



**HORLLYS GOMES BARRETO**

**ANÁLISE DOS GENES HOMÓLOGOS DE  
FLORESCIMENTO EM CAFÉ ARÁBICA**

**LAVRAS – MG**

**2011**

**HORLLYS GOMES BARRETO**

**ANÁLISE DOS GENES HOMÓLOGOS DE FLORESCIMENTO EM  
CAFÉ ARÁBICA**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Biotecnologia Vegetal, área de concentração em Biotecnologia Vegetal, para a obtenção do título de Mestre.

Orientador

PhD. Antonio Chalfun Júnior

**LAVRAS-MG**

**2011**

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da  
Biblioteca da UFLA**

Barreto, Horllys Gomes.

Análise dos genes homólogos de florescimento em café arábica /  
Horllys Gomes Barreto. – Lavras : UFLA, 2011.  
73 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2011.  
Orientador: Antonio Chalfun Júnior.  
Bibliografia.

1. *Coffea arabica*. 2. Expressão gênica. 3. Bioinformática. 4.  
MADS-box. 5. Regulação do florescimento. I. Universidade Federal  
de Lavras. II. Título.

CDD – 631.523

**HORLLYS GOMES BARRETO**

**ANÁLISE DOS GENES HOMÓLOGOS DE FLORESCIMENTO EM  
CAFÉ ARÁBICA**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-Graduação em Biotecnologia Vegetal, área de concentração em Biotecnologia Vegetal para a obtenção do título de Mestre.

APROVADA em 23 de fevereiro de 2011.

Dr. Luciano Vilela Paiva

UFLA

PhD. Vagner Augusto Benedito

WVU

PhD. Alan Carvalho Andrade

EMBRAPA

PhD. Antonio Chalfun Júnior  
Orientador

**LAVRAS – MG**

**2011**

Dedico a Deus, por acreditar em mim, mesmo com todos os meus defeitos, e fazer por mim o que tão pouco mereço.

DEDICO

## **AGRADECIMENTOS**

A Deus, pela vida e por todas as pessoas que Ele colocou em minha vida para que este sonho se tornasse realidade.

À minha família, por simplesmente ser a minha família; aos meus pais, Deuzamar Gomes Pereira e Francisca das Chagas Correia Barreto, que pela graça de Deus me deram a vida; porém, gostaria de agradecer especialmente ao meu Pai, Rosenberg Pereira de Abreu, que me proporcionou e me proporciona saber realmente o que é amor de pai. Ele, juntamente com minha mãe, são as duas figuras mais importantes da minha vida, são Jesus para mim, uma vez que estão dispostos a se crucificarem por minha vida. Só Deus e eu sabemos o quanto eles foram e são importantes. Simplesmente os amo.

Aos Professores da Universidade Federal do Tocantins (UFT), pelos ensinamentos necessários para pleitear um curso de mestrado, em especial aos professores Alan Barbiero, Valéria Momenté, Eduardo Erasmo e Jacinto Santos, que me incentivaram a seguir a carreira científica e acompanharam-me durante o mestrado. Além do Professores, quero deixar minha gratidão a meu amigo Adeilson (Motorista – UFT), de quem também contei com o apoio.

À Universidade Federal de Lavras (UFLA) e ao Programa de Pós-graduação em Biotecnologia Vegetal, pela oportunidade concedida para a realização do mestrado, em especial ao professor Luciano Paiva, que brilhantemente tem conduzido o programa, “oxigenando” todas as pesquisas que são desenvolvidas no Laboratório Central de Biologia Molecular (LCBM).

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), pela concessão da bolsa de estudos.

Aos professores do Programa de Pós-graduação em Biotecnologia Vegetal, pelos ensinamentos transmitidos e harmoniosa convivência, em especial aos professores Chalfun Júnior, Luciano Paiva e Alan Andrade, pela amizade e total apoio no desenvolvimento do meu trabalho de dissertação.

Ao Professor Chalfun Júnior, pela oportunidade, orientação, paciência, dedicação, amizade, companheirismo e ensinamentos, que foram de grande importância para a realização deste trabalho e meu crescimento pessoal e profissional. Deixo aqui não só os agradecimentos, mas também meus parabéns, pois o professor sempre tratou todos os integrantes do grupo de pesquisa igualmente, independentemente do nível de conhecimento, e tem buscado sempre o melhor para o grupo mesmo quando isso não vai de encontro aos interesses pessoais. O meu muito obrigado.

A West Virginia University (WVU), pela oportunidade de fazer um treinamento em Biologia Molecular durante o curso de mestrado, em especial aos professores Luciano Paiva, Chalfun Júnior e Vagner Benedito, pela confiança e credibilidade.

Ao professor Vagner Benedito, com quem trabalhei e aprendi bastante durante os últimos seis meses do mestrado, o qual sempre esteve muito próximo, “oxigenando” os trabalhos que desenvolvi na West Virginia University, porém, foi mais que um côorientador, também foi amigo e companheiro durante minha estada nos Estados Unidos.

Por fim, aos amigos Fabrício Lelis e Eula Tamiris, em nome de todos do Laboratório Central de Biologia Molecular (LCBM), aos amigos Anderson Tadeu, Márcia Fabiana, Solange Ságio, Andre Lima e Gabriel Alves, em nome de todos aqueles que contribuíram para o desenvolvimento deste trabalho, aos amigos Anderson Tadeu e Luiz Gustavo, pela irmandade, às amigas Solange Ságio e Kamila Dázio, em nome do Grupo de Pesquisa do Laboratório de Fisiologia Molecular (LFM) e aos amigos do Programa de Pós-graduação em Fisiologia Vegetal. Finalmente o meu obrigado a todos aqueles que contribuíram diretamente ou indiretamente para o desenvolvimento deste trabalho.

## RESUMO GERAL

Apesar dos vários estudos que abordaram a fisiologia e os processos de desenvolvimento do café, muitos aspectos diretamente relacionados com a qualidade do café ainda requerem uma melhor compreensão, como o processo de regulação do florescimento. A maturação desuniforme dos frutos do cafeeiro, gerada pelo florescimento sequencial encontrado nessa espécie, afeta diretamente a sua produtividade e leva a aumentos nos custos de produção e também a uma bebida de menor qualidade. Assim, objetivando uma melhor compreensão do processo de florescimento do café (*Coffea arabica*), realizou-se neste trabalho a identificação e caracterização dos genes *FLOWERING LOCUS C*, *FRIGIDA*, *REDUCED VERNALIZATION RESPONSE 2* e *VERNALIZATION INSENSITIVE 3* em café. Análises *in silico* realizadas no banco de ESTs (*Expressed Sequence Tags*) CAFEST permitiram a identificação dos genes estudados, os quais tiveram suas sequências confirmadas. As análises de expressão realizadas pela técnica da PCR quantitativa em tempo-real (qPCR) mostraram que os quatro genes foram expressos em todos os tecidos analisados (raízes, folhas de plantas submetidas ao estresse hídrico, folhas, e flores) e sugeriram que todos estejam envolvidos na regulação do florescimento em café. Estudos futuros que envolvam mutantes para esses genes em plantas-modelo e em café possibilitarão um melhor entendimento das funções desempenhadas por cada um desses genes no florescimento, assim como uma melhor compreensão desse processo como um todo.

Palavras-chave: Bioinformática. *Coffea arabica*. Expressão gênica. *MADS-box* Regulação do florescimento.

## GENERAL ABSTRACT

Although extensive studies have been carried out on coffee physiology and developmental processes, many aspects directly associated to coffee quality require better understanding, such as the flowering regulation process. The asynchronous fruit ripening showed by coffee trees, which is triggered by the sequential flowering found in this species, greatly affects its productivity and leads to higher production costs and a lower cup quality. Thus, in order to get a better understanding of the coffee flowering process, this work identified and characterized the coffee (*Coffea arabica* L.) *FLOWERING LOCUS C*, *FRIGIDA*, *REDUCED VERNALIZATION RESPONSE 2* e *VERNALIZATION INSENSITIVE 3* in coffee. *In silico* analysis in the coffee expressed sequence tag database (CAFEST) database allowed the identification of the genes studied in this work, which had their identity confirmed by sequencing. Expression analysis performed by quantitative real-time PCR (qPCR) showed that the four genes studied are expressed in all tissues analyzed (roots, leaves of plants subjected to water stress, leaves and flowers), and suggested that they are directly involved in the coffee flowering process. Further functional studies comprising mutants for these genes in a model species and in coffee trees will lead to a better understanding of the specific functions displayed by each one of these genes in the coffee flowering induction, as well as a better understanding of this process as a whole.

Keywords: Bioinformatics. *Coffea arabica*. Flowering regulation. Gene expression. *MADS-box*.

## SUMÁRIO

	<b>CAPÍTULO 1 Fenologia do florescimento do cafeeiro e genética do florescimento .....</b>	
<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>14</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO.....</b>	<b>16</b>
<b>2.1</b>	<b>Fenologia do florescimento do cafeeiro.....</b>	<b>16</b>
<b>2.2</b>	<b>Genética do florescimento.....</b>	<b>19</b>
<b>3</b>	<b>CONSIDERAÇÕES GERAIS.....</b>	<b>23</b>
	<b>REFERÊNCIAS .....</b>	<b>24</b>
	<b>CAPÍTULO 2 Análise <i>in silico</i> dos genes <i>FLC</i>, <i>FRI</i>, <i>VRN2</i> e <i>VIN3</i> em café arábica.....</b>	<b>28</b>
<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>31</b>
<b>2</b>	<b>MATERIAL E MÉTODOS.....</b>	<b>33</b>
<b>2.1</b>	<b>Análise <i>in silico</i>.....</b>	<b>33</b>
<b>2.2</b>	<b>Análise filogenética.....</b>	<b>34</b>
<b>2.3</b>	<b>Identificação de <i>motifs</i> comuns de agrupamento.....</b>	<b>34</b>
<b>2.4</b>	<b><i>Northern</i> Eletrônico .....</b>	<b>35</b>
<b>3</b>	<b>RESULTADOS E DISCUSSÃO.....</b>	<b>36</b>
<b>3.1</b>	<b>Análises de <i>motifs</i> comuns e filogenéticas.....</b>	<b>36</b>
<b>3.1.1</b>	<b><i>FLC</i>.....</b>	<b>36</b>
<b>3.1.2</b>	<b><i>FRI</i>.....</b>	<b>39</b>
<b>3.1.3</b>	<b><i>VRN2</i>.....</b>	<b>41</b>
<b>3.1.4</b>	<b><i>VIN3</i>.....</b>	<b>42</b>
<b>3.2</b>	<b><i>Northern</i> eletrônico.....</b>	<b>44</b>
<b>3.2.1</b>	<b><i>FLC</i>, <i>FRI</i>, <i>VRN2</i> e <i>VIN3</i>.....</b>	<b>44</b>
<b>4</b>	<b>CONCLUSÃO.....</b>	<b>48</b>
	<b>REFERÊNCIAS.....</b>	<b>49</b>
	<b>CAPÍTULO 3 Sequenciamento e análise quantitativa dos genes <i>FLC</i>, <i>FRI</i>, <i>VRN2</i> e <i>VIN3</i> em café arábica.....</b>	<b>53</b>
<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>55</b>
<b>2</b>	<b>MATERIAL E MÉTODOS.....</b>	<b>57</b>
<b>2.1</b>	<b>Coleta de material vegetal.....</b>	<b>57</b>
<b>2.2</b>	<b>Extração de RNA e síntese de cDNA.....</b>	<b>57</b>
<b>2.3</b>	<b>Sequenciamento.....</b>	<b>58</b>
<b>2.4</b>	<b>Análises das sequências em banco de dados.....</b>	<b>59</b>

2.5	qPCR.....	59
3	RESULTADOS E DISCUSSÃO.....	62
3.1	Sequenciamento.....	62
3.2	qPCR.....	67
4	CONCLUSÃO.....	70
5	CONSIDERAÇÕES FINAIS.....	71
	REFERÊNCIAS.....	72

## RESUMO

O café é uma das mercadorias mais comercializadas no mundo. *Coffea arabica* L. e *Coffea canephora* L. são as espécies mais importantes economicamente. *Coffea arabica* L. é uma espécie tropical caracterizada por apresentar uma floração gregária, isto é, todas as plantas, em uma certa extensão geográfica, florescem simultaneamente. De acordo com as condições ambientais, o cafeeiro pode apresentar números variáveis de floradas dentro de um mesmo ano, podendo variar de poucas floradas, em regiões com estações secas bem definidas (Brasil), a várias floradas, comum em regiões equatoriais (Colômbia). Dessa forma, há um grande interesse em melhor compreender os diferentes fatores que regulam esse processo, objetivando um florescimento mais uniforme. O estudo de genes relacionado à regulação do florescimento pode fornecer as bases para estudos aplicados que possam melhorar a cadeia produtiva do café e permitir um controle do tempo de florescimento melhor e, assim, aumentar a produtividade do café.

Palavras-chave: *Coffea arabica*. Florescimento sequencial. MADS-box. Regulação do florescimento. Vernalização.

## ABSTRACT

Coffee is the most traded commodity in the worldwide. Among the cultivated species, *Coffea arabica* L. and *Coffea canephora* L. are the most economically important ones. *Coffea arabica* L. is a tropical species characterized by a gregarious flowering process, in which all flowers within a certain geographical extension bloom simultaneously. According to environmental factors, coffee trees may display different numbers of flowering events through the same year, ranging from a few in regions with a well defined dry season, such as Brazil, to several flowering events, common in equatorial rainy regions, like Colombia. Thus, there has been a great interest to understand the different factors that regulate this process, aiming to promote a uniform flowering. The study of genes related to flowering regulation can provide the basis for applied studies to improve the coffee production chain and allow a better control of the flowering time, thus increasing coffee productivity.

Keywords: *Coffea arabica*. Flowering sequence. MADS-box. Regulation of flowering. Vernalization.

## **CAPÍTULO 1**

### **Fenologia do florescimento do cafeeiro e genética do florescimento**

## 1 INTRODUÇÃO

O gênero *Coffea* pertence à família Rubiaceae e apresenta cerca de 100 espécies identificadas (DAVIS et al., 2006). A maioria delas são árvores e arbustos crescendo a baixas latitudes nas florestas tropicais da África e Ásia. Todas as espécies conhecidas desse gênero são diploides ( $2n = 22$ ) e apresentam fecundação cruzada obrigatoriamente com sistema de autoincompatibilidade. Uma exceção a essas características é a espécie *Coffea arabica*, que é alotetraploide ( $2n=4x=44$ ) e apresenta autofecundação em aproximadamente 90% (CHARRIER; BERTHAUD, 1985). Essa espécie é a que apresenta maior valor comercial.

O café é um dos principais produtos agrícolas, sendo considerado uma das principais *commodities* internacional. Das espécies cultivadas, *Coffea arabica* L. e *Coffea canephora* L. são as mais importantes economicamente. *Coffea arabica* L. é uma espécie tropical de floração gregária, isto é, todas as plantas, em uma certa extensão geográfica, florescem simultaneamente, com surtos de floradas variáveis, desde poucos nas regiões cafeeiras de latitudes médias, com época seca bem definida (Brasil), até vários ao longo do ano, em regiões equatoriais chuvosas (Colômbia) (GUIMARÃES; MENDES, 1998). Dessa forma, há um interesse em desenvolver estudos que contribuam para uma floração mais uniforme, mediante compreensão da interação de fatores que afetam essa fase de desenvolvimento da planta.

Sabe-se que o início da floração depende da expressão equilibrada de uma complexa rede de genes, que é regulada por fatores endógenos e ambientais. No entanto, na planta do café, alguns recursos adicionais são associados, tais como o ciclo bienal, gemas dormentes, indução da floração por estresse hídrico, maturação sequencial de botões florais (CAMARGO, 1985; RENA; MAESTRI, 1985; PEZZOPANE et al., 2008), e que dificulta a

compreensão desse processo. Assim, embora os fatores ambientais tenham sido amplamente estudados, muito pouco é conhecido sobre o processo de regulação gênica do florescimento nessa espécie.

## 2 REFERENCIAL TEÓRICO

### 2.1 Fenologia do florescimento do cafeeiro

O ciclo fenológico do cafeeiro (*Coffea sp*) apresenta duas fases que ocorrem simultaneamente: vegetativa e reprodutiva. As plantas de café são classificadas como sendo bienais, ou seja, demoram dois anos para completar o ciclo, diferentemente da maior parte das plantas, que florescem e frutificam no mesmo ano fenológico.

Com o intuito de facilitar a descrição dessas duas fases, Camargo e Camargo (2001, p.65) subdividiram-nas em seis fases distintas, sendo duas delas no primeiro ano fenológico, que compreende a fase vegetativa, e as quatro últimas no segundo ano fenológico ou na fase reprodutiva, adaptadas às condições climáticas do Brasil (Figura 1).

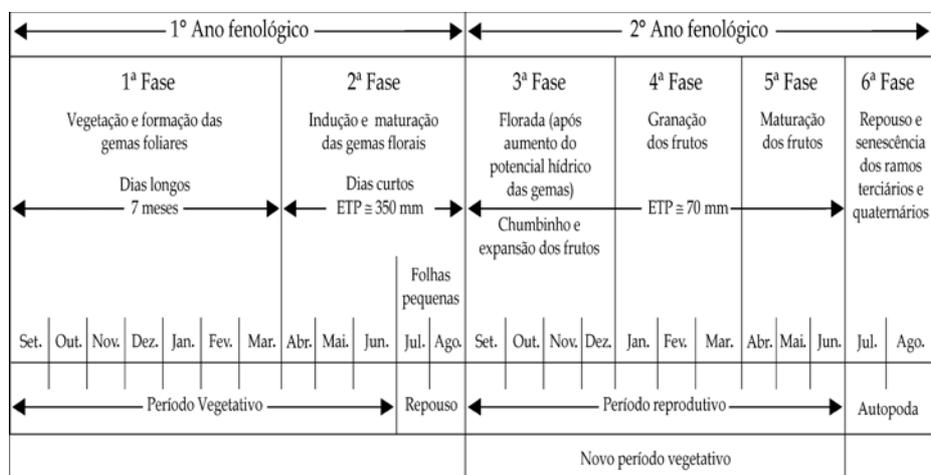


Figura 1 Fenologia do cafeeiro arábica, durante 24 meses, nas condições climáticas tropicais do Brasil (Adaptado de CAMARGO; CAMARGO, 2001)

No primeiro ano fenológico, a primeira fase vegetativa está relacionada à formação das gemas vegetativas e ocorre normalmente de setembro a março. São sete meses de dias longos, com fotoperíodo acima de 13 a 14 horas de luz efetiva (CAMARGO, 1985, p.831).

Já na segunda fase vegetativa, ocorre a maturação das gemas florais, sendo caracterizada por dias curtos, indo normalmente de abril a agosto, período durante o qual é observado um crescimento das gemas florais existentes. Após o completo desenvolvimento, entram em dormência e ficam prontas para a antese, que ocorrerá quando houver um aumento substancial de seu potencial hídrico, causado pela chuva ou irrigação. Nos dois meses finais dessa etapa, julho a agosto, as gemas dormentes produzem um par de folhas pequenas, separando o primeiro ano fenológico do segundo (GOUVEIA, 1984).

No cafeeiro, o florescimento é promovido por dias curtos e pelo choque hídrico. Sendo assim, no segundo ano fenológico, período reprodutivo, a terceira fase inicia-se com a florada após um aumento do potencial hídrico nas gemas florais maduras (choque hídrico). Após a fecundação, ocorre o processo de formação de frutos (chumbinhos) e a expansão dos frutos. Essa etapa compreende quatro meses, de setembro a dezembro (CAMARGO; CAMARGO, 2001).

Cafeeiros que recebem, na terceira fase, água com muita frequência têm a floração indefinida. Uma florada principal ocorre quando se verifica um período de restrição hídrica, seguido de chuva ou irrigação abundante (RENA; MAESTRI, 1985).

A temperatura ambiente elevada associada a um intenso déficit hídrico, durante o início da florada, provoca a morte dos tubos polínicos pela desidratação, causando o abortamento das flores, resultando nas conhecidas “estrelinhas” (CAMARGO; CAMARGO, 2001).

A quarta fase está relacionada à granação dos frutos, quando os líquidos internos solidificam-se, dando formação aos grãos, ocorrendo em pleno verão, de janeiro a março. Porém, a maturação dos frutos se dá na quinta fase, compreendendo normalmente os meses de abril, maio e junho.

Para uma melhor visualização de algumas etapas do desenvolvimento floral do cafeeiro, Morais et al. (2008) elaboraram uma escala que representa desde a formação da gema até o desenvolvimento dos frutos (Figura 2).

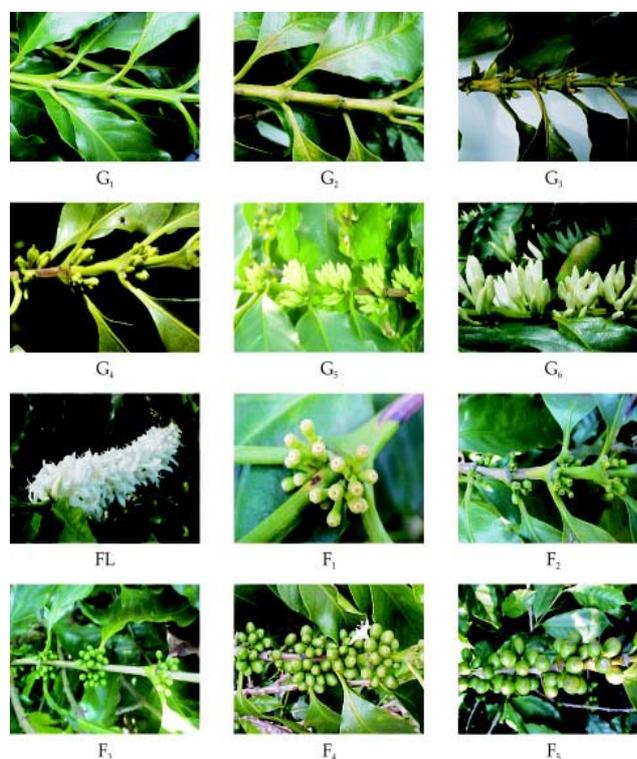


Figura 2 Fases reprodutivas do cafeeiro. G1 - refere-se aos nós com gemas indiferenciadas; G2 – nós com gemas intumescidas; G3 - gemas com até 3 mm de comprimento; G4 – gemas medindo 3,1 a 6 mm de comprimento; G5 – gemas de 6,1 a 10 mm (coloração verde-clara); G6 – gema maior que 10 mm (coloração branca); FL – abertura do botão floral (antese); F1 – F5 diferentes fases de maturação dos frutos. Adaptado de Morais et al. (2008)

A passagem da gema à fase reprodutiva (indução) compreende uma sequência de eventos de natureza bioquímica e morfológica que começa com o estímulo indutivo. Em *Coffea arabica*, um leve achatamento do meristema e sua elevação acima do nível original são os primeiros indícios visuais de transição floral (RENA; MAESTRI, 1985). Na sequência, são lançadas pequenas gemas que vão crescendo e ficando visivelmente diferenciadas (G1 – G6). Na fase representada como FL ocorre a antese e as fases caracterizadas como F são representações do desenvolvimento dos frutos.

Durante a fase reprodutiva, as etapas podem ocorrer mais de uma vez, ou seja, o cafeeiro arábica, além da florada principal, apresenta sucessivas floradas em razão da variação climática, da variabilidade genética e do manejo (RIBEIRO et al., 2004; GARCIA, 2006; LUNZ, 2006). Isso faz com que a maturação seja desigual, inclusive dentro da mesma florada, dificultando, assim, a colheita, que se torna uma das maiores dificuldades na produtividade e qualidade do café.

## **2.2 Genética do florescimento**

Estudos genéticos estabeleceram quatro rotas de desenvolvimento distintas que controlam o florescimento em *Arabidopsis*, permitindo a sincronização do tempo de florescimento: comprimento do dia (fotoperíodo), tratamento de frio (vernalização), via autônoma e das giberelinas (BLASQUEZ, 2000, p.3547). Sendo assim, o início do florescimento é induzido pelo cruzamento dessas distintas vias genéticas (Figura 3).

Três classes de genes que regulam o desenvolvimento floral foram identificadas por mutações: genes de identidade de órgãos florais, genes cadastrais e genes de identidade de meristemas.

Genes de identidade de meristemas controlam a transição da fase vegetativa para meristemas florais. Dentro dos meristemas florais, genes cadastrais marcam o limite da função de genes de identidade floral, definindo, então, os diferentes verticilos florais. Genes de identidade de órgãos florais (genes seletores homeóticos) especificam a identidade do órgão dentro de cada verticilo da flor, pela ativação de “genes realizadores” (THEISSEN et al., 2000).

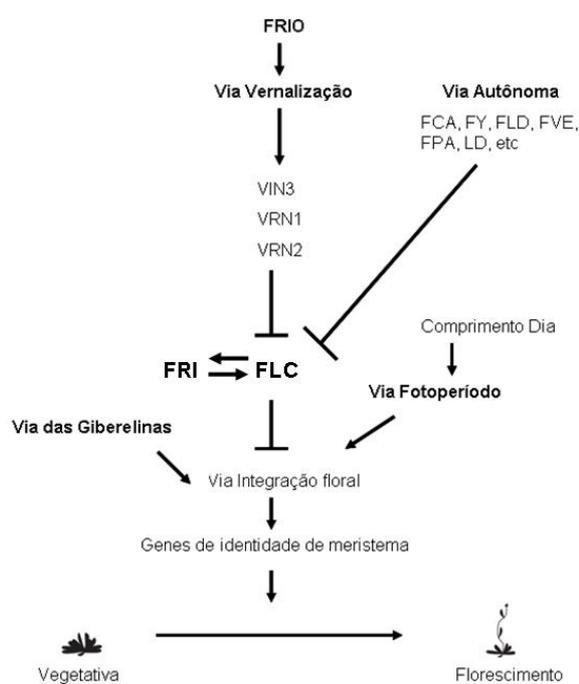


Figura 3 Vias de controle do tempo de florescimento em *Arabidopsis* (Ciannamea, 2007)

Muitos dos genes que determinam a identidade dos órgãos florais são *MADS-box* (SCHWARZ-SOMMER et al., 1990). Genes *MADS-box* constituem uma família de fatores de transcrição que atuam como reguladores-chave em

muitos processos no desenvolvimento celular de diversos organismos, tais como fungos, plantas e animais. Esses genes codificam um domínio de ligação de DNA dos respectivos fatores de transcrição do domínio MADS.

Além de determinar funções homeóticas florais, genes *MADS-box* têm muitas outras funções dentro de uma rede de genes que controlam o desenvolvimento reprodutivo em angiospermas, como *Arabidopsis*, tais como a iniciação do florescimento, identidade do óvulo, desenvolvimento da semente, formação da raiz, embriogênese, desenvolvimento do fruto, identidade do meristema floral, multiplicação celular e desenvolvimento das folhas (PARENICOVA et al. 2003; RIJPKEMA et al. 2007; THEISSEN et al., 2000).

O principal gene *MADS-box* responsável pela regulação do florescimento é o *Flowering Locus C (FLC)*, que de maneira positiva e reciprocamente é regulado por *FRIGIDA (FRI)*, como também inibido pela vernalização por meio dos genes como *REDUCED VERNALIZATION RESPONSE 1 (VRN1)*, *REDUCED VERNALIZATION RESPONSE 2 (VRN2)* e *VERNALIZATION INSENSITIVE 3 (VIN3)* (CIANNAMEA, 2007).

O gene *FRI* eleva a expressão de *FLC* para níveis que inibem o florescimento, ao passo que a vernalização promove a inibição do gene *FLC* e induz o florescimento (JOHANSON et al., 2000; MICHAELS; AMASINO, 1999;). O gene *FLC* inibe a transcrição de dois genes promotores do florescimento; *Flowering Locus T (FT)* e *Supressor of Over-expression of Constans 1 (SOC1)* (LEE et al., 2000; MICHAELS et al., 2005). Em plantas não-vernalizadas, *FLC* é expresso em altos níveis e sua proteína liga-se a regiões intrônicas do gene *FT* e à região promotora do gene *SOC1* para reprimir a transcrição de ambos os genes (HELLIWELL et al., 2006).

A vernalização inibe a transcrição do gene *FLC* (MICHAELS; AMASINO, 1999; SHELDON et al., 1999) e isso permite a expressão do gene *FT* em condições de dias longos, o qual promove a transição para o

desenvolvimento reprodutivo. O grau com que o gene *FLC* é reprimido depende da duração de exposição das plantas ao frio, permitindo mudanças quantitativas no florescimento em resposta a diferentes períodos de vernalização (SHELDON et al., 2000).

Muitos trabalhos com elo entre tratamento de frio e genes envolvidos no florescimento têm sido feitos. Foi demonstrado que o efeito quantitativo da vernalização é mediado pelo nível de repressão de *FLC*, e que a atividade de *VRN1* e *VRN2* é necessária para a repressão eficiente de *FLC* durante o tratamento de frio; contudo, eles não são necessários para a manutenção do estado de repressão durante o crescimento depois do tratamento de frio (SHELDON et al., 2006).

Os produtos de *VRN2* e *VRN1* mantêm a repressão de *FLC*, desde que os níveis de transcritos *FLC* estejam em declínio (LEVY et al., 2002), mas não é conhecido como esse declínio é inicialmente alcançado.

A maioria dos efeitos moleculares precedentes à vernalização são indutores da expressão do domínio-PHD da proteína VIN3 (SUNG; AMASINO, 2004). *VIN3* é expresso somente a baixas temperaturas e ambos os níveis de transcritos e proteínas diminuem gradualmente com o prolongamento da exposição. Desde que quantidades suficientes da proteína VIN3 estejam presentes, o sinal de frio é traduzido a modificações na cromatina em locus *FLC*.

Embora o florescimento tem sido bastante estudado em *Arabidopsis thaliana*, e grandes esforços têm sido feitos para descrever o processo em outras espécies relacionadas, em café o conhecimento sobre a regulação do florescimento ainda é escasso.

### **3 CONSIDERAÇÕES GERAIS**

O florescimento sequencial é um grande problema que afeta tanto a produtividade quanto a qualidade do café. O estudo dos genes relacionados com a regulação do florescimento podem proporcionar a identificação de genes de interesse para estudos aplicados que possam melhorar a cadeia produtiva do café, o que pode permitir um controle do tempo de florescimento e aumento da produtividade.

## REFERÊNCIAS

BLAZQUEZ, M.A. Flower development pathways. **Journal of Cell Science**, London, v.113, n.20, p.3547-3548, Oct. 2000.

CAMARGO, A.P. Florescimento e frutificação do café arábica nas diferentes regiões cafeeiras do Brasil. **Pesquisa Agropecuária Brasileira**, Brasília, v.20, n.7, p.831-839, jul. 1985.

CAMARGO, A.P.; CAMARGO, M.B.P. Definição e esquematização das fases fenológicas do cafeeiro arábica nas condições tropicais do Brasil. **Bragantia**, Campinas, v.60, n.1, p.65-68, jan. 2001.

CHARRIER, A.; BERTHAUD, J. Botanical classification of coffee. In: CLIFFORD, M.N.; WILLSON, K.C. (Ed.). **Coffee: botany, biochemistry and production of beans and beverage**. London: Croom Helms, p.167-195, 1985.

CIANNAMEA, S. **Molecular characterization of the vernalization response in *lilium perenne***. 2007. 147p. Tese (Doctor in Plant Science) – Wageningen University, Wageningen.

DAVIS, A. et al. An annotated checklist of the genus *Coffea* L. (Rubiaceae). **Botanical Journal of the Linnean Society**, Malden, v.152, p. 465–512, Dec. 2006.

GARCIA, A.L.A.; PADILHA, L.; FAGUNDES, A.V. Índice de abortamento ao longo da frutificação influenciado pela face de exposição ao sol das plantas de *C. arábica*. In: CONGRESSO BRASILEIRO DE PESQUISAS CAFEEIRAS, 32., 2006, Poços de Caldas. **Anais...** Poços de Caldas: SBPC, 2006. p.67.

GOUVEIA, N.M. **Estudo da diferenciação e crescimento de gemas florais de *Coffea arabica* L:** observações sobre a antese e maturação dos frutos. 1984. 237f. Dissertação (Mestrado em Fisiologia)-Universidade Estadual de Campinas, Campinas.

GUIMARÃES, R.J.; MENDES, A.N.G. **Morfologia e fisiologia do cafeeiro.** Lavras: UFLA, 1998. 38p.

HELLIWELL, C.A. et al. The *Arabidopsis FLC* protein interacts directly in vivo with *SOCI* and *FT* chromatin and is part of a high-molecular-weight protein complex. **The Plant Journal**, Oxford, v.46, n.2, p.183-192, Apr. 2006.

JOHANSON, U. et al. Molecular analysis of FRIGIDA, a major determinant of natural variation in *Arabidopsis* flowering time. **Science**, Washington, v.290, n.5490, p.344-347, Oct. 2000.

LEE, H. et al. The AGAMOUS-like 20 MADS domain protein integrates floral inductive pathways in *Arabidopsis*. **Genes and Development**, Norwich, v.14, n.18, p.2366-2376, Aug. 2000.

LEVY, Y.Y. et al. Multiple roles of *Arabidopsis* VRN1 in vernalization and flowering time control. **Science**, Washington, v.297, n. 5579, p.243-246, July 2002.

LUNZ, A. **Crescimento e produtividade do cafeeiro sombreado e a pleno sol.** 2006. 94p. Tese (Doutorado em Fitotecnia)-Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba.

MICHAELS, S.D.; AMASINO, R.M. FLOWERING LOCUS C a novel MADS domain that acts as a Repressor of Flowering. **The Plant Cell**, Rockville, v.11, n.5, p.949-956, May. 1999.

MICHAELS, S.D. *et al.* Integration of flowering signals in winter-annual *Arabidopsis*. **Plant Physiology**, Rockville v.137, n.1, p.149-156, Sept. 2005.

MORAIS, H. *et al.* Escala fenológica detalhada da fase reprodutiva de *coffea arabica*. **Bragantia**, Campinas, v.67, n.1, p.257-260, jan. 2008.

PARENICOVA, L. *et al.* Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: New openings to the MADS world. **The Plant Cell**, Rockville, v. 15, n.7, p.1538-1551, July. 2003.

PEZZOPANE, J.R.M. *et al.* Exigência térmica do café arábica cv Mundo Novo no subperíodo do florescimento-colheita. **Ciência e Agrotecnologia**, Lavras, v 32, p.1781-1786, dez. 2008

RENA, A.B., MAESTRI, M. Fisiologia do cafeeiro. **Informe agropecuário**, Belo Horizonte, v.11, n.126, p.26-40, jun.1985.

RIBEIRO, M.N.O. *et al.* Influência dos diferentes processos de condução da lavoura sobre a maturação dos frutos de cafeeiro (*Coffea arabica* L.). In: CONGRESSO BRASILEIRO DE PESQUISAS CAFEEIRAS, 30. 2004, São Lourenço. **Anais...** São Lourenço: EMBRAPA, 2004. p.81-82.

RIJPKEMA, A.S.; GERATS T.; VANDENBUSSCHE, M. Evolutionary complexity of MADS complexes. **Current Opinion in Plant Biology**, London, v. 10, n.1, p.32-38, Feb. 2007.

SCHWARS-SOMMER, Z. *et al.* Genetic control of flower development by homeotic genes in *Antirrhinum majus*. **Science**, Washington, v.250, n.4983, p.931-936, Nov. 1990.

SHELDON, C.C. *et al.* The *FLF* MADS box gene: a repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. **The Plant Cell**, Rockville, v.11, n.3, p.445-458, Mar. 1999.

SHELDON, C.C. et al. The molecular basis of vernalization: The central role of *FLOWERING LOCUS C (FLC)*. **Proceedings of the National Academy of Science**, Washington, v.97, n.7, p.3753-3758, Mar. 2000.

SHELDON, F. et al. Quantitative effects of vernalization on *FLC* and *SOC1* expression. **Plant Journal**, Oxford, v. 45, n.2, p.871-883, Apr. 2006.

SUNG, S; AMASINO, R.M. Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. **Nature**, London, v. 427, p.159–164, Jan. 2004.

THEISSEN, G. et al. A short history of MADS-box genes in plants. **Plant Molecular Biology**, Dordrecht, v. 42, n.1, p.115–149, Jan. 2000.

## **CAPÍTULO 2**

**Análise *in silico* dos genes *FLC*, *FRI*, *VRN2* e *VIN3* em café arábica**

## RESUMO

O café é um dos principais produtos agrícolas comercializados no mundo, sendo as espécies *Coffea arabica* L. e *Coffea canephora* L. as mais importantes economicamente. O café apresenta uma floração sequencial, a qual afeta de maneira substancial a qualidade final dessa bebida. O florescimento é um estágio de desenvolvimento vegetal regulado por vários fatores endógenos e ambientais. Dessa forma, há um grande interesse em desenvolver estudos que contribuam para a obtenção de um florescimento mais uniforme em cafeeiro e, embora os fatores ambientais tenham sido amplamente estudados, pouco se sabe sobre a regulação gênica do florescimento nessa espécie. Neste capítulo verificam-se os resultados da busca no banco de *ESTs* (*Expressed Sequence Tags*) CAFEST pelos genes *FLOWERING LOCUS C* (*FLC*), *FRIGIDA* (*FRI*), *VERNALIZATION REDUCED RESPONSE 2* (*VRN2*) e *VERNALIZATION INSENSITIVE 3* (*VIN3*) de café (*Coffea arabica*). Foram realizadas buscas por palavras-chave e por similaridade no CAFEST, e as sequências apresentando similaridade confiável foram clusterizadas, anotadas e analisadas quanto à presença de domínios conservados. Alinhamentos múltiplos entre as sequências encontradas neste trabalho e sequências de proteínas dos genes *FLC*, *FRI*, *VRN2* e *VNR3* de outras espécies foram realizados pelo programa ClustalW, a filogenia entre essas sequências foi avaliada pela construção de árvores filogenéticas pelo programa MEGA4 e a identificação de motivos comuns foi avaliada pelo programa MEME (*Multiple Expectation Minimization for Motif elicitation*). O perfil de expressão *in silico* dos prováveis genes *FLC*, *FRI*, *VRN2* and *VNR3* de café foi analisado por *Northern*s eletrônicos gerados pelo programa Cluster e visualizado pelo programa TreeView. A busca no banco de dados CAFEST permitiu a identificação de fortes candidatos para todos os genes avaliados neste trabalho. A análise filogenética e de motivos conservados confirmaram a grande similaridade entre os prováveis genes *FLC*, *FRI*, *VRN2* e *VNR3* de café e as sequências desses genes de outras espécies e o *Northern* eletrônico corroborou com as funções desempenhadas por esses genes, mostrando a sua expressão durante a fase reprodutiva.

Palavras-chave: Análises filogenéticas. Bioinformática. *Coffea arabica*. *MADS-box*.

## ABSTRACT

Coffee is one of the most commercialized agricultural products in the world and *Coffea arabica L.* and *Coffea canephora L.* are the two most important coffee species. Coffee has a sequential flowering process, which substantially affects its cup quality. Flowering is a plant developmental stage regulated by several endogenous and environmental factors. Thus, there is a great interest in developing studies aiming at achieving a more uniform flowering of coffee trees and, although environmental factors have been extensively studied, little about flowering gene regulation is known in this species. This chapter reports the results of the search in the coffee expressed sequence tag CAFEST database for coffee (*Coffea arabica L.*) homologs of *FLOWERING LOCUS C (FLC)*, *FRIGIDA (FRI)*, *VERNALIZATION REDUCED RESPONSE 2 (VRN2)* and *VERNALIZATION INSENSITIVE 3 (VIN3)* genes. Keyword and similarity searches were performed in the CAFEST, and the sequences showing reliable similarity were clustered, annotated and analyzed for the presence of conserved domains. Multiple alignments comprising the sequences found in this work and FLC, FRI, VRN2 and VNR3 protein sequences from other species were performed by the ClustalW program, their phylogeny was assessed by the MEGA 4 software and the identification of common motifs among these sequences was assessed by MEME (*Multiple Expectation Minimization for Motif Elicitation*) program. The expression profile of putative coffee *FLC*, *FRI*, *VRN2* and *VNR3* homologs was analyzed by *in silico* Northern performed with the Cluster program and visualized with TreeView. The search in the CAFEST database allowed identification of strong candidates for all the genes studied in this work. The phylogenetic and motif analyses confirmed the high similarity between the putative coffee FLC, FRI, VRN2 and VRN3 homologs and the protein sequences of these genes from other species. The electronic northern corroborated the proposed function of these genes in flowering from their expression pattern during the reproductive phase.

Keywords: Bioinformatics. *Coffea arabica*. MADS-box. Phylogenetic analysis.

## 1 INTRODUÇÃO

O cafeeiro (*Coffea* sp.), pertencente à família Rubiaceae, é responsável pela produção da segunda bebida mais consumida pela população brasileira, o café, perdendo apenas para a água (TNS INTERSCIENCE, 2008). O Brasil ocupa o primeiro lugar na produção e exportação dessa cultura, com 48.094,8 mil sacas de 60 kg/ano e 34,28% respectivamente, com área plantada de 2.076,6 hectares e produtividade média de 23,16 sacas/ha (COMPANHIA NACIONAL DE ABASTECIMENTO, 2011).

Embora existam cerca de 100 espécies de cafeeiro, apenas duas têm importância econômica: *Coffea arabica* e *Coffea canephora*, com destaque para o primeiro, que corresponde cerca de 76,57% da produção nacional, e tem como maior produtor o Estado de Minas Gerais, com 67,63% (24.903,0 mil sacas) de café beneficiado (COMPANHIA NACIONAL DE ABASTECIMENTO, 2011).

Em relação à produção, mesmo o cafeeiro sendo uma cultura bastante estudada, ainda existem problemas com a produtividade, reflexo principalmente do florescimento sequencial, que causa uma maturação desuniforme dos frutos, dificultando a colheita e causando perdas de produção. Sincronizar o florescimento e amadurecimento dos frutos pode contribuir para reduções no custo de produção.

A compreensão e identificação do mecanismo de florescimento, assim como os genes envolvidos nos processos de regulação floral, são de fundamental importância para elucidar como esses mecanismos ocorrem em café. Vários genes já foram identificados em outras espécies como peças-chave na regulação molecular do florescimento.

Em *Arabidopsis thaliana*, estudos genéticos e moleculares estabeleceram a presença de quatro rotas distintas que controlam e sincronizam o desenvolvimento floral: fotoperíodo, via autônoma, via das giberelinas e vernalização (BLASQUEZ, 2000, p.3547). A vernalização é controlada por dois

locus principais: *FLOWERING LOCUS C (FLC)* e *FRIGIDA (FRI)* (SHINDO et al., 2005), sendo o *FLC* o maior repressor floral da via de integração floral em *Arabidopsis* e, quando ativo, conduz a ativação de genes de identidade de meristemas.

De maneira positiva e reciprocamente, a expressão de *FLC* é regulada por *FRIGIDA*, como também pela vernalização por meio dos genes como REDUCED VERNALIZATION RESPONSE 1 (*VRN1*), REDUCED VERNALIZATION RESPONSE 2 (*VRN2*) e VERNALIZATION INSENSITIVE 3 (*VIN3*) (CIANNAMEA, 2007).

Hoje, com o sequenciamento do transcriptoma do cafeeiro (VIEIRA et al., 2006) e a criação do banco CAFEST de ESTs (*Expresssed Sequence Tags*), tornou-se possível a realização de buscas por sequências formadoras dos prováveis genes relacionados às características de interesse. Dessa forma, estudos comparativos de genômica funcional surgem como ferramentas poderosas para a identificação de genes envolvidos na regulação de vias metabólicas. Esses estudos, relacionados à expressão gênica no cafeeiro, são ainda escassos.

Diante disso, com o presente trabalho objetivou-se identificar os genes *FLC*, *FRI*, *VRN2* e *VIN3* em café arábica, mediante uso de ferramentas de bioinformática.

## 2 MATERIAL E MÉTODOS

### 2.1 Análise *in silico*

As sequências dos genes *FLC*, *FRI*, *VRN2* e *VIN3* foram identificadas, tendo-se como fonte de dados o banco de ESTs (*Expressed Sequence Tags*), gerado pelo projeto Genoma Brasileiro Café (CAFEST; VIEIRA et al., 2006). Por meio da interface *Gene Project* (<http://www.lge.ibi.unicamp.br/cafe/>), foi possível a procura por *reads*, montagem de *clusters* e análise dos prováveis genes.

Inicialmente, foi feita uma busca por palavra-chave, tendo em vista que todos os *reads* são previamente anotados automaticamente por comparação com o banco de genes do NCBI (National Center for Biotechnology: <http://www.ncbi.nlm.nih.gov/>).

Uma segunda de busca foi realizada por meio da ferramenta BLAST (*Basic Local Alignment Search*), utilizando-se as sequências dos genes em estudo disponíveis no GenBank (<http://www.ncbi.nlm.nih.gov/>). Assim, por meio da ferramenta tBLASTN (ALTSCHUL et al., 1997), foram comparados os aminoácidos (*query*) com os nucleotídeos traduzidos do banco de dados, e selecionados todos os *reads* que apresentaram alinhamento significativo (e-value  $> 10^{-5}$ ).

Posteriormente, os reads encontrados foram agrupados (*clustering*), formando os EST-*contigs*. Todos os EST-*contigs* formados foram utilizados para uma nova busca no banco CAFEST, visando-a encontrar novas sequências dos prováveis genes, bem como corrigir clusters incompletos. Esse processo, também chamado de saturação, foi repetido até não ser encontrado nenhum novo *read* significativo.

As sequências de nucleotídeos de interesse foram comparadas com as sequências de proteínas de banco de dados públicos (NCBI) utilizando-se para

isso o algoritmo BLASTx. Para classificar a parte funcional da proteína, bem como sua matriz de leitura, foi utilizada a ferramenta ORF FINDER (*Open Reading Frame Finder*). O alinhamento da sequência da proteína de interesse com as proteínas disponíveis no NCBI foi realizado pelo algoritmo BLASTp (ALTSCHUL et al., 1997).

## 2.2 Análise filogenética

O alinhamento dos EST-contigs foi feito pelo programa ClustalW (THOMPSON et al., 1994) com os parâmetros padrões (default), utilizando-se as sequências de nucleotídeos traduzidas em aminoácidos. As sequências foram visualmente inspecionadas e manualmente corrigidas, sendo removidos segmentos cuja homologia não pode ser acertada. A árvore final foi feita utilizando-se o programa MEGA 4.0 (TAMURA et al., 2007), com o modelo de comparação *Neighbor-Joining* (SAITOU; NEI, 1987), método de distância  $p$  e supressão *pair-wise*. A validade da árvore quanto à distância filogenética dos *clusters* foi medida pelo teste probabilístico de bootstraps, originado de 100 replicatas (SITNIKOVA et al., 1995).

## 2.3 Identificação de motivos comuns de agrupamento

Para descobrir *motifs* de agrupamento entre as sequências dos genes *FLC*, *FRI*, *VRN2* e *VIN3* selecionados no CAFEST, foi utilizado o programa MEME (*Multiple Expectation Minimization for Motif Elicitation*, <http://meme.sdsc.edu/meme/meme.html>) versão 3.5.4 (BAILEY; ELKAN, 1994). A sequência total de aminoácidos das proteínas foi comparada de acordo com a ordem de agrupamento observada na árvore filogenética, facilitando, dessa forma, a observação dos *motifs* de agrupamento em cada subclasse. Os

parâmetros utilizados foram: número de repetições qualquer, máximo número de *motifs* 20 e amplitude ótima entre 6 e 200 (PARENICOVA et al., 2003).

#### **2.4 Northern eletrônico**

Para a análise espacial de expressão, foram calculadas as frequências de *reads* formadores de cada EST-*contig* nas bibliotecas em que apareceram expressos. Em seguida, os dados foram normalizados, para dar uma idéia exata do grau de expressão dos prováveis genes em cada tratamento e local da planta. O procedimento para a normalização consistiu em multiplicar cada *read* pela razão entre o número total de *reads* de todas as bibliotecas e o número de *reads* da biblioteca em que ele aparecia expresso. As sequências foram reorganizadas de acordo com os locais de expressão e seus dados foram lançados em uma matriz, relacionando genes e bibliotecas. Os EST-*contigs* e bibliotecas foram agrupados por *hierachical clustering* utilizando-se os programas Cluster e TreeView (EISEN et al., 1998). Os resultados de expressão foram apresentados em uma escala cinza, em que expressão zero ou negativa foi representada por coloração mais clara, sendo aumentada gradativamente até atingir o preto, que representa o grau máximo de expressão positiva.

### 3 RESULTADOS E DISCUSSÃO

#### 3.1 Análises de motivos comuns e filogenéticas

##### 3.1.1 FLC

A clusterização (agrupamento) realizada com os 57 *reads* encontrados no banco de dados CAFEST formaram 12 EST-*contigs* e 18 *singlets*. Após a análise das sequências proteicas, da parte funcional de cada uma delas e seus domínios conservados, foi possível observar a eficiência do método de seleção de *reads*, pois todos os 12 EST-*contigs* apresentaram o domínio conservado MADS-*box* (*motif* 1); a exceção foi observada em relação aos *singlets*, pois somente a metade (9 *singlets*) apresentou o domínio conservado para o gene em estudo (Figura 4).

Diante disso, os *singlets* que não apresentaram o domínio conservado foram excluídos das análises filogenéticas.

Análises filogenéticas entre genes *FLC* têm sido relatadas em vários estudos (BECKER; THEISSEN, 2003; REEVES et al., 2007). Nesse caso, quando submetemos os 12 EST-*contigs* e os 9 *singlets* às análises filogenéticas em comparação com outras espécies, em que o gene *FLC* já foi identificado e caracterizado, observou-se que o EST-*contig* 9 (621 nucleotídeos) encontrado no CAFEST apresentou similaridade com as demais espécies pertencente à subfamília *FLC*, com destaque para as espécies *Beta vulgaris* e *Poncirus trifoliata*, que apresentaram maior grau de similaridade com ele e sendo assim, agruparam-se (Figura 5).

Esses resultados são interessantes, pois estas duas espécies (*Beta vulgaris* e *Poncirus trifoliata*) agruparam-se com EST-*contig* 9 de *Coffea arábica*; em estudos recentes, foi observado que a expressão do gene *FLC* em ambas é semelhante a espécies como *Arabidopsis*, *Thellungiella* e *Brassica*, pois esse

gene quando expresso em plantas transgênicas de *Arabidopsis*, atrasou o florescimento, ou seja, está relacionado com o controle da transição da fase vegetativa para meristemas florais (REEVES et al. 2007, ZHANG et al. 2009).

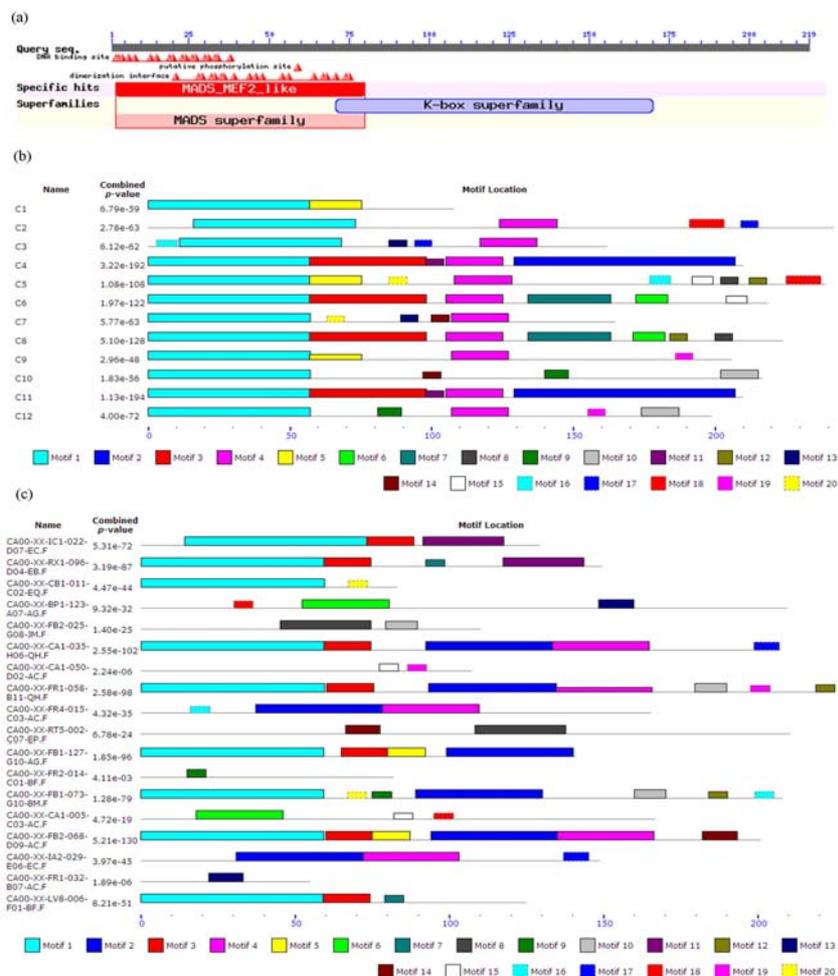


Figura 4 (a) Os dois principais domínios para o gene *FLC*, utilizando-se o programa (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) *Conserved Domains*. (b e c) - Motivos de agrupamento entre os prováveis *EST-contigs* e *singlets* respectivamente, para o gene *FLC*, encontrados no CAFEST, utilizando-se o programa MEME (*Multiple Expectation Minimization for Motif Elicitation*, (<http://meme.sdsc.edu/meme/meme.html>))

Isso demonstra que o EST-*contig* 9 é um forte candidato para o gene *FLC* em cafeeiro, mesmo não apresentando alta similaridade com *Arabidopsis*, *Thellungiella* e *Brassica*, que antes acreditava-se que genes *FLC* eram restritos a estas espécies.

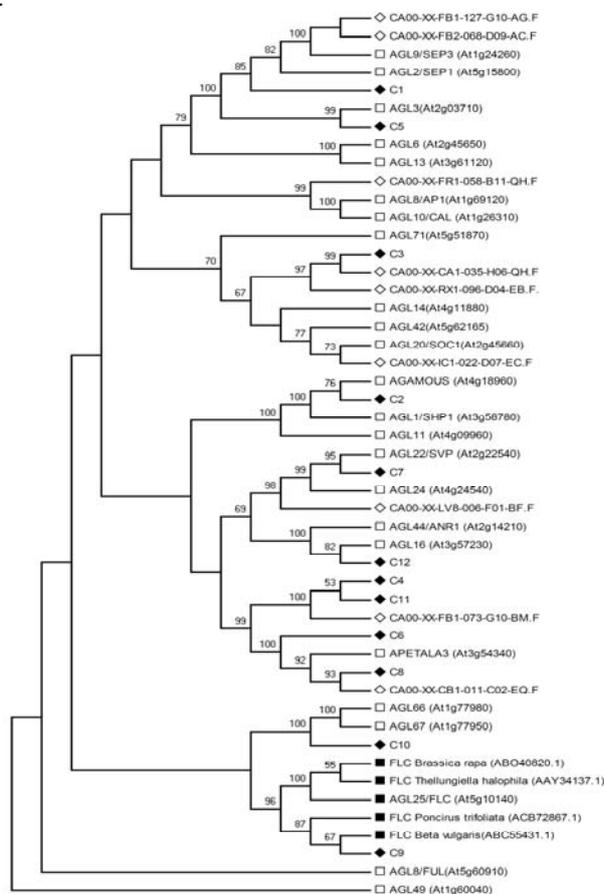


Figura 5 Árvore filogenética relacionando as 21 prováveis sequências de *FLC* de *C. arabica* encontradas no CAFEST, e as proteínas de genes MADS de espécies pertencente à subfamília *FLC* publicados e depositados no NCBI. (◆) EST-*contigs*, (◇) *singlets*, (□) *proteínas de genes MADS* de *Arabidopsis thaliana* (■) *FLC* de outras espécies já descritas. Foi utilizado o modelo de comparação *Neighbor-joining*, pelo método de distância *p* e supressão *pair-wise*. Valores de Bootstrap menores do que 50% foram omitidos

### 3.1.2 FRI

Com relação ao gene *FRI*, o resultado foi semelhante ao *FLC*. Foram obtidos 57 reads no banco de dados CAFEST e após clusterização (agrupamento) formaram-se 6 EST-*contigs* e 7 *singlets*. Após a análise das seqüências protéicas da parte funcional de cada uma delas e de seus domínios conservados, observou-se que dois EST-*contigs* (1 e 5) e um *singlet* apresentaram o domínio *FRIGIDA* (*motif* 1) muito conservado (Figura 6). Os EST-*contigs* e *singlets* incompletos (sem o domínio) foram excluídos das análises filogenéticas.

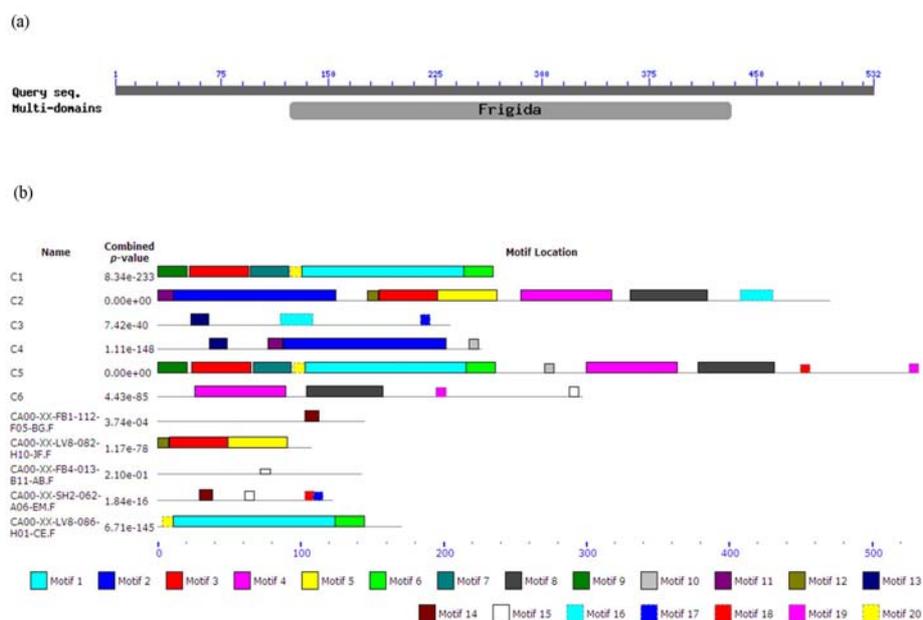


Figura 6 (a) O principal domínio para o gene *FRI*, utilizando-se o programa (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) *Conserved Domains*. (b) - Motivos de agrupamento entre os prováveis EST-*contigs* e *singlets*, para o gene *FRI*, encontrados no CAFEST, utilizando-se o programa MEME (*Multiple Expectation Minimization for Motif Elicitation*, <http://meme.sdsc.edu/meme/meme.html>)

As análises filogenéticas demonstram a comparação entre os EST-*contigs* 1 e 5 e o *singlet* com a superfamília *FRIGIDA* descrito por Risk et al., (2010), em que se observou que os EST-*contigs* 1 e 5 e o *singlet* agruparam-se no mesmo grupo e apresentaram similaridade com espécies como *Medicago truncatula*, *Glycine max*, *Vitis vinifera* e *Arabidopsis thaliana*, pertencentes à família IV; os membros dessa família ainda não foram caracterizados funcionalmente (Figura 7).

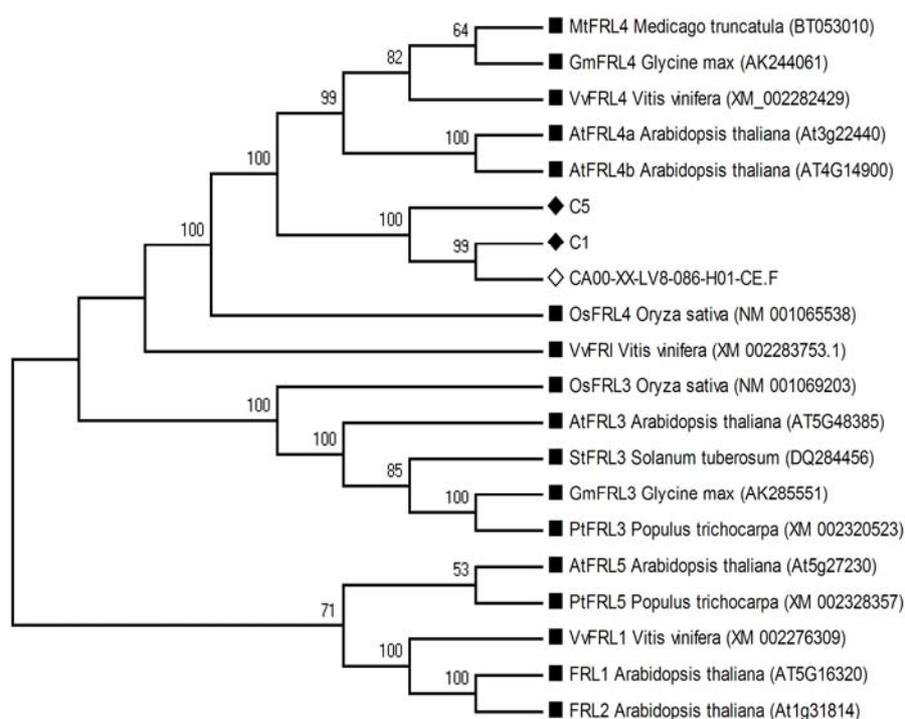


Figura 7 Árvore filogenética relacionando as 3 prováveis sequências de *FRI* de *C. arabica* encontradas no CAFEST, e as proteínas de espécies pertencente à subfamília *FRI* publicados e depositados no NCBI. (◆) EST-*contigs*, (◇) *singlets* e (■) espécies já descritas. Foi utilizado o modelo de comparação *Neighbor-joining*, pelo método de distância *p* e supressão *pair-wise*. Valores de Bootstrap menores do que 50% foram omitidos

### 3.1.3 VRN2

Após a clusterização (agrupamento) foi encontrados 10 *reads* no banco de dados CAFEST, que formaram 2 EST-*contigs* e 2 *singlets*. Com a análise das sequências proteicas da parte funcional de cada uma delas e seus domínios conservados, foi possível observar que o EST-*contig 1* e o *singlet* CA00-XX-FB1-011-F05-QH.F apresentaram o domínio conservado VEFS-*box* (*motif 1*) característico do gene (Figura 8).

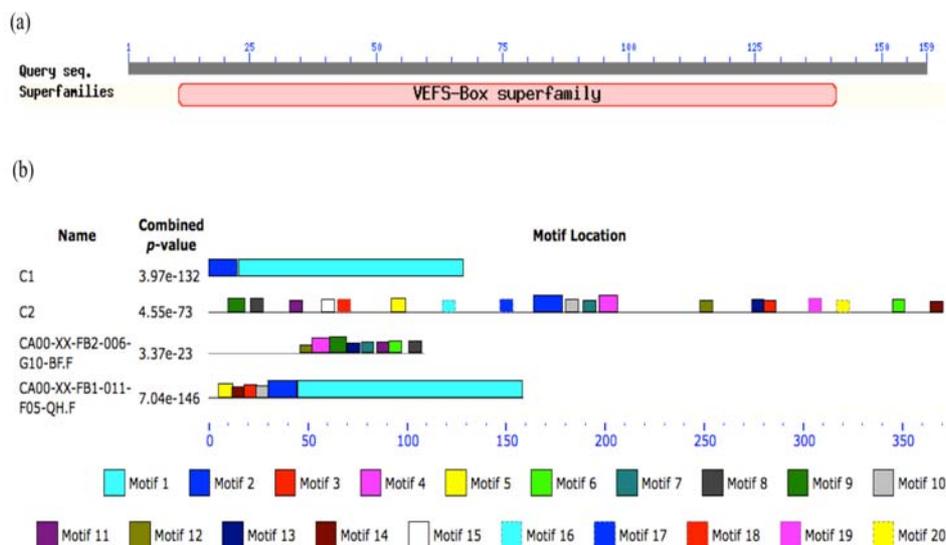


Figura 8 (a) O principal domínio para o gene *VRN2*, utilizando-se o programa (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) Conserved Domains. (b) - Motivos de agrupamento entre os prováveis EST-*contigs* e *singlets*, para o gene *VRN2*, encontrados no CAFEST, utilizando-se o programa MEME (*Multiple Expectation Minimization for Motif Elicitation*, <http://meme.sdsc.edu/meme/meme.html>)

Com as análises filogenéticas (Figura 9), foi possível a confirmação do EST-*contig 1* e o *singlet* CA00-XX-FB1-011-F05-QH.F como fortes candidatos para o gene *VRN2* de *Coffea arabica*, pois ambos apresentaram alta similaridade com o gene *VRN2* de *Arabidopsis thaliana* e *VRN2-like* da espécie *Allium sativum*. Pode ser observado também que o EST-*contig 2* e o *singlet* CA00-XX-FB2-006-G10-BF.F, que embora não apresentaram o domínios conservados para o gene em estudo, alinharam-se com espécie *Triticum aestivum*.

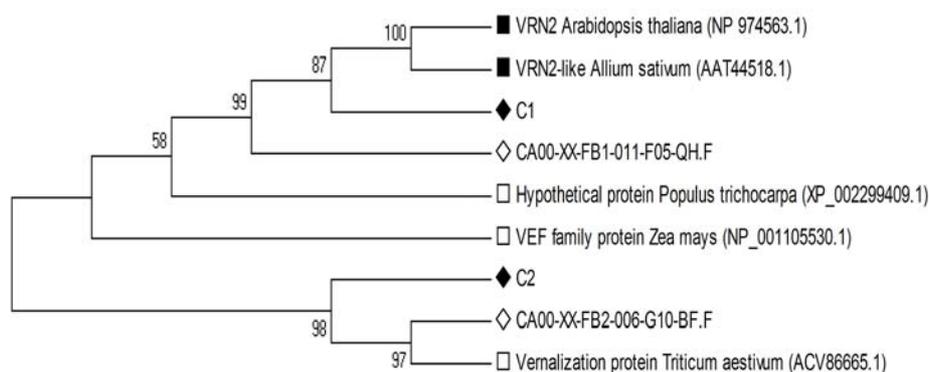


Figura 9 Árvore filogenética relacionando as 4 prováveis sequências de *VRN2* de *C. arabica* encontradas no CAFEST, e as proteínas de espécies pertencente à subfamília *VRN2* publicados e depositados no NCBI. (◆) EST-*contigs*, (◇) *singlets* e (■) espécies já descritas (□) espécies semelhantes as já descritas. Foi utilizado o modelo de comparação *Neighbor-joining*, pelo método de distância *p* e supressão *pair-wise*. Valores de Bootstrap menores do que 50% foram omitidos

### 3.1.4 VIN3

Para o gene *VIN3*, foram encontrados 23 *reads* no banco de dados CAFEST que formaram 4 EST-*contigs* e 4 *singlets*. Com a análise das sequências proteicas da parte funcional de cada uma delas e seus domínios

conservados, foi possível observar que o EST-*contig 1* e o *singlet* CA00-XX-EA1-019-D02-EC.F apresentaram o domínio conservado DUF1423 (*motif 4*) característico do gene (Figura 10).

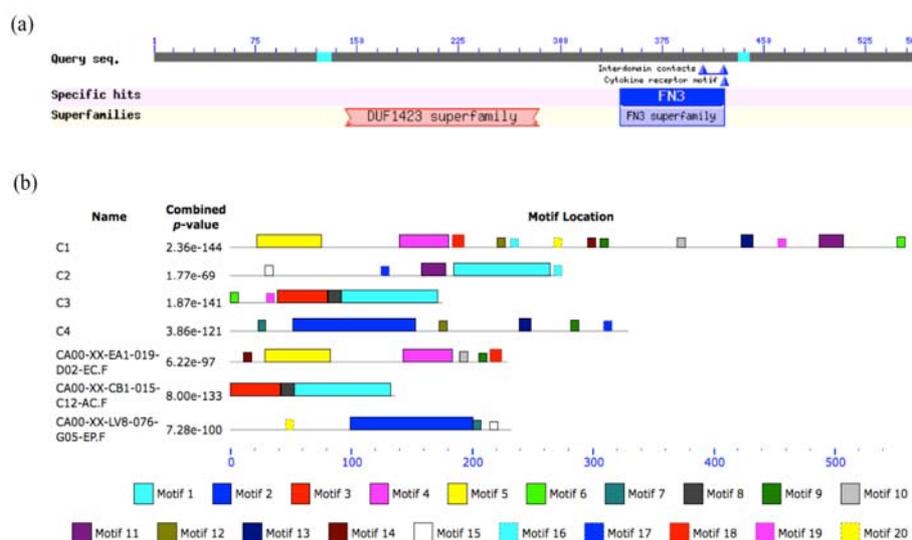


Figura 10 (a) Os dois principais domínios para o gene *VIN3*, utilizando-se o programa (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) Conserved Domains. (b) - Motivos de agrupamento entre os prováveis EST-*contigs* e *singlets*, para o gene *VIN3*, encontrados no CAFEST, utilizando-se o programa MEME (*Multiple Expectation Minimization for Motif Elicitation*, <http://meme.sdsc.edu/meme/meme.html>)

Após as análises filogenéticas (Figura 11), observou-se resultado semelhante ao do estudo dos domínios conservados, pois o EST-*contig 1* e o *singlet* CA00-XX-EA1-019-D02-EC.F, que apresentaram o domínio conservado DUF1423, também foram os que se alinharam com as espécies *Arabidopsis thaliana* e *Brassica rapa* para o gene em estudo (*VRN2*) de *Coffea arabica*, pois ambos apresentaram alta similaridade.

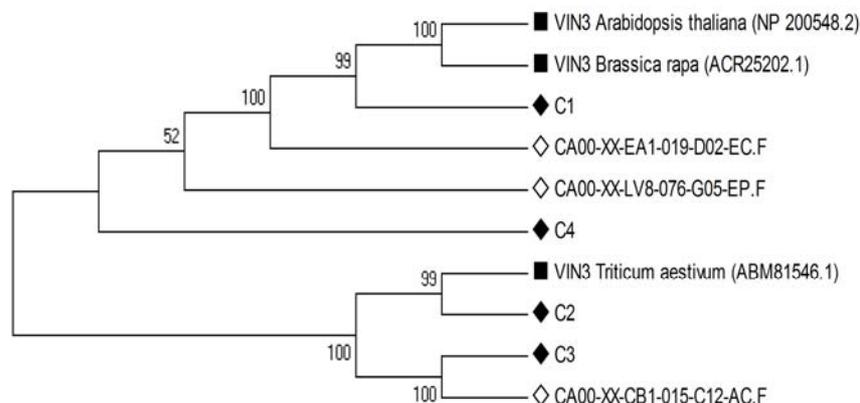


Figura 11 Árvore filogenética relacionando as 4 prováveis sequências de *VIN3* de *C. arabica* encontradas no CAFEST, e as proteínas de espécies pertencente à subfamília *VIN3* publicados e depositados no NCBI. (◆) EST-contigs, (◇) singlets e (■) espécies já descritas. Foi utilizado o modelo de comparação *Neighbor-joining*, pelo método de distância *p* e supressão *pair-wise*. Valores de Bootstrap menores do que 50% foram omitidos

### 3.2 Northern eletrônico

#### 3.2.1 *FLC*, *FRI*, *VRN2* e *VIN3*

Com análise do *northern* eletrônico, pode-se confirmar os resultados obtidos com a filogenia, pois observou-se que o EST-contig 9 apresentou o maior padrão de expressão em bibliotecas de caules infectados com *Xylella* spp. (RX1) e de tecidos sob estresse hídrico (SH2), ambas amostras oriundas de tecidos vegetativos submetidos a estresse (Figura 12a). Esse fato é muito relevante, inferindo-se que a expressão de *FLC* pode ser detectável quando a planta sofre estresse, incluindo biótico (como por ataque de um patógeno) ou abiótico (déficit hídrico).

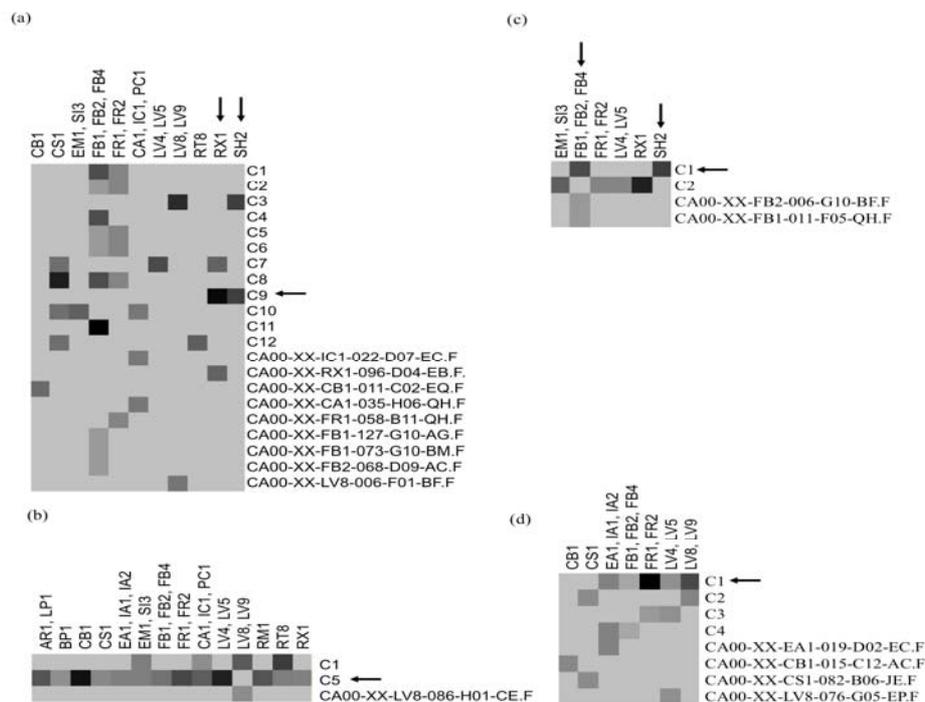


Figura 12 *Northern* Eletrônico representando os níveis de expressão dos EST-contigs nas diferentes bibliotecas por meio de uma escala cinza, onde quanto mais escura for à cor, maior a expressão. (a) – Gene *Flowering Locus C*. (b) – Gene *Frigida*. (c) Gene *Reduced Vernalization Response 2*. (d) – *Vernalization Insensitive 3*. Bibliotecas: AR1/LP1, folhas tratadas com ácido araquidônico; BP1, células em suspensão tratadas com acibenzolar-S-metil; CB1, células em suspensão tratadas com acibenzolar-S-metil and brassinoesteroids; CS1, células em suspensão tratadas com NaCl; EA1/IA1/IA2, calos embriogênicos; EM1/SI3, germinação de sementes (sementes inteiras e embriões zigóticos); FB1/FB2/FB4, botões florais em diferentes estágios do florescimento; FR1/FR2, botões florais e frutos de diferentes estágios; CA1/IC1/PC1, calos não embriogênicos com e sem 2,4 D; LV4/LV5, folhas jovens de ramos ortotrópicos; LV8/LV9, folhas maduras de ramos plagiotrópicos; PA1, calos embriogênicos primários; RM1, folhas infectadas com bicho-mineiro e ferrugem; RT8, células em suspensão com estresse ao alumínio; RX1, caules infectados com *Xylella* spp.; SH2, pool de tecidos de plantas sob estresse hídrico no campo;. As bibliotecas utilizadas foram descritas por Vieira et al., 2006. As setas indicam EST contigs selecionados para qPCR

Sendo o cafeeiro uma cultura que necessita de um período de déficit hídrico, seguido de uma elevação da umidade para que ocorra o florescimento, isso reforça o EST-*contig* 9 como um forte candidato para o gene *FLC*. Além disso, a biblioteca SH2 foi obtida por meio de um pool de diferentes tecidos (excluindo-se raiz); assim, espera-se que o *FLC* (*Contig* 9) seja expresso constitutivamente em todas as partes da planta, inclusive em raiz, uma vez que entre os 214,773 mil reads do banco de dados do café, apenas 1,6 % (3,440) dos *reads* está relacionado com a raiz (VIEIRA et al., 2006).

Esses resultados são relevantes, pois Zhang et al. (2009) ao analisarem o nível de expressão tecido-específico do gene *PtFLC* em *Poncirus trifoliata*, observaram a expressão do gene *FLC* em tecidos de folha, flor, raiz e fruto. De fato, esta espécie foi uma que se agrupou com EST-*contig* 9 de *Coffea arabica* durante as análises filogenéticas (Figura 5).

Com relação ao *FRI*, observou-se que o EST-*contig* 5, pertencente à família IV da superfamília *FRIGIDA*, foi o que obteve maior padrão de expressão, pois apresentou perfil de expressão em todas as bibliotecas em estudo, exceto em folhas maduras de ramos plagiotrópicos, muito embora a biblioteca de folhas jovens de ramos ortotrópicos foi uma das bibliotecas com maior perfil de expressão (Figura 12b). Esses resultados corroboram com a literatura, pois a família IV tem altos níveis de expressão em diversos tecidos, inclusive durante o desenvolvimento da semente (RISK et al., 2010).

Para os gene *VRN2*, o EST-*contig* 1 apresentou o perfil de expressão nas bibliotecas de botões florais em diferentes estágios de desenvolvimento e (FB1/FB2/FB4) e de tecidos sob estresse hídrico (SH2), o que faz desse um forte candidato para o estudo do gene *VRN2* em *Coffea arabica* (Figura 12c). Agora, com relação ao *VIN3* (Figura 12d), o EST-*contig* 1 apresentou padrões de expressões em bibliotecas de calos embriogênicos (EA1/IA1/IA2) botões florais em diferentes estágios de desenvolvimento (FB1/FB2/FB4), botões florais mais frutos em diferentes estágios de desenvolvimento (FR1/FR2), todos tecidos

relacionados à parte produtiva da planta, o que faz desse um EST-*contig* para o estudo do gene VIN3 em *Coffea arabica*.

#### 4 CONCLUSÃO

Com base em uma busca direcional no banco de dados *CAFEST*, foi possível identificar prováveis sequências para os genes *FLC*, *FRI*, *VRN2* e *VIN3*, que apresentaram com as análises de *motifs* porções determinantes para identificação desses genes em *Coffea arabica*.

Com a construção da árvore filogenética, foi possível demonstrar o correto agrupamento das sequências dos genes em estudo, com proteínas deles em outras espécies, pois todos apresentaram alta homologia, observada pelos valores de *bootstrap*.

Na análise de expressão pelo *Northern* eletrônico, foi possível observar uma expressão característica desses genes em tecidos relacionados com o desenvolvimento reprodutivo (florescimento).

## REFERÊNCIAS

ALTSCHUL, S. F. et al. Gapped BLAST and PSI BLAST: A new generation of protein database search programs. **Nucleic Acids Research**, Oxford, v.25, n.17, p.3389-3402, Sept. 1997.

BAILEY, T.L.; ELKAN, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In: INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEMS FOR MOLECULAR BIOLOGY, Menlo Park. **Proceeding**. Menlo Park: AAI, 1994. p. 28–36.

BECKER, A.; THEISSEN G. The major clades of MADS-box genes and their role in the development and evolution of flowering plants. **Molecular Phylogenetic Evolution**, San Diego, v. 29, n.3, p.464-489, Dec. 2003.

BLAZQUEZ, M.A. Flower development pathways. **Journal of Cell Science**, London, v.113, n.20, p.3547-3548, Oct. 2000.

CIANNAMEA, S. **Molecular characterization of the vernalization response in *lilium perenne***. 2007. 147p. Tese (Doctor in Plant Science) – Wageningen University, Wageningen.

CONAB: COMPANHIA NACIONAL DE ABASTECIMENTO. **Primeira estimativa**: janeiro 2011. Disponível em: < [www.conab.gov.br/](http://www.conab.gov.br/) >. Acesso em: 20 jan. 2011.

EISEN, M.B. et al. Cluster analysis and display of genome-wide expression patterns. **Proceedings of the national academy of sciences of the United States of America**, Washington, v.95, n.25, p.14863-14868, Nov. 1998.

PARENICOVA, L. et al. Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: New openings to the MADS world. **The Plant Cell**, Rockville, v. 15, n.7, p.1538-1551, July 2003.

REEVES, P.A. et al. Evolutionary conservation of the FLC-mediated vernalization response: evidence from the sugarbeet (*Beta vulgaris*). **Genetics**, Baltimore, v.176, n.1, p.295-307, May 2007.

SHINDO, C. et al. Role of FRIGIDA and FLOWERING LOCUS C in determining variation in flowering time of *Arabidopsis*. **Plant Physiology**, Rockville, v.138, n.2, p.1163- 1173, Jun. 2005.

RISK, J.M. et al. FRIGIDA and related proteins have a conserved central domain and family specific N- and C- terminal regions that are functionally important. **Plant Mol Biology**, New Zealand, v.73, p.493–505, July. 2010.

SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, Lawrence, v. 4, n.4, p.406–425, July. 1987.

SITNIKOVA, T.; RZHETSKY, A.; NEI, M. Interior-branch and bootstrap tests of phylogenetics trees. **Molecular Biology and Evolution**, Lawrence, v.12, p.319-33, Apr. 1995.

TAMURA, K. et al. MEGA4: molecular evolutionary genetics analysis: software: version 4.0. **Molecular Biology and Evolution**, Oxford, v. 24, n. 8, p. 1596-1599, Aug. 2007.

THOMPSON, J.D. et al. Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Research**, Oxford, v. 22, p.4673–4680, Nov. 1994.

TNS RESEARCH INTERNATIONAL. **Primeira estimativa** 2008: janeiro de 2008 <<http://www.abic.com.br/estatisticas.html#intro>>. Acesso em 20 jan 2010.

VIEIRA, L.G.E. et al. Brazilian coffee genome project: an EST-based genomic resource. **Brazilian Journal Plant Physiology**, Londrina, v.18, n.1, p.95-108, Jan./Mar. 2006.

ZHANG, J.Z. et al. *PtFLC* homolog from trifoliate orange (*Poncirus trifoliata* ) is regulated by alternative splicing and experiences seasonal fluctuation in expression level. **Plant Journal**, Oxford, v. 299, n. 4, p. 847-859, Mar. 2009.

### **CAPÍTULO 3**

**Sequenciamento e análise quantitativa dos genes *FLC*, *FRI*, *VRN2* e *VIN3*  
em café arábica**

## RESUMO

Neste capítulo relatam-se os resultados do sequenciamento e da análise de expressão dos genes *FLC*, *FRI*, *VRN2* and *VIN3* de café (*Coffea arabica*). O material vegetal foi coletado em um campo experimental da Universidade Federal de Lavras (UFLA). Cinco tecidos vegetais diferentes *Coffea arabica* cv. Rubi foram utilizados para extração de RNA. Amostras de cDNA de folhas foram amplificadas utilizando-se *primers* específicos para os genes *FLC*, *FRI*, *VRN2* and *VIN3* de café foram desenhados para a síntese de cDNA, realizada usando o kit *High-Capacity cDNA Reverse Transcription* (*Applied Biosystems*). Os fragmentos obtidos foram sequenciados e as sequências obtidas foram, então, comparadas com aquelas encontradas no CAFEST e também com o GenBank (NCBI). Para a análise de expressão por qPCR, *primers* para cada gene foram desenhados no programa *Primer Express v2.0* (*Applied Biosystems*) e as análises de expressão foram realizadas no aparelho ABI PRISM 7500 Real-Time PCR (*Applied Biosystems*), utilizando o sistema de detecção SYBR Green e o ensaio de quantificação relativa. Os resultados obtidos neste estudo possibilitaram a caracterização de quatro genes de grande importância no processo de florescimento do café, e estudos futuros compreendendo mutantes para esses genes em species-modelo e também em plantas de café certamente fornecerão uma melhor compreensão das funções desempenhadas por cada um desses genes durante a indução do florescimento em café.

Palavras-chave: *Coffea arabica*. Expressão gênica. MADS-box. qPCR. Regulação do florescimento.

## ABSTRACT

This chapter reports the results of sequencing and expression analysis of the coffee (*Coffea arabica*) *FLC*, *FRI*, *VRN2* and *VIN3* putative homologs. Plant material was collected in an experimental field located at Federal University of Lavras (UFLA). Five different plant tissues from *Coffea arabica* cv. Rubi were used for RNA extraction. Samples of leaf cDNA were amplified with specific primers for full-length coffee *FLC*, *FRI*, *VRN2* and *VIN3* putative homolog transcripts. The amplification fragments were sequenced, and the transcript sequences obtained were then compared to those found in the CAFEST database and also in the GenBank (NCBI). For the qPCR analysis, primers for each gene were designed in the Primer Express v2.0 (*Applied Biosystems*) program and expression analysis were performed in an ABI PRISM 7500 Real-Time PCR (*Applied Biosystems*), using the SYBR Green detection system and the relative quantification assay. The results obtained from this study enabled the characterization of four genes of paramount importance of the *Coffea arabica* flowering process, and further studies comprising mutants for these genes in model species and also in coffee trees will certainly provide better comprehension of the functions displayed by each one of these during coffee flowering induction.

Keywords: *Coffea arabica*. Gene expression. MADS-box. qPCR. Regulation of flowering.

## 1 INTRODUÇÃO

A transição da fase vegetativa para reprodutiva é particularmente importante na agricultura em termos de produtividade. No cafeeiro, o florescimento desuniforme causa danos de produção e muitos prejuízos.

A compreensão e identificação do mecanismo de florescimento, assim como os genes envolvidos nos processos de regulação floral, são de fundamental importância para elucidar como esses mecanismos ocorrem em café. Vários genes já foram identificados como peças-chave na regulação molecular do florescimento em *Arabidopsis*.

Para maximizar o sucesso reprodutivo, plantas têm desenvolvido mecanismos para permitir que o florescimento ocorra sob condições favoráveis (RATCLIFFE et al., 2001). Análises de mutantes de atraso de florescimento e ecotipos de *Arabidopsis* têm revelado que esses mecanismos dependem da interação de várias vias genéticas, como fotoperíodo e vernalização (KOORNNEEF et al., 1998; FINNEGAN et al., 2004).

Estudos referentes à influência da temperatura na floração do cafeeiro indicam que aparentemente temperaturas baixas estimulam a iniciação floral sob condições controladas em interação com o fotoperíodo (GUIMARÃES; MENDES, 1998).

A aceleração do florescimento pela exposição a um longo período de frio é um processo denominado vernalização. A exigência da vernalização é uma estratégia reprodutiva adaptada por muitas espécies de plantas (SIMPSON; DEAN, 2002). Em *Arabidopsis*, a vernalização é controlada por dois loci principais: *FLOWERING LOCUS C (FLC)* e *FRIGIDA (FRI)* (SHINDO et al., 2005). Sendo o *FLC* o maior repressor floral da via de integração floral em *Arabidopsis* e quando ativo, conduz a ativação de genes de identidade de meristemas.

De maneira positiva e reciprocamente, a expressão de *FLC* é regulada por *FRIGIDA*, como também pela vernalização por meio dos genes, como REDUCED VERNALIZATION RESPONSE 2 (*VRN2*) e VERNALIZATION INSENSITIVE 3 (*VIN3*) (CIANNAMEA, 2007).

Dessa forma, objetivou-se com esse trabalho analisar a expressão gênica dos genes *FLC*, *FRI*, *VRN2* e *VIN3* em café arábica.

## **2 MATERIAL E MÉTODOS**

### **2.1 Coleta de material vegetal**

O material vegetal foi coletado no campo experimental da Universidade Federal de Lavras (UFLA) em Lavras-MG (21°14'42"S 45°00'00"W) e o experimento foi desenvolvido no Laboratório Central de Biologia Molecular (LCBM-UFLA). Para a extração de RNA de café, foram utilizados cinco tecidos vegetais de *Coffea arabica*, cv. Rubi: raízes (secundárias), folhas (plantas submetidas a 90 dias de estresse hídrico antes do florescimento), folhas (durante o florescimento) flores (estádios G4-G5 descritos por Moraes et al. (2008) e frutos (em diferentes estádios de desenvolvimento). As amostras coletadas foram conservadas imediatamente ao nitrogênio líquido e posteriormente armazenado a -80 °C.

### **2.2 Extração de RNA e síntese de cDNA**

Para extração do RNA, utilizou-se o Kit Concert™ Plant RNA Reagent (Invitrogen), seguindo as recomendações do manual, com pequenas alterações. Em seguida, as amostras foram tratadas com o inibidor Turbo DNA-free Kit (Ambion), para eliminação de contaminação de DNA residual. A integridade do RNA foi verificada em gel de agarose 0,8%. As amostras foram quantificadas em espectrofotômetro (Nanodrop® Espectrophotometer ND-1000) a  $A_{260}$ . As amostras que apresentaram alto grau de integridade e pureza foram usadas para a síntese de cDNA, com o Kit High-Capacity cDNA Reverse Transcription (Applied Biosystems). Após a síntese de cDNA, as amostras foram armazenadas no freezer a -20° C até o uso.

### 2.3 Sequenciamento

A partir das prováveis sequências dos genes *FLC*, *FRI*, *VRN2* e *VIN3* de café que foram obtidas do estudo *in silico* (capítulo 2), foram desenhados os *primers* para PCR (Tabela 1) utilizando o programa Primer 3 (<http://frodo.wi.mit.edu/primer3/>).

Tabela 1 *Primers* utilizados na PCR e sequenciamento (*FLC*-EST-*contig* 9, *FRI*-EST-*contig* 5, *VRN2*-EST-*contig* 1 e *VIN3*-EST-*contig* 1)

Gene	Sequência do <i>primer</i>	<i>Amplicon</i> ( <i>pb</i> )
FLC F	5'-CACCGAAGGAATTAAGGA	621
FLC R	5'-TTAAAGTAGTTCAAGCATA	
FRI F	5'-CACCAATCTAAAGAGATGTTGG	1599
FRI R	5'-ATCTTTCCTACTGATTAGTGCA	
VRN2 F	5'-CACCTCCAACCTGTGCTG	390
VRN2 R	5'-GCAACAATCGCTTTCACC 3'	
VIN3 F	5'-GCTCCTCAAAAGAAGCTGGT	1680
VIN3 R	5'-CCCAGTCTGCTAGGGAGTATT	

As amostras de cDNA de folhas foram amplificadas utilizando uma enzima de alta fidelidade, que possui atividade de exonuclease 3'-5', a *Pfx* DNA polymerase (Invitrogen), e *primer* específicos (“forward” e “reverse”) para todos genes em estudo.

Os produtos das ampliações foram submetidos à eletroforese em gel de agarose 1,2% (m/v) corado com 2 µL de Brometo de Etídio (0,5 µg/mL), sob corrente elétrica de 110 V em Tampão TAE (1 mM EDTA pH 8,0; 40 mM TRIS

pH 8,0; 20 mM ácido acético), por 40 min, que foi visualizado sob luz ultravioleta e a imagem captada pelo fotodocumentador EDAS 290 (Kodak®). Os fragmentos obtidos foram eluídos do gel de agarose por meio do Kit MinElute (Qiagen®).

Os fragmentos foram submetidos ao sequenciador automático MegaBace 1000 (Amersham Biosciences), pelo método enzimático baseado na síntese de DNA *in vitro*, na presença de nucleosídeos trifosfatados terminadores de cadeia (SANGER et al., 1977), com o kit “Big Dye Terminator” (Applied Biosystems).

#### **2.4 Análises das sequências em banco de dados**

Para pesquisa de similaridade, as sequências dos fragmentos *FLOWERING LOCUS C (FLC)*, *FRIGIDA (FRI)*, *REDUCED VERNALIZATION RESPONSE 2 (VRN2)* e *VERNALIZATION INSENSITIVE 3 (VIN3)* obtidas mediante sequenciamento foram comparadas com as previamente obtida *in silico* e com do GenBank via BlastX (ALTSCHUL et al., 1997).

#### **2.5 qPCR**

A partir das prováveis sequências dos genes *FLC*, *FRI*, *VRN2* e *VIN3* de café que foram obtidas do estudo *in silico* e confirmadas pelo sequenciamento, foram desenhados os *primers* para qPCR (Tabela 2) utilizando o programa Primer Express v2.0 da Applied Biosystems.

Para a análise da expressão gênica quantitativa por qPCR, foi utilizado o modelo ABI PRISM 7500 Real-Time PCR (Applied Biosystems), usando SYBR Green e o cDNA obtido a partir de RNA extraído dos cinco tecidos em estudo.

As condições térmicas da reação foram 2 minutos a 50°C, 10 minutos a 95°C, seguidos por 40 ciclos de 15 segundos a 95°C e 1 minuto a 60°C, e finalizando com 15 segundos a 95°C. Os dados foram coletados e armazenados no programa 7500 Fast Software (Versão 2.1). Para cada reação, foram utilizados 80ng de cDNA, 1,5 µM de cada primer e 5,0 µL de Master Mix SYBR green UDG com ROX (Invitrogen) para um volume final 10,0 µL/amostra. As amostras foram processadas em réplicas e os resultados foram normalizados usando  $C_T$ s (Ciclo *Threshold*), obtidos pela expressão dos genes endógenos,  $\beta$ -actin (*ACTINA*) e *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*). O  $C_T$  foi determinado pelo número de ciclos no qual a fluorescência gerada dentro de uma reação cruza a linha de base (*threshold*). O método que foi usado é o do  $C_T$  comparativo.

Tabela 2 *Primers* utilizados na qPCR (*FLC-EST-contig 9*, *FRI-EST-contig 5*, *VRN2-EST-contig 1* e *VIN3-EST-contig 1*)

Gene	Sequencia do primer	Eficiência %
FLC F	5'-GCTGAGGGAAGAAAACAAGC	97,80
FLC R	5'-CGTGTGTAAGGTCCCCAAAT	
FRI F	5'-CATTTATGGACACGGAAGCA	100
FRI R	5'-GGACTGCCATAGCCTCCAT	
VRN2 F	5'-GGGCTTGTGAGGCGTTTTTC	92,5
VRN2 R	5'-ACAGTCTCCAACACCAAAGCAA	
VIN3 F	5'-GCCGTAAGTATGATGACAACAAAGA	100
VIN3 R	5'-TAGGGCACAATCAAGGTGACAT	
ACT F	5'-AATTGTCCGTGACATCAAGGAA	91,01
ACT R	5'-TGAGCTGCTTTGGCTGTTC	
GAPDH F	5'-TTGAAGGGCGGTGCAAA	99,02
GAPDH R	5'-AACATGGGTGCATCCTTGCT	

A normalização foi realizada utilizando-se a equação  $\Delta C_T = C_T$  (gene alvo) -  $C_T$  (controle endógeno). A calibração foi determinada pela fórmula  $\Delta\Delta C_T = \Delta C_T$  (amostra) -  $\Delta C_T$  (calibrador). O calibrador é uma amostra usada como base para resultados de expressão comparativa. A quantificação relativa foi obtida pela fórmula  $2^{-\Delta\Delta C_T}$ .

### 3 RESULTADOS E DISCUSSÃO

#### 3.1 Sequenciamento

Após o sequenciamento, foi possível confirmar as prováveis sequências dos genes *FLC*, *FRI*, *VRN2* e *VIN3* de café, que foram identificadas do estudo *in silico* (capítulo 2). Para todos os genes em estudo, foram obtidas sequências menores após o sequenciamento do que as identificadas nas análises *in silico*, resultado esse já esperado, pois foi realizado sequenciamento direto, ou seja, a partir do produto de PCR, o que gerou uma perda no tamanho dos amplicons, devido à baixa qualidade das partes iniciais e finais das sequências; porém, isso não interferiu no resultado, uma vez que todas as sequências apresentaram alta identidade.

Para o gene *FLC* (*contig* 9) de 621 pares de bases, foi obtida uma sequência de 573 pares de bases após o sequenciamento, em que pelo alinhamento das duas sequências pelo programa ClustalW (Thompson et al., 1994), observou-se alta identidade (Figura 13). Uma vez que a sequência foi confirmada, ela foi comparada com o banco de dados (NCBI) utilizando o programa BLASTx (ALTSCHUL et al., 1997). A sequência do *FLC* apresentou homologia com genes MADS de várias espécies e principalmente com os genes *BvFLC* (*B. vulgaris*), cujos alinhamentos demonstraram 56% de identidade, *PtFLC* (*P. Trifoliata*) com 51% de identidade, *ThFLC* (*T. halophila*) com 49% de identidade e *AGL24/FLC* 49% de identidade.

Com relação ao *FRI* (*contig* 5), foi selecionada apenas parte da sequência obtida *in silico* para a confirmação pelo sequenciamento, pois a sequência identificada tem 1599 pares de bases e o sequenciador (MegaBace) produz de 600 a 750 pares de bases com qualidade por amostra. Analisando os 720 pares de bases obtidos pelo sequenciamento, foi possível fazer a

confirmação da sequência obtida *in silico* (Figura 14). Quando analisamos essa sequência em comparação ao banco de dados (NCBI), observou-se que o contig 5 (*FRI*) apresentou uma identidade de 69% com o gene *VvFRL4* de *Vitis vinifera*, 67% com *GmFRL4* de *Glycine max*, 66% com *MtFRL4* de *Medicago truncatula*, 66% e 65% com *AtFRL4a* e *AtFRL4b* respectivamente (*Arabidopsis thaliana*).

```

C9          ATGGGGCGGAGGAAGTGGAGATTAAGAAAA-TCGAGGACAAGAACAGCAGGCAAGTCAC 59
Seq         ---GGCGGAGGAAGTGGAGATTAAGAAAAATCGAGGACAAGAACAGCAGGCAAGTCAC 56
          *****

C9          GTTTCCAAAGCGGAGAAGCGGACTGATGAAGAAAGCCAAGGAACTTCCGTTCTCTGCGA 119
Seq         GTTTCCAA-CGGAGAA-CGGACTGATGAAGAAAGCCAAGGAACTTCCGTTCTCTGCGA 114
          *****

C9          CGTGGATGTGCTGTCTCATCTTCTCTGGTCGCGGCAAGCTCTACGACTTCTGCAGCAC 179
Seq         CGTGGATGTGCTGTCTCATCTTCTCTGGTCGCGGCAAGCTCTACGACTTCTGCAGCAC 174
          *****

C9          CAACAGTTGGCCAAGATCCTACAACGATATCGCAACTACGCAGAAGCAGAAGACGGGTC 239
Seq         CAACAGTTGGCCAAGATCCTACAACGATATCGCAACTACGCAGAAGCAGAAGACGGGCC 234
          *****

C9          TGCAAGAATTAGCGGCTAGAGAAACGTAACCTGAAGGCAGAAATGTCGTGACAATCAG 299
Seq         TGCAAGAAT-AGCGGCTAGAGAAACGTAACCTGAAGGCAGAAATGTCGTGACAATCAG 293
          *****

C9          AAAGCTGCTGGAAAAAGT-TGAAAGGGATCTAGAGGAGCCAGATGTTGACCACCTTAACC 358
Seq         AAAGCTGCTGGAAAAAGTCTGAAAGGGATCTAGAGGAGCCAGATGTTGACCACCTTAACC 353
          *****

C9          TGAGTGAAGTACTGCAATTGGAAGAACAACCTGAAGATGCACTCATTCAAACAAGATCTA 418
Seq         TGAGTGAAGTACTGCAATTGGAAGAACAACCTGAAGATGCACTCATTCAAACAAGATCTA 413
          *****

C9          GGAAGACACGATTACTGATGGAATCAATAACCAGTCTAAGTGAAGTGGAAAAGATGCTGA 478
Seq         GGAAGACACGATTACTGATGGAATCAATAACCAGTCTAAGTGAAGTGGAAAAGATGCTGA 473
          *****

C9          GGAAGAAAACAAGCTTCTGCAAAAATAAGGTAGCTGCAGGTACATCCAATGAGA-AGAGG 537
Seq         GGAAGAAAACAAGCTTCTGCAAAAATAAGGTAGCTGCAGGTACATCCAATGAGATAGAGG 533
          *****

C9          AATGACTTGATCCTTGAATTTGGGGACCTTACACACGTTGGAATGATTTCTGGGCAGCGA 597
Seq         AATGACTTGATCCTTGAATTTGGGGACCTTACACACGTTGGAATGATTTCTGGGCAGCGA 597
          AATGACTTGATCCTTGAATTTGGGGACCTTAAAC-----CGGGC----- 573
          ***** * ***** * ** ** * ****

C9          CAGGCTATGCTTGAACACTTTAA 621
Seq         -----

```

Figura 13 Alinhamento das sequências *in silico* e sequenciada do *Contig 9* (candidato para o gene *FLC*) pelo programa ClustalW (THOMPSON et al., 1994)

```

Part_C5      GCTG-TCTGTGCGATTACTTCTGCTCTGGAGAAGCATAGGAAGTTATCTTCTGATGAACC 59
Seq          GCTGGTCTGTGCGATTACTTCTGCTCTGGAGAAGCATAGGAAGTTATCTTCTGATGAACC 60
          ****

Part_C5      TGCTTTTGATAAATAGCAAGGACCATGGTGGGGCACCAGATAAGGAGGATAAACCTCCTGA 119
Seq          TGCTTTTGATAAATAGCAAGGACCATGGTGGGGCACCAGATAAGGAGGATAAACCTCCTGA 120
          ****

Part_C5      TGCTATGGCTGCTGAGAGCAACATAAGACAGTTGACGGCACCTTCTGAAATTGAGAATGC 179
Seq          TGCTATGGCTGCTGAGAGCAACATAAGACAGTTGACGGCACCTTCTGAAATTGAGAATGC 180
          ****

Part_C5      AGTGGTGAAGTTGTATCCAGAGCTTGTCAAGTTGTGTGTCAGGAGATGGATTGAGAAGGCCT 239
Seq          AGNGGTGAAGTTGTATCCAGAGCTTGTCAAGTTGTGTGTCAGGAGATGAATTCAGAAGGCCT 240
          **

Part_C5      GCACAAATTT-ATCTCAGACAACCGTAAAAACCTTGCTGTTATGAGGGAGGAAATCCAA 298
Seq          GCACAAATTTTATCTCAGACAACCGTAAAAACCTTGCTGTTATGAGGGAGGAAATCCAA 300
          ****

Part_C5      ATGCACTAAAGGCTGCAAAATGATCCGGCCTCGCTGGTTCGGACTCACTCAATGGTTTCT 358
Seq          ATGCACTAAAGGCTGCAAAATGATCCGGCCTCGCTGGTTCGGACTCACTCAATGGTTTCT 360
          ****

Part_C5      ACAGCATGGACATGCCTAGCTCAGATGCCAAAAAGGACTCAAACCTCCTAGGCCTTCGTC 418
Seq          ACAGCATGGAAATGCCTAGCTCAGATGCCAAAAAGGACTCAAACCTCCTAGGCCTTCGTC 420
          ****

Part_C5      -GAAGCTGTATCATGCTGATGGAATGCCTTAGCACTTCATTCACAAA--TTTGATCAGG 475
Seq          CGAAGCTGTATCATGCTGATGGAATGCCTTAGCACTTCATTCACAAAATTTGATCAGG 480
          ****

Part_C5      AACTCATCTCAGGTATGATCTCATCTGATGTCAAGGAA--CGGG--CAAAAGTTATT-GC 530
Seq          AACTCATCTCAGGTATGATCTCATCTGATGCCAAGGAAACCGGGCCAAAAGTTATTGCG 540
          ****

Part_C5      TGAAGAGTGAAGCCGAAGTTAGATGAAC TTGA-TGTTGATG-CCAGCAG-TGGGAA--- 584
Seq          TGAAGAGTGAAGCCGAAGTTAGATGATCTTGAATGTTGATGCCAGCAGGTTGGGAAACT 600
          ****

Part_C5      CTCCTTAGAGGCAC---ATGCATTC--TTACAA--CTTCTA-GCTACTTTT---GGTA-T 632
Seq          CCCTTTGGAGGGGCTCCATGCTTTCCCTTTACGAAGTTCTAAGCTACCTTTTGGGTA 660
          * * * * *

Part_C5      CAATTCT---GATTTT--GA--CCAGGAGAGCTTATCCA-----AGCT-AATACCTATGG 679
Seq          CACTTCTCTGGATTTTGGAAACCAGGAGAGCTTTTTCACAAGAGGCTTANTACCTATGG 720
          **

Part_C5      --TTTCTCGTCGTCGCCAG-ACAG- 700
Seq          GTTTCCTCTCCGTCGCCAGGACAGT 745
          **

```

Figura 14 Alinhamento das seqüências *in silico* e sequenciada do *Contig 5* (candidato para o gene FRI) pelo programa ClustalW (THOMPSON et al., 1994)

Como pode ser observado na Figura 15, o alinhamento obtido da seqüência *in silico* com a do sequenciamento do *contig 1* (*VRN2*) possibilitou a

confirmação do *contig* e, após a comparação com sequências depositadas no banco de dados do NCBI, foi encontrado um valor de identidade da sequência do *contig* 1 de 75% com *VRN2* de *Arabidopsis thaliana*.

```

C1          ATGGCATTGGAGCAAGTCTTATCTGATCGGGACAGTGGAGATGAAGTTGATGATGATGTA 60
Seq         ATGGCATTGGAGCAAGTCTTATCTGATCGGGACAGTGGAGATGAAGTTGATGATGATGTA 60
*****

C1          GCAGATTTGGAAGATCGAAGGATGCTAGATGACTTTGTGGATGTTACGAAAGATGAAAAG 120
Seq         GCAGATTTGGAAGATCGAAGGATGCTAGATGACTTTGTGGATGTTACGAAAGATGAAAAG 118
*****

C1          CAAATGATGCATCTCTGGAACCTTTTGTTCGAAAACAAAGGGTCTGGCAGATGGTCAT 180
Seq         CAAATGATGCATCTCTGGAACCTTTTGTTCGAAAACAAAGGGTCTGGCAGATGGTCAT 178
*****

C1          GTTCCTGGGCTTGTGAGGCGTTTTCAAGGTTGCATGCACAGAATCTTGTCCAGGCACCT 240
Seq         GTTCCTGGGCTTGTGAGGCGTTTTCAAGGTTGCATGC-CAGNATCTTGTCCAGGCACCT 237
*****

C1          GCATTGCTTTGGTGTGGAGACTGTTTCATGATAAAGTTATGGAACCATGGTCTGCTTGAT 300
Seq         GCATTGCTTTGGTGTGGAGACTGTTTCATGATAAAGTTATGGAACCATGGTCTGCTTGAT 297
*****

C1          GCGCGTACTATGAACAATTGTAATATAATTCTTGAACAATGCCAAAACAAGACAATGAC 360
Seq         GCGCGTACCATGAACAATTGTAATATAATTCTTGAACAATGCCAAAACAAGACAATGAC 357
*****

C1          CCTATGACAAGCAGTAAAGGC-GGTGATTAG 390
Seq         CCTATGACAAGCAGTAAAGGCCGGTATTAG 388
*****

```

Figura 15 Alinhamento das sequências *in silico* e sequenciada do *Contig* 1 (candidato para o gene *VRN2*) pelo programa ClustalW (THOMPSON et al., 1994)

Com relação ao *VIN3* (*contig* 1), foi usada parte da sequência obtida *in silico* para confirmação do gene, assim como o gene *FRI* (*contig* 5) descrito anteriormente. Após a confirmação da sequência (Tabela 16), essa foi comparada com sequências do banco de dados (NCBI) e observou-se identidade de 44% com *VIN3* de *Arabidopsis thaliana* e 46% com *VIN3* de *Brassica rapa*.

Todos os resultados obtidos com o sequenciamento e análises das sequências são relevantes, pois corroboram com as análises filogenéticas realizadas neste trabalho (capítulo 2), em que todos os genes em estudo têm

apresentado alta similaridade e identidade com os genes (homólogos) descritos em outras espécies.

```

Parte_C1      CATCATCTGGCCATCTGGTGAGGGCTGGCTCTCAAGATTCACAGCATCAGTATTTGGTGC 60
Seq           CATCATCTGGCCATCTGGTGAGGGCTGGCTCTCAAGATTCACAGCATCAGTATTTGGTGC 60
*****

Parte_C1      AGAACAAATTTTCCCATGCTATGTTCCCTCATCCAACAAGGATATCGTCTCACCAGTAAG 120
Seq           AGAACAAATTTTCCCATGCTATGTTCCCTCATCCAACAAGGATATCGTCTCACCAGTAAG 120
*****

Parte_C1      TCCTCTTCAATAGTATCTGTATAGCCATTATACACCTTTGTAGAGACCATTTTTTCAAG 180
Seq           TCCTCTTCAATAGTATCTGTATAGCCATTATACACCTTTGTAGAGACCATTTTTTCAAG 180
*****

Parte_C1      GGTATTGTGATGGTCATGATAGTTATCCCTCTAGTCTCATCCTCATTGCTGCACGGAAT 240
Seq           GGTATTGTGATGGTCATGATAGTTATCCCTCTAGTCTCATCCTCATTGCTGCACGGAAT 240
*****

Parte_C1      GACATTATTAGTTTCATCTTCCACCGATGAGGGATTAGACAGGCTACTACAGTTGGTCCG 300
Seq           GACATTATTAGTTTCATCTTCCACCGATGAGGGATTAGACAGGCTACTACAGTTGGTCCG 300
*****

Parte_C1      TGGACTTTCACTCCTTCCAACCTCCATATCTTTAGA-GTTGGGATTAGTAGCTTCAGTTT 359
Seq           TGGACTTTCACTCCTTCCAACCTCCATATCTTTAGAAGTTGGGATTAGTAGCTTCAGTTT 360
*****

Parte_C1      CAGTAGCGGCTGTTTGGAACTAACTTCACACATCCCAGCTCCCTATTACTATCAAGGG 419
Seq           CAGTAGCGGCTGTTTGGAACTAACTTCACACATCCCAGCTCCCTATTACTATCAAGGG 420
*****

Parte_C1      ATACAACCTTGAGATGATAATCTGTATCTGGAGATAGACCAGAGAGATAAAATTCGAGT 479
Seq           ATACAACCTTGAGATGATAATCTGTATCTGGAGATAGACCAGAGAGATAAAATTCGAGT 480
* *****

Parte_C1      TTGGAGCAAAACAATGTACAA-GTAGGTCTGTCCGATAGTCCAATCATCAGATTTGCCA 538
Seq           TTGGAGCAAAACAATGTACAAAGTAGGTCTGTCCGATAGTCCAATCATCAGATTTGCCA 540
*****

Parte_C1      TGCCATAAGGTATAACCAACTAAATTTCC-AAGAGATGAATCTTCAAATTC-AGAGTCA 596
Seq           TGCCATAAGGTATAACCAACTAAATTTCCCAAGAGATGAATCTTCAAATTCAGAGTCA 600
*****

Parte_C1      C--AGTGACAGATGAGGTGCAGACATCTT--CCAACG-TCACC-AATTTAGCACTCATCA 650
Seq           CCAGGTGACAGATGAGGTGCAGACATCTTCCCAACGGTCACCCAATTTAGCACTCATCA 660
* *****

Parte_C1      CCTTGCAATCTGACATATCAGGTACTCTTTCTGAAAGC-ATGGTGTCAA-GTGACTCAAT 708
Seq           CCTTGCAATCTGACATATCAGGTACTCTTTCTGAAAGCCATGGTGTCAAAGTCAATCAAT 720
*****

Parte_C1      GGCAAAAGCACAGAG---TCTCTGAA--CCTCTGGGCC-AGAAGATAGCC-GGTTG- 757
Seq           GGCAAAAGCACAGAGTTCTCTGGAACCTTCTGGGCCAGAAAGATAGCCAGGTTGA 777
*****

```

Figura 16 Alinhamento das seqüências *in silico* e sequenciada do *Contig 1* (candidato para o gene VIN3) pelo programa ClustalW (THOMPSON et al., 1994)

### 3.2 qPCR

Todos os genes em estudo (referência e alvo) apresentaram eficiência dos *primers* acima de 90% (Tabela 2). Os genes de referência, utilizados no trabalho, foram descritos na literatura em diferentes condições experimentais para a cultura do café (BARSALOBRES-CAVALLARI et al. 2009).

Analisando a expressão relativa dos genes alvo através da utilização dos genes de referência *GAPDH* e *ACTINA* (Figura 17), observou-se que os genes alvos expressaram em todos os tecidos estudados, para o *FRI*, esse resultado é semelhante à literatura, pois os genes da família IV da superfamília *FRIGIDA* são expressos em diversos tecidos (RISK et al., 2010). Com relação ao *FLC*, os resultados indicam uma expressão constitutiva maior em folha e flor. Zhang et al. (2009) ao analisarem o padrão de expressão do *PtFLC* em *Poncirus*, observaram um perfil muito semelhante ao do café. Agora para o *VIN3* observou-se um maior padrão de expressão em folha e raiz, enquanto o *VRN2* teve uma expressão constante para todos os tecidos.

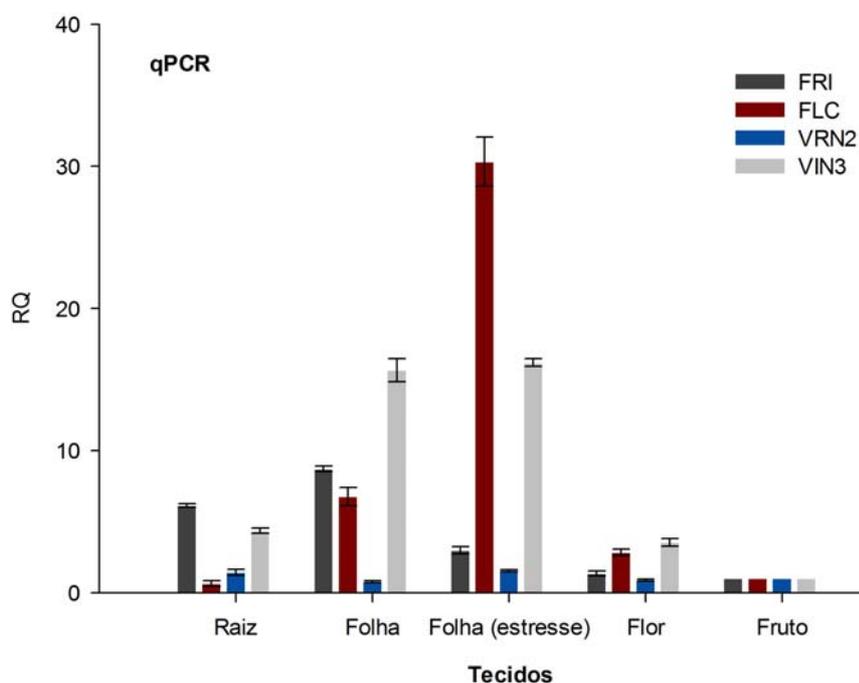


Figura 17 Perfil da expressão quantitativa relativa por qPCR dos genes *FLC*, *FRI*, *VRN2* e *VIN3* identificados em *Coffea arabica*. As colunas representam a expressão de transcritos gênicos em diferentes tecidos de plantas de café, com valores dados a partir da média dos valores de expressão das réplicas biológicas. Os transcritos gênicos foram normalizados pela expressão de múltiplos genes de referência (*GAPDH* e *ACTINA*)

Os resultados obtidos para os genes alvos *FRI* e *FLC* alvo são relevantes, pois a maior expressão de *FLC* foi observada para folhas de plantas submetidas a noventa dias de estresse hídrico antes do florescimento, enquanto o *FRI* apresentou maior expressão em folhas coletadas durante o florescimento. Isso é um indicativo de que o gene *FLC* desempenha a mesma função em cafeeiro do que os genes *FLC* em outras espécies, uma vez que sua expressão é elevada antes do florescimento, bem como que o gene *FRI* pode estar

relacionado a regulação da expressão de *FLC*, pois em folhas antes do florescimento a expressão de *FRI* é dez vezes menor do que a de *FLC*, ao passo que após o florescimento a expressão de *FRI* é superior.

#### 4 CONCLUSÃO

Por meio do sequenciamento, foi confirmada a sequência de todos os genes em estudo (*FLC*, *FRI*, *VRN2* e *VIN3*). Após a análise da expressão gênica, observou-se que todos os genes apresentam expressão em todos os tecidos analisados, com maior expressão do gene *FLC* em folha de plantas submetidas ao estresse hídrico antes do florescimento, enquanto o *FRI* apresentou maior expressão em folhas coletadas durante o florescimento.

## **5 CONSIDERAÇÕES FINAIS**

Com este estudo, possibilitou-se a caracterização de quatro dos principais genes envolvidos na regulação do florescimento em *Coffea arabica*, o que com estudos complementares com mutantes, em espécies-modelo, possibilitarão uma maior compreensão dos processos regulatórios bem como a função específica de cada gene durante a indução floral.

## REFERÊNCIAS

ALTSCHUL, S. F. et al. Gapped BLAST and PSI BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, Oxford, v. 25, n. 17, p. 3389-3402, Sept. 1997.

BARSALOBRES-CAVALLARI, C. F. B. et al. Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions. **Bmc Molecular Biology**, v. 10, n. 1, p. 1-11, Jan. 2009.

CIANNAMEA, S. Molecular characterization of the vernalization response in *lotion perenne*. 2007. 147p. Tese (Doctorate in Plant Science) - Wageningen University, Wageningen.

FINNEGAN, E. J. et al. Cluster of *Arabidopsis* genes with a coordinate response to an environmental signal. **Current Opinion in Plant Biology**, London, v. 14, n. 10, p. 911-916, May. 2004.

GUIMARÃES, R. J.; MENDES, A. N. G. **Morfologia e fisiologia do cafeeiro**. Lavras: UFLA, 1998. 38p.

KOORNNEEF, M. et al. Genetic control of flowering time in *Arabidopsis*. **Annual Review of Plant Physiology and Plant Molecular Biology**, Palo Alto, v. 49, p. 345-370, Jun. 1998.

MORAIS, H. et al. Escala fenológica detalhada da fase reprodutiva de *coffea arabica*. **Bragantia**, Campinas, v. 67, n. 1, p. 257-260, jan. 2008.

RATCLIFFE, O. J. et al. Regulation of flowering in *Arabidopsis* by an FLC Homologue. **Plant Physiology**, Rockville, v. 126, p. 122-132, May. 2001.

RISK, J. M. et al. FRIGIDA and related proteins have a conserved central domain and family specific N- and C- terminal regions that are functionally

important. **Plant Molecular Biology**, New Zealand, v. 73, p. 493–505, Apr. 2010.

SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 74, n. 12, p. 5463-5467, Jun. 1977.

SHINDO, C. et al. Role of FRIGIDA and FLOWERING LOCUS C in determining variation in flowering time of *Arabidopsis*. **Plant Physiology**, Rockville, v. 138, n. 2, p. 1163- 1173, Jun. 2005.

SIMPSON, G.G.; DEAN, C. *Arabidopsis*, the rosetta stone of flowering time. **Science**, Norwich, v. 296, n. 5566, p. 285-289, Apr. 2002.

THOMPSON, J.D. et al. Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Research**, Oxford, v. 22, p.4673–4680, Nov. 1994.

ZHANG, J.Z. et al. *PtFLC* homolog from trifoliolate orange (*Poncirus trifoliata* ) is regulated by alternative splicing and experiences seasonal fluctuation in expression level. **Plant Journal**, Oxford, v. 299, n. 4, p. 847-859, Mar. 2009.