  
 gaaagaagctgctgagatgaacaagaggtcattcaaatagcctgggtgcttgacaagcttaaggctgagcgtga  
 gagaggaattaccattgatattgcattatggaagtttgagaccacaaataactactgcactgttattgatgctcc  
 aggacatcgtgatttcatcaagaacatgattactggaacttcccaggctgattgtgctgtccttatcattgattc  
 cacaactggtggttttgaagctggtatttccaagatggacagactcgcgagcatgctctacttgccttaccct

tgggtgcaagcaaatgatctgttgctgtaacaagatggatgccactaccccaaagtactctaaggccaggtatga  
 agaaattgtgaaggaagtttcttctacttgaagaaggtaggttacaaccagacaagattccatttgttcccat  
 ctcaggttttgagggtgacaacatgattgagaggtccaccaatcttgattggtacaagggaccaacccttctcga  
 ggctcttgaccagatcaatgagcccaagagaccctcagacaagcctttgaggcttcccttgaggatgtctacaa  
 gattgggtggtattggaacagtgccagtgggacgagttgaaactgggtgcttgaagcccggatgggtgggtgacttt  
 tgctcccactggcctgacaactgaagttaagtctgtggagatgcacatgaagctctccaggaggcccttccagg  
 tgacaacgtaggatttaatgtgaagaatgttgcagtcaaggatctcaagcgtggttttgttgcctcaaactctaa  
 ggatgatcctgccaaaggaggctgccaaacttcacatcccaagtattataatgaaccatcctggccaaattggaaa  
 tggttatgctcctgtgcttgactgccacacttctcacattgccgtcaagtttgccgaacttgtgaccaagattga  
 caggcgatctggtaaggagattgagaaggagcccaaattcttgaaaaatggagatgcagggttatgttaagatgat  
 tcccaccaagcctatgggtggtcgaaaactttctccgagatccctcctcttggctggttttgccgtgagggacatgcg  
 acaaactgtggctgttggagtcacaaagagtggtggagaagaaggaccaactggagccaaggtcaccaaggctgc  
 acagaagaagaagtgaatacatttttgtggttcttcaagtaacatgcatagttacatttatagttacacaca  
 acagtggttccttattagttattttattttttgagagtataggtgcttttacttggttatagttttacttttatgta  
 tcagccatttctattgcagccacctaccctttcaaaactgggttcttgatgggctgggtgacttttctattttata  
 agctgtatgagaactctttttgcacaagcagctacagggttttagtt

>GmEF1 $\alpha$  3

gctacactccttctcttctcgcggcttctttctccaagattaagtttcagtagtaatgggtaaggaaaaggttca  
 cattaacatcgtggtcattggccatgtcgaactctgggaagtcaaccaccaccgggtcaccttatctacaagctcgg  
 aggtattgacaagcgtgtcatcgagagattcgagaaggaagctgctgagatgaataagaggcttttcaagtatgc  
 ctgggtgcttgacaagctcaaggctgagcgtgaaagaggaatcaccattgatattgctctgtggaagtttgagac  
 cacaaagtactactgcacagtcattgatgccctggacatagggatttcatcaagaacatgattactggaacatc  
 ccaagctgactgtgctgttcttattattgattccaccactgggtggttttgaagctggtatttccaaggatggaca  
 gactcgtgagcatgctctccttgacattcactcttgggtgtgaggcagatgatttgcctgcaacaagatggatgc  
 cactacacctaaagtactccaaggcaaggtaacgatgaaattgtcaaggaagtttcttctacatgaagaaggtcgg  
 atacaaccctgacaaaaattccttttgttcccatctctggttttgaggagataacatgattgagaggtctacaaa  
 ccttgactggtaacaaggtccaacccttcttgaggcccttgatcagatcaatgagcccaagaggccatcggacaa  
 gcctctccgattgccccttcaggatgtctacaagattgggtgggataggaactgtgctgttggacgggttgaac  
 cgggtcttgaaacctggaatgggtggttacctttgcaccaactggactaacaactgaagttaagtctgtggagat  
 gcaccatgaaactctcactgaggcccatcctgggtgacaatgtgggattcaatgtaagaatgttctgtcaagga  
 tctcaagcgtggttttgttgcgtctaactctaaggatgaccctgccaaggaggctgtaacttcacatcccaagt  
 tatcatcatgaaccacctggacagattggaatggctatgcacctgttcttgattgccacacctccacattgc  
 cgtcaagtttgcgtgaactcttgaccaagattgacaggcgtctggttaaggaacttgagaaggagcccaaattcct  
 gaagaatgggtgatgctggtttgggtcaagatgattcccaccaagccatgggtggttgaaccttctccgagtatcc  
 tccacttggctggttttgcgtgagagacatgcgtcaaactgtggctgtgggagtcacaaagcgtcgagaagaa  
 ggacctactggagctaagggtcaccaaggctgcacagaagaaggccaaatgaattgtgctggatgggtttatcaag  
 gaaactctattattatgactacaataaatggtttcttgttctgtagtacttcattttttcgtgtagggaacttgtt  
 tatcagattttgtgtcaggtttgctccgtcatctcgcaactttgttccagaactgggttcttgatcgacgggtg  
 gcaaggctatttctttttcattgttatgcttggtttctattacaatgtgtctctgtaagaacaataagctgcata  
 gtactagtcctatgctgcaggtttggcttttgaattcaagtctatattgtgtttataat

**Anexo 02:** Alinhamento das sequências de aminoácidos deduzidas dos genes *EF1α* membros da família multigênica obtidas através do CLUSTAL W.

```

Gm EF1α 1a      MGKEKVHINIVVIGHVDSGKSTTTGHLIYKLGIDKRVIERFEKEAAEMNKRSFKYAWVL 60
Gm EF1α 1b      MGKEKTHINIVVIGHVDSGKSTTTGHLIYKLGIDKRVIERFEKEAAEMNKRSFKYAWVL 60
Gm EF1α 1c      MGKEKVHINIVVIGHVDSGKSTTTGHLIYKLGIDKRVIERFEKEAAEMNKRSFKYAWVL 60
Gm EF1α 3       MGKEKVHINIVVIGHVDSGKSTTTGHLIYKLGIDKRVIERFEKEAAEMNKRSFKYAWVL 60
Gm EF1α 2a      MGKEKVHISIVVIGHVDSGKSTTTGHLIYKLGIDKRVIERFEKEAAEMNKRSFKYAWVL 60
Gm EF1α 2b      MGKEKVHISIVVIGHVDSGKSTTTGHLIYKLGIDKRVIERFEKEAAEMNKRSFKYAWVL 60
                ***** . ** .*****

Gm EF1α 1a      DKLKAERERGITIDIALWKFETTKYYCTVIDAPGHRDFIKNMITGTSQADCAVLIIDSTT 120
Gm EF1α 1b      DKLKAERERGITIDIALWKFETTKYYCTVIDAPGHRDFIKNMITGTSQADCAVLIIDSTT 120
Gm EF1α 1c      DKLKAERERGITIDIALWKFETTKYYCTVIDAPGHRDFIKNMITGTSQADCAVLIIDSTT 120
Gm EF1α 3       DKLKAERERGITIDIALWKFETTKYYCTVIDAPGHRDFIKNMITGTSQADCAVLIIDSTT 120
Gm EF1α 2a      DKLKAERERGITIDIALWKFETTKYYCTVIDAPGHRDFIKNMITGTSQADCAVLIIDSTT 120
Gm EF1α 2b      DKLKAERERGITIDIALWKFETTKYYCTVIDAPGHRDFIKNMITGTSQADCAVLIIDSTT 120
                *****

Gm EF1α 1a      GGFEAGISKDGQTRHALLAFTLVGKQMICCCNKMDATTPKYSKARYDEIVKEVSSYLKK 180
Gm EF1α 1b      GGFEAGISKDGQTRHALLAFTLVGKQMICCCNKMDATTPKYSKARYDEIVKEVSSYLKK 180
Gm EF1α 1c      GGFEAGISKDGQTRHALLAFTLVGKQMICCCNKMDATTPKYSKARYEEIVKEVSSYLKK 180
Gm EF1α 3       GGFEAGISKDGQTRHALLAFTLVGKQMICCCNKMDATTPKYSKARYDEIVKEVSSYLKK 180
Gm EF1α 2a      GGFEAGISKDGQTRHALLSFTLVGKQMICCCNKMDATTPKYSKARYDEIVKEVSSYLKK 180
Gm EF1α 2b      GGFEAGISKDGQTRHALLSFTLVGKQMICCCNKMDATTPKYSKARYDEIVKEVSSYLKK 180
                *****:*****:*****:*****:

Gm EF1α 1a      VGYNPDKIPFVPIISGFEGDNMIERSTNLDWYKGPPTLLEALDQINEPKRPSDKPLRLPLQD 240
Gm EF1α 1b      VGYNPDKIPFVPIISGFEGDNMIERSTNLDWYKGPPTLLEALDQINEPKRPSDKPLRLPLQD 240
Gm EF1α 1c      VGYNPDKIPFVPIISGFEGDNMIERSTNLDWYKGPPTLLEALDQINEPKRPSDKPLRLPLQD 240
Gm EF1α 3       VGYNPDKIPFVPIISGFEGDNMIERSTNLDWYKGPPTLLEALDQINEPKRPSDKPLRLPLQD 240
Gm EF1α 2a      VGYNPDKIPFVPIISGFEGDNMIERSTNLDWYKGPPTLLEALDQISEPKRPSDKPLRLPLQD 240
Gm EF1α 2b      VGYNPDKIPFVPIISGFEGDNMIERSTNLDWYKGPPTLLEALDQISEPKRPSDKPLRLPLQD 240
                *****:*****:*****

Gm EF1α 1a      VYKIGGIGTVPVGRVETGVVKGPMVVFAPGTLTTEVKSVEMHHEALTEALPGDNVGFNV 300
Gm EF1α 1b      VYKIGGIGTVPVGRVETGVVKGPMVVFAPGTLTTEVKSVEMHHEALTEALPGDNVGFNV 300
Gm EF1α 1c      VYKIGGIGTVPVGRVETGVVKGPMVVFAPGTLTTEVKSVEMHHEALTEALPGDNVGFNV 300
Gm EF1α 3       VYKIGGIGTVPVGRVETGVVKGPMVVFAPGTLTTEVKSVEMHHEALTEALPGDNVGFNV 300
Gm EF1α 2a      VYKIGGIGTVPVGRVETGVVKGPMVVFAPGTLTTEVKSVEMHHEALTEALPGDNVGFNV 300
Gm EF1α 2b      VYKIGGIGTVPVGRVETGVVKGPMVVFAPGTLTTEVKSVEMHHEALTEALPGDNVGFNV 300
                *****:*****:*****:*****:

Gm EF1α 1a      KNVAVKDLKRGFVASNSKDDPAKEAANFTSQVIIMNHGQIGNGYAPVLDCHTSHIAVKF 360
Gm EF1α 1b      KNVAVKDLKRGFVASNSKDDPAKEAANFTSQVIIMNHGQIGNGYAPVLDCHTSHIAVKF 360
Gm EF1α 1c      KNVAVKDLKRGFVASNSKDDPAKEAANFTSQVIIMNHGQIGNGYAPVLDCHTSHIAVKF 360
Gm EF1α 3       KNVAVKDLKRGFVASNSKDDPAKEAANFTSQVIIMNHGQIGNGYAPVLDCHTSHIAVKF 360
Gm EF1α 2a      KNVAVKDLKRGYVASNSKDDPAKEAANFTAQVIIMNHGQIGNGYAPVLDCHTSHIAVKF 360
Gm EF1α 2b      KNVAVKDLKRGYVASNSKDDPAKEAANFTAQVIIMNHGQIGNGYAPVLDCHTSHIAVKF 360
                *****:*****:*****:*****:

Gm EF1α 1a      SEILTKIDRRSGKELEKEPKFLKNGDAGVMKMPVKPMVVETTFSEYPPLGRFAVRDMRQT 420
Gm EF1α 1b      SEILTKIDRRSGKELEKEPKFLKNGDAGVMKMIPTKPMVVETTFSEYPPLGRFAVRDMRQT 420
Gm EF1α 1c      AELVTKIDRRSGKEIEKEPKFLKNGDAGYVKMIPTKPMVVETTFSEYPPLGRFAVRDMRQT 420
Gm EF1α 3       AELLTKIDRRSGKELEKEPKFLKNGDAGLVKMIPTKPMVVETTFSEYPPLGRFAVRDMRQT 420
Gm EF1α 2a      AELMTKIDRRSGKELEKEPKFLKNGDAGFVKMIPTKPMVVETTFSEYPPLGRFAVRDMRQT 420
Gm EF1α 2b      AELMTKIDRRSGKELEKEPKFLKNGDAGFVKMIPTKPMVVETTFSEYPPLGRFAVRDMRQT 420
                :*:*****:*****:*****:*****:

Gm EF1α 1a      VAVGVIKSVEKKDPTGAKVTKAAAKKK- 447
Gm EF1α 1b      VAVGVIKSVEKKDPTGAKVTKAAAKKK- 447
Gm EF1α 1c      VAVGVIKSVEKKDPTGAKVTKAAQKKK- 447
Gm EF1α 3       VAVGVIKSVEKKDPTGAKVTKAAQKKAK 448
Gm EF1α 2a      VAVGVKNVEKKDPTGAKVTKAAQKKK- 447
Gm EF1α 2b      VAVGVKNVEKKDPTGAKVTKAAQKKK- 447
                ***** .***** **

```





EF1A 2b TGGTGGTGACTTTTGCACCAACTGGACTGACAACTGAAGTCAAGTCTGTGGAGATGCACC 927  
EF1A 3 TGGTGGTTACCTTTGCACCAACTGGACTAACAACCTGAAGTTAAGTCTGTGGAGATGCACC 905  
\*\*\*\*\* \*\* \*\*\*\* \*\* \*\*\*\*\* \*\* \*\*\*\*\* \*\* \*\* \*\*\*\*\* \*\* \*\*\*\*\*

EF1A 1a ATGAGGCTCTCACAGAGGCTCTTCCAGGTGACAATGTTGGATTTAATGTGAAGAATGTTG 1010  
EF1A 1b ATGAGGCTCTCACAGAGGCTCTTCCAGGTGACAATGTTGGCTTTAATGTGAAGAATGTTG 1006  
EF1A 1c ATGAAGCTCTCCAGGAGGCCCTTCCAGGTGACAACGTAGGATTTAATGTGAAGAATGTTG 1007  
EF1A 2a ATGAAGCTCTCACAGAGGCTCTTCCCGGTGATAATGTTGGATTCAATGTTAAGAATGTTG 984  
EF1A 2b ATGAATCTCTTACAGAGGCACATCCTGGTGACAATGTGGGATTCAATGTTAAGAATGTTG 987  
EF1A 3 ATGAAACTCTCACTGAGGCCATCCTGGTGACAATGTGGGATTCAATGTTAAGAATGTTG 965  
\*\*\*\* \*\*\*\*\* \*\*\*\*\* \* \*\* \*\*\*\*\* \*\* \*\* \* \* \* \*\*\*\*\* \*\*\*\*\*

EF1A 1a CAGTCAAGGATCTCAAGCGTGGTTTTGTTGCATCCAACCTCCAAGGACGACCCTGCCAAGG 1070  
EF1A 1b CAGTCAAGGATCTCAAGCGTGGTTTTGTTGCATCCAACCTCCAAGGATGACCCTGCCAAGG 1066  
EF1A 1c CAGTCAAGGATCTCAAGCGTGGTTTTGTTGCCTCAAACCTCAAGGATGATCCTGCCAAGG 1067  
EF1A 2a CTGTTAAGGATCTCAAGCGTGGTTATGTTGCCTCGAACTCAAAGGATGATCCTGCCAAGG 1044  
EF1A 2b CTGTTAAGGATTTGAAGCGTGGTTATGTTGCCTCAAACCTCAAAGGATGACCCTGCCAAGG 1047  
EF1A 3 CTGTCAAGGATCTCAAGCGTGGTTTTGTTGCGTCTAACTCAAGGATGACCCTGCCAAGG 1025  
\* \*\* \*\*\*\*\* \* \*\*\*\*\* \*\*\*\*\* \*\* \*\*\*\*\* \*\*\*\*\* \*\* \*\*\*\*\* \*\*\*\*

EF1A 1a AAGCTGCCAACTTCACATCCCAAGTCATTATCATGAACCATCCTGGCCAGATTGGTAATG 1130  
EF1A 1b AAGCTGCCAACTTCACATCTCAAGTCATTATCATGAACCACCCTGGCCAGATTGGAAATG 1126  
EF1A 1c AGGCTGCCAACTTCACATCCCAAGTCATTATAATGAACCATCCTGGCCAAATTGGAAATG 1127  
EF1A 2a AGGCTGCTAACTTCACTGCCAGGTTATCATCATGAACCATCCTGGTCAGATTGGAAATG 1104  
EF1A 2b AGGCTGCTAACTTCACTGCCAGGTTATCATCATGAACCACCCTGGTCAGATTGGAAATG 1107  
EF1A 3 AGGCTGCTAACTTCACTGCCAGGTTATCATCATGAACCACCCTGGACAGATTGGAAATG 1085  
\* \*\*\*\*\* \*\*\*\*\* \* \*\* \* \* \* \*\*\*\*\* \*\*\*\*\* \*\* \*\*\*\*\* \*\*\*\*

EF1A 1a GATACGCACCAGTCCCTTGACTGCCACACTTCTCACATTGCTGTGAAGTTTTCTGAAATCT 1190  
EF1A 1b GATACGCACCGGTGCTTGATTGCCACACTTCTCACATTGCTGTCAAGTTTTCTGAAATCT 1186  
EF1A 1c GTTATGCTCCTGTGCTTGACTGCCACACTTCTCACATTGCTGTCAAGTTTGCCGAACTTG 1187  
EF1A 2a GCTATGCCCTGTCTTGACTGCCACACTTCCACATTGCTGTCAAGTTTGCTGAACTCA 1164  
EF1A 2b GCTATGCCCTGTCTCGACTGCCACACTTCTCACATTGCTGTCAAGTTTGCTGAACTCA 1167  
EF1A 3 GCTATGCACCTGTCTTGATTGCCACACTCCACATTGCCGTCAAGTTTGCTGAACTCT 1145  
\* \*\* \* \* \* \* \* \* \* \*\*\*\*\* \*\* \*\*\*\*\* \*\* \*\*\*\*\* \* \* \* \*

EF1A 1a TGACCAAGATTGACAGGCGATCTGGTAAGGAGCTTGAGAAGGAGCCCAAATTTTTGAAGA 1250  
EF1A 1b TGACCAAGATTGACAGGCGATCTGGTAAGGAGCTTGAGAAGGAGCCCAAATTTTTGAAGA 1246  
EF1A 1c TGACCAAGATTGACAGGCGATCTGGTAAGGAGATTGAGAAGGAGCCCAAATTTCTGAAAA 1247  
EF1A 2a TGACCAAGATTGACAGGCGATCTGGCAAAGAGCTTGAGAAGGAACCCAAAGTTTTTTGAAGA 1224  
EF1A 2b TGACCAAGATTGACAGGCGATCCGGCAAAGAGCTTGAGAAGGAGCCCAAAGTTTTTTGAAGA 1227  
EF1A 3 TGACCAAGATTGACAGGCGATCTGGTAAGGAACCTTGAGAAGGAGCCCAAATTTCTGAAAGA 1205  
\*\*\*\*\* \*\*\*\*\* \*\* \*\* \* \* \*\*\*\*\* \*\*\*\*\* \*\* \*\*\*\*\* \*

EF1A 1a ATGGTGATGCTGGTATGGTTAAGATGGTTCCAACCAAGCCCAAGGTGGTTGAAACTTTCT 1310  
EF1A 1b ATGGTGATGCAGGTATGGTTAAGATGATTCCAACCAAGCCCAAGGTGGTTGAAACTTTCT 1306  
EF1A 1c ATGGAGATGCAGGTATGTTAAGATGATTCCACCAAGCCTAAGGTGGTCGAAACTTTCT 1307  
EF1A 2a ATGGTGATGCTGGTTTTGTTAAGATGATTCCGACCAAACCCAAAGGTGGTTGAAACTTTCT 1284  
EF1A 2b ACGGTGATGCTGGTTTTGTTAAGATGATTCCAACCAAACCCAAAGGTGGTTGAAACTTTCT 1287  
EF1A 3 ATGGTGATGCTGGTTTTGGTCAAGATGATTCCACCAAGCCCAAGGTGGTTGAAACTTTCT 1265  
\* \*\* \*\*\*\*\* \*\* \*\* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\*

EF1A 1a CTGAGTATCCTCCCTTGGTTCGTTTTGCTGTGAGGGACATGCGTCAGACCGTAGCTGTTG 1370  
EF1A 1b CTGAGTATCCTCCCTTGGTTCGTTTTGCTGTGAGGGACATGCGTCAAACCTGTGGCTGTCG 1366  
EF1A 1c CCGAGTATCCTCCTCTTGGTTCGTTTTGCCGTGAGGGACATGCGACAAACTGTGGCTGTTG 1367  
EF1A 2a CTGAGTACCCCCACTTGGTTCGTTTTGCTGTGAGGGATATGCGTCAAACCTGTTGCTGTGG 1344  
EF1A 2b CCGAGTATCCTCCACTTGGTAGGTTTTGCTGTGAGGGACATGCGTCAAACCTGTTGCTGTGG 1347  
EF1A 3 CCGAGTATCCTCCACTTGGTTCGTTTTGCTGTGAGAGACATGCGTCAAACCTGTTGCTGTGG 1325  
\* \*\*\*\*\* \*\* \* \* \*\*\*\*\* \* \*\*\*\*\* \*\* \* \* \* \*\*\*\*\* \*\* \* \* \* \*\*\*\*\* \*

EF1A 1a GAGTCATCAAGAGTGTGAGAAGAAGACCCACCGGAGCCAAGGTCACAAAGGTCGCCG 1430  
EF1A 1b GAGTCATCAAGAGTGTGAGAAGAAGACCCACCGGAGCCAAGGTCACAAAGGTCGCCG 1426  
EF1A 1c GAGTCATCAAGAGTGTGAGAAGAAGGACCCAACTGGAGCCAAGGTCACAAAGGTCGCAC 1427  
EF1A 2a GAGTCATCAAGAACGTGGAGAAGAAGGATCCTACTGGAGCCAAGGTCACAAAGGTCGCC 1404  
EF1A 2b GAGTCATCAAGAACGTGGAGAAGAAGGATCCTACCGGAGCCAAGGTCACAAAGGTCGCC 1407  
EF1A 3 GAGTCATCAAGAGCGTCGAGAAGAAGGACCCACTGGAGCTAAGGTCACAAAGGTCGCAC 1385  
\*\*\*\*\* \*\*\*\*\* \*\* \*\*\*\*\* \*\* \* \* \* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \*\*



```

EF1A 1a      CCAAGAAG---AAGTGATTGCATTTTGGCAATTTTGTCTAGCACATGT-GATCAT-----C 1481
EF1A 1b      CCAAGAAG---AAGTGATTGCATTTGGGCAATTTTGTCTAGCACATGT-GATCAT-----C 1477
EF1A 1c      AGAAGAAG---AAGTGAATACATTTTGTGTTTCTTCAAGTACATGC-GATCATAGTTAC 1483
EF1A 2a      AGAAGAAG---AAGTGAAT-CGTGCGGTTTGGTTCATCAGGGGATGTCGTTTCT----- 1454
EF1A 2b      AGAAGAAG---AAGTGAAT-CGTGCGGGCTGGTTCATCAGGGGATGT----- 1450
EF1A 3       AGAAGAAGGCCAATGAAT-TTGTGCTGGATGGTTTATCAAGCAAACCTCTATTAT----- 1438
              *****  **  ***  *      *      *      *

```

```

EF1A 1a      AAC-GTGGTTTCAAAAAAAGTTGTTCCCTTTAC-AGTAGTTT---ATCCTTGCAGAGTCTT 1536
EF1A 1b      ATC-GTGGTTAC-----TCCTTTAT-AGTAGTTT---ATCCTTGCAGAGTCTT 1521
EF1A 1c      ATTTATAGTTACACACAACAGTGGTTCCTTATTAGTATTTT---ATTTTTGAGAGTAT- 1539
EF1A 2a      ---TATGGTTACAATAAATGTTGGTTCCTT-----GCCCTTGTGTCTTCGT---TTCT- 1501
EF1A 2b      -----TGGTTACAATAAATGTTGGTTCCTTTTCTGTACTCTTGTGTCTTCTT---TTCT- 1501
EF1A 3       ---TATGACTACAATAAATG---GTTTCTTGTTCGTAGTACTTCATTTTTTTC---GTGT- 1488
              *  *  *      *  **      *  *      *  *

```

```

EF1A 1a      AGGTG-TTTGTTTACC--AGTTATATTTTGAAGTGTCCGCCGATTCAT-GTAGCCG-- 1590
EF1A 1b      AGGTG-TTTTGTFTA---AGTTATATTTTTAAGTTTCTGCCGATTCAT-GTAGCCG-- 1573
EF1A 1c      AGGTGCTTTTACTTGTATAGTTTACTTTTATGTATCAGCC-ATTCATTGCAGCCACC 1598
EF1A 2a      AGGTA-GCTTGTTTTTT--GGACATAGTTTGAAGTCTCCACC-ATCATCTCGCAACTT-- 1555
EF1A 2b      AGGTA-GCTTGTTTTTT--GGACAAAGTTTGAAGTCTCCACC-ATCATCTCGCAACTG-- 1555
EF1A 3       AGGGA-ACTTGTTTATC--AGATTTTGTGTCTCAGGTTTGTCCGATCTCGCAACTT-- 1542
              ***  *  **      *      *  *  *  *  *  *  *  *  *

```

```

EF1A 1a      TAGCCTTCAAAGTGGGTTCTTGATCGGCGGTAAC---ATT-TTCGTTGCTGTTTGT-TT 1645
EF1A 1b      TAACTTTCAAAGTAGGTTCTTGATCGGCGGTGGTC-AATT-TTCATTGCTGTTTGT-TT 1630
EF1A 1c      TACCTTTCAAAGTGGGTTCTTGATGGGCGGTGGT--GACT-TT-----TCTAT-TT 1646
EF1A 2a      TTGTTCCCAGAATTTGGGTTCTTGATCGACGGTGGCAAGACTCCTTTTATCATTCTGT--T 1613
EF1A 2b      TTGTTCCCAGAAGTGGGTTCTTGATCGACGGTGGCAAAATGCTTTTATTTATCTGTGT 1615
EF1A 3       TTGTTCCCAGAAGTGGGTTCTTGATCGACGGTGGCAAGGCTATTTCTTTTTCATGT--- 1599
              *      **  **  *  *****  *  ***      *  *      *  *

```

```

EF1A 1a      TTGATGAGTACTGTTTTTGTTTTGTATGGTAAAA---GTCTGAGAT-TTTCAAATT-CAC 1700
EF1A 1b      TTGATGAGTACTGTTTTTGTTTTGTATGGTAAGAAGAGTCTGAGATATTTCGAATTTAC 1690
EF1A 1c      T--ATAAG--CTGT-----ATGA-----GAACTCTTT---TTGCAC 1675
EF1A 2a      TTAATGTG--TTGTGTTT-----GTGAGAACCCTGATTACATTTTGTAA--GCGC 1662
EF1A 2b      TTAATGTG--TTGTGTTT-----GTGAGAACCCTGATTACATTTTGTAA--GCGC 1664
EF1A 3       --TATGCT--TTGTTTCT-----ATTACAA-----TGTGTCTCTGT-AA--GAAC 1637
              **      ***      *      *

```

```

EF1A 1a      AAGCAGCCATAGGGTTTTAGTCCATTTCCCTTTGCTGCTGAGGAGGGATGTCTTAAATTTG 1760
EF1A 1b      AAGCAGCTATAGGGTTTTAGTCCATTTCCCTTTGCTGCTGA---GGGATGTTTTAAGTT-G 1746
EF1A 1c      AAGCAGCTACAGGTTTTAGTT----- 1697
EF1A 2a      AGCGAGTTTTAGGGCTTTGC-----CGTTGCGTTGTTGG---TTTGCTTTTTAAATG-- 1711
EF1A 2b      AGCGAGTTTCAGGACTTTGCTG---CGTTGTGTGCT-----TTGGTTTATTAATG-- 1713
EF1A 3       AATAAGCTGCATAGTACTAG-----TCCTATGCTGCAGG---TTTGGCTTTTGAATT-- 1686
              *  **  *      *

```

```

EF1A 1a      CATTTAATTTATAAGGAAGTTTTGTTAACTGTT----- 1793
EF1A 1b      CATTTAATTTATAACGAAGTTTTATAAACTGTTTATGGTTTTAAAGGCTATTATTCCTTT-- 1804
EF1A 1c      -----
EF1A 2a      --TCAACTTTATATT---TGIGTTCAATTTTTGCTTGGTTTGTCTTTAAAAAATCAAA 1764
EF1A 2b      --TCAACTTTCTATT---TGT---AGTGTTT----- 1736
EF1A 3       --TCAAGTCTATATTG---TGT---TGT-----ATAAT--- 1712

```

```

EF1A 1a      -----
EF1A 1b      -----
EF1A 1c      -----
EF1A 2a      TTTATGTG 1772
EF1A 2b      -----
EF1A 3       -----

```