

**IDENTIFICAÇÃO E EXPRESSÃO DO GENE  
*FLOWERING LOCUS C (FLC)* EM CAFÉ  
ARÁBICA**

**FABIANE LAZZARI**

**2009**

FABIANE LAZZARI

**IDENTIFICAÇÃO E EXPRESSÃO DO GENE  
*FLOWERING LOCUS C (FLC)* EM CAFÉ  
ARÁBICA**

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



Orientador

Prof. Antonio Chalfun Júnior, PhD



LAVRAS  
MINAS GERAIS – BRASIL  
2009

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

Lazzari, Fabiane.

Identificação e expressão do gene *Flowering Locus C (FLC)* em  
café arábica / Fabiane Lazzari. – Lavras : UFLA, 2009.

72 p. : il.

Dissertação (Mestrado) – Universidade Federal de Lavras, 2009.

Orientador: Antonio Chalfun Júnior.

Bibliografia.

1. Florescimento. 2. *Coffea arabica*. 3. MADS-box. 4. Expressão  
gênica. 5. Desenvolvimento. I. Universidade Federal de Lavras. II.  
Título.

CDD – 633.7389

FABIANE LAZZARI

**IDENTIFICAÇÃO E EXPRESSÃO DO GENE  
*FLOWERING LOCUS C (FLC)* EM CAFÉ  
ARÁBICA**

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

APROVADA em 13 de fevereiro de 2009

Prof. Dr. Luciano Vilela Paiva

UFLA

Prof. Dr. Antônio Paulino da Costa Netto

UEMG



Prof. Antonio Chalfun Júnior, PhD

UFLA

(Orientador)

2009

À Deus, pela força e serenidade nos momentos mais difíceis,  
Aos meus avós Idaci e Carlos (*in memoriam*) e Cidália e João,

## Ofereço

Aos meus pais, Roberto e Salete

Aos meus irmãos, Luciane e Roberto Jr,

Com amor dedico.

## SUMÁRIO

RESUMO GERAL.....	i
ABSTRACT.....	ii
CAPÍTULO 1.....	1
INTRODUÇÃO.....	1
1 INTRODUÇÃO.....	2
2 REFERENCIAL TEÓRICO.....	4
2.1 <i>Coffea arabica</i> .....	4
2.2 Fisiologia do cafeeiro: Floração.....	4
2.2.1 Fotoperíodo.....	7
2.2.2 Vernalização.....	8
2.3 Genética do Florescimento.....	9
2.4 Fatores de Transcrição: Genes MADS- <i>box</i> .....	11
2.5 <i>Flowering Locus C (FLC)</i> .....	14
3 REFERÊNCIAS BIBLIOGRÁFICAS.....	20
CAPÍTULO 2.....	27
IDENTIFICAÇÃO <i>IN SILICO</i> DO GENE <i>FLOWERING LOCUS C</i> NO BANCO CAFEST.....	27
RESUMO.....	28
ABSTRACT.....	29
1 INTRODUÇÃO.....	30
2 MATERIAIS E MÉTODOS.....	32
2.1 Identificação de sequências <i>FLC</i> : análise <i>in silico</i> .....	32
2.2 Filogenia.....	33
2.3 Identificação de Motivos Comuns de Agrupamento.....	33
2.4 <i>Northern</i> eletrônico.....	34
3 RESULTADOS E DISCUSSÃO.....	35

4 CONCLUSÕES.....	44
5 REFERÊNCIAS BIBLIOGRÁFICAS.....	44
CAPÍTULO 3.....	48
ISOLAMENTO E CLONAGEM DO GENE <i>FLOWERING LOCUS C</i> EM CAFÉ.....	48
RESUMO.....	49
ABSTRACT.....	50
1 INTRODUÇÃO.....	51
2 MATERIAIS E MÉTODOS.....	52
2.1 Extração e quantificação de RNA.....	52
2.2 Desenho dos primers.....	54
2.3 Obtenção da sequência cDNA.....	55
2.4 Eletroforese e Purificação do produto da amplificação (cDNA) por PCR.....	56
2.5 Ligação com vetor.....	57
2.6 Transformação e clonagem bacteriana.....	58
2.7 Extração do DNA plasmidial Miniprep.....	60
2.8 Digestão com enzimas de restrição.....	61
2.9 Sequenciamento.....	61
2.10 Análises das sequências em banco de dados.....	63
2.11 Análise de expressão tecido específico do gene <i>FLC</i> por qRT-PCR.....	63
3 RESULTADOS E DISCUSSÃO.....	65
4 CONCLUSÕES.....	70
5 REFERENCIAS BIBLIOGRÁFICAS.....	71

## AGRADECIMENTOS

À Deus, por capacitar-me em todos os momentos da condução deste trabalho.

À Universidade Federal de Lavras, em especial ao Curso de Biotecnologia Vegetal, pela oportunidade de realização do curso.

À CAPES, pela concessão de bolsa.

Ao Laboratório Central de Biologia Molecular da UFLA, pela realização do experimento, e aos laboratoristas.

Ao meu orientador, professor Dr. Antonio Chalfun Júnior, pela dedicação, respeito, conselhos e amizade, dedicando seu precioso tempo para orientar-me.

Ao prof. Dr. Luciano, pela oportunidade e por não medir esforços para que os seus objetivos e o de seus alunos fossem alcançados.

Aos professores do curso de Biotecnologia vegetal, pelos ensinamentos.

À Embrapa Recursos Genéticos e Biotecnologia/Cenargen e ao Dr. Alan Carvalho de Andrade, pela oportunidade.

Ao Dr. Ivan Schuster e à Dr. Elisa Serra Negra Vieira, por incentivarem-me a fazer o mestrado e pelos ensinamentos, eles que são pesquisadores nos quais me espelho.

Aos membros da banca examinadora.

À minha família, meus pais Roberto e Salete e meus irmãos Luciane e Roberto Junior, pelo apoio e carinho constantes.

Ao meu amor Pedro Ivo, pelo companheirismo, confiança e amor, me apoiando e dando força pra continuar todos os dias.

À minha amiga Sandra, pela amizade, carinho, paciência e companhia.

À Fabiana e ao Anderson, por toda ajuda nos experimentos, pela amizade e paciência.

Aos meus queridos amigos Felipe, Raphael, Geraldo, Rosane, Suzana, Glacy, Danielle, Evânia, Solange e Gustavo pela convivência, amizade e apoio.

Aos colegas de curso e de laboratório, pelo convívio e amizade.

Por fim, a todos que, de alguma forma, contribuíram para a realização desse trabalho.

**MUITO OBRIGADA**

## RESUMO GERAL

LAZZARI, Fabiane. **Identificação e expressão do gene *Flowering Locus C* (*FLC*) em café arábica.** 2009. 72p. Dissertação (Mestrado em Agronomia/Biotecnologia Vegetal) – Universidade Federal de Lavras.\*

Este estudo teve como objetivo estudar um dos fenômenos envolvidos no florescimento em cafeeiros, através da análise do gene *Flowering Locus C* (*FLC*), principal repressor da transição floral em *Arabidopsis*. Os experimentos foram conduzidos no Laboratório Central de Biologia Molecular da Universidade Federal de Lavras, MG e na EMBRAPA Recursos Genéticos e Biotecnologia/Cenargen em Brasília. A primeira etapa envolveu uma análise *in silico* para obtenção da sequência *FLC* de *C. arabica* através do banco CAFEST de EST (*Expressed Sequence Tags*). As prováveis sequências *FLC* obtidas foram anotadas e submetidas à análise filogenética, e validadas quanto aos domínios conservados e bibliotecas de expressão. O único EST-*contig* de café pertencente à família *FLC* foi o C15. Primers específicos foram desenhados para clonar o gene e para análise de expressão tecido específica de qRT-PCR. O RNA total foi obtido a partir de tecidos de folha, raiz, fruto e flor de plantas de café. A partir do RNA total, o cDNA foi sintetizado através da transcriptase reversa e a sequência obtida após o sequenciamento foi comparada com aquela obtida *in silico*, apresentando similaridade, o que demonstra que a análise preliminar utilizando-se da bioinformática é uma ferramenta muito poderosa em estudos moleculares. Avaliando-se os níveis de expressão nos diferentes tecidos, foi possível observar que a expressão quantitativa relativa do gene *FLC* se dá diferencialmente entre os tecidos, sendo que em folhas foram observados os mais elevados níveis de expressão.

---

\* Orientador: Antonio Chalfun Júnior, PhD - UFPA

## ABSTRACT

LAZZARI, Fabiane. **Identification and expression of the gene *Flowering Locus C (FLC)* arabica coffee.** 2009. 72p. Dissertation (Master in Agronomy/ Plant Biotechnology) - Federal University of Lavras.\*

This work aimed at to investigate one of the phenomena involved in flowering in coffee plants, studying the *Flowering Locus C* gene (*FLC*), the major repressive floral transition in *Arabidopsis*. The experiments were carried out at the Central Laboratory of Molecular Biology, at the Federal University of Lavras, in Lavras, MG and at EMBRAPA Genetic Resources and Biotechnology/Cenargen in Brasilia, DF. *In silico* analysis was performed to obtain *FLC C. arabica* sequences using the CAFEST (Coffee EST database). The putative *FLC* sequences obtained were annotated and submitted to phylogenetic analysis. Conserved domains were identified and electronic Northern was performed. The only EST-contig which had similarity to *FLC* family was the C15. Specific primers were designed to clone the gene and tissue-specific expression analyses were carried out by qRT-PCR. Total RNA was extracted from leaf, root, flower and fruit of coffee plants to obtain the cDNA sequence that was used to compare to the *FLC in silico* one. Based on that, it was possible to demonstrate similarity of both sequences, indicating that the preliminary analysis using bioinformatics is a very powerful tool in molecular studies. In addition, based on the qRT-PCR results, it was possible to observe that the relative quantitative expression of the *FLC* occurs differentially among the tissues and in leaves, the highest level of expression was detected.

---

\* Adviser: Antonio Chalfun Júnior, PhD - UFLA

## **CAPÍTULO 1**

### **INTRODUÇÃO**

# 1 INTRODUÇÃO

O cafeeiro (*Coffea sp.*) pertence à família Rubiaceae e é formado por mais de noventa espécies. É amplamente cultivado em países tropicais e, das espécies cultivadas, *Coffea arabica L.* e *Coffea canephora* são as mais importantes economicamente.

Dentre as *commodities* naturais, o café tem seu valor ultrapassado apenas pelo petróleo, pois é a segunda mercadoria mais comercializada do mundo. O Brasil destaca-se como o maior produtor mundial de café e sua produção está relacionada diretamente ao desenvolvimento e à industrialização do país. Entretanto, o problema do florescimento sequencial dos botões florais causa uma maturação desuniforme dos frutos, dificultando a colheita e causando perdas econômicas de produção.

A identificação e compreensão dos mecanismos do florescimento, bem como o conhecimento de técnicas viáveis de uniformização da floração são de fundamental importância para a organização e redução dos custos da colheita e para manutenção da produção agrícola brasileira e mundial.

Estudos genéticos e moleculares estabeleceram a presença de quatro rotas de desenvolvimento distintas que controlam o florescimento em *Arabidopsis thaliana*, permitindo a sincronização do tempo de florescimento, que são: comprimento do dia (fotoperíodo), tratamento de frio (vernalização), via autônoma e das giberelinas (Blasquez, 2000).

Além do comportamento fenológico da cultura, vários fatores do ambiente, tais como, suprimento de água, temperatura do ar e níveis de irradiância influenciam no florescimento, desenvolvimento dos frutos e produtividade do cafeeiro. No entanto a mudança da fase vegetativa para a fase reprodutiva é uma alteração crítica no ciclo das plantas, e a regulação molecular dessa transição é essencial para o desenvolvimento e produção de frutos e

sementes. Assim, as plantas evoluíram para integrar sinais endógenos e ambientais, para controlar precisamente o tempo de florescimento.

Um amplo número de genes regulatórios que fazem parte de vias de interação que controlam o tempo de florescimento em *A. thaliana* tem sido identificado (Simpson & Dean, 2002). Em *Arabidopsis*, a vernalização é controlada em por dois locos principais: *FLOWERING LOCUS C (FLC)* e *FRIGIDA (FRI)* (Shindo et al., 2005).

*Flowering Locus C* é um fator de transcrição *MADS-box* e altos níveis de sua expressão desencadeiam a repressão da transição floral. Em diferentes ecotipos responsivos à vernalização e em mutantes para tempo de florescimento em *Arabidopsis* os níveis de mRNA e proteína estão relacionados com o tempo de florescimento em resposta ao tratamento de frio (Michaels & Amasino, 1999). Segundo Dennis & Peacock (2007), a redução da atividade de *FLC* e o florescimento precoce também podem ser induzidos através da ação de mecanismos epigenéticos, tal como uma redução no nível de metilação no DNA.

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.

Neste contexto, o objetivo do presente trabalho foi estudar um dos principais fenômenos endógenos envolvidos no florescimento em cafeeiros, através de análise do gene *FLOWERING LOCUS C*, principal repressor da transição floral em *Arabidopsis*.

## 2 REFERENCIAL TEÓRICO

### 2.1 *Coffea arabica*

O cafeeiro (*Coffea sp.*) é um arbusto da família Rubiaceae e do gênero *Coffea L.* e conta com cerca de cem espécies identificadas. Todas as espécies conhecidas, exceto *Coffea arabica*, são diplóides ( $2n=2x=22$  cromossomos). *Coffea arabica L.* ( $2n=4x=44$ ), a única espécie poliplóide no gênero *Coffea*, é um alotetraplóide contendo dois subgenomas diplóides, e cuja origem vem de duas diferentes espécies diplóides, *C. canephora* e *C. eugenioides* (Lashermes et al., 1999).

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, numa 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 a partir da compreensão da interação de fatores que afetam essa fase de desenvolvimento da planta.

### 2.2 Fisiologia do cafeeiro: Floração

O florescimento do café envolve uma complexa sequência de eventos bioquímicos, fisiológicos e morfológicos, os quais são afetados por vários fatores como temperatura, luz, solo, ar e disponibilidade de água (Rena & Barros, 2004).

A mudança da fase vegetativa para a fase reprodutiva é uma alteração crítica no ciclo das plantas, crucial para a produção de frutos e sementes.

O cafeeiro, *Coffea arabica* L., completa seu ciclo produtivo em dois anos. Um esquema detalhado da fenologia do cafeeiro é apresentado (FIGURA 1) conforme Camargo & Camargo (2001). São seis fases fenológicas distintas, sendo duas vegetativas e quatro reprodutivas que ocorrem em dois anos seguidos, detalhadas a seguir.

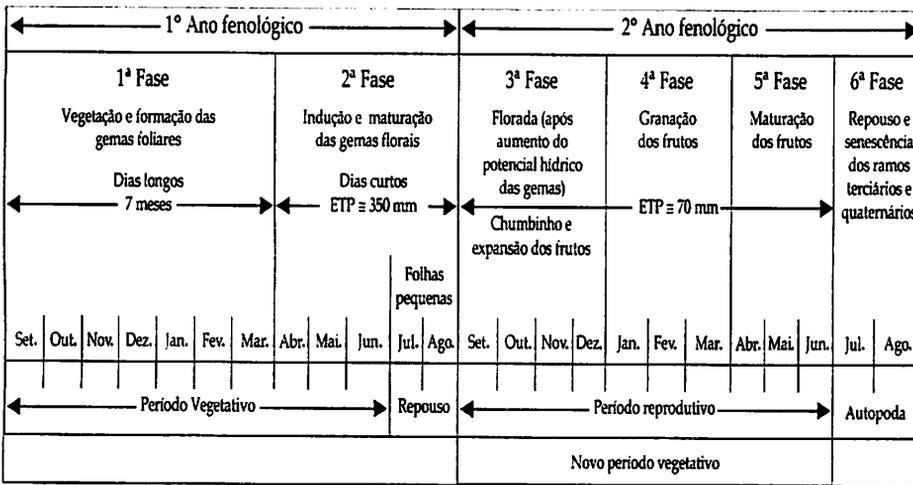


FIGURA 1. Esquemática das seis fases fenológicas do cafeeiro arábica, durante 24 meses, nas condições climáticas tropicais do Brasil (Adaptado de Camargo & Camargo, 2001).

No primeiro ano fenológico, a fase inicial que consiste na vegetação e formação das gemas foliares, vai de setembro a março. São sete meses de dias longos, com fotoperíodo acima de 13 e 14 horas de luz efetiva ou acima de 12 horas de luz (CAMARGO, 1985). A segunda fase, indução, maturação e dormência das gemas florais, é de dias curtos (CAMARGO & FRANCO, 1985).

A terceira fase é a primeira do segundo ano fenológico. Inicia-se com a florada após um aumento do potencial hídrico nas gemas florais maduras (choque hídrico). Após a fecundação, vem os chumbinhos e a expansão dos

frutos. Essa etapa compreende os quatro meses, de setembro a dezembro (Camargo & Camargo, 2001). No cafeeiro, o florescimento é promovido por dias curtos e por um período de frio (vernalização).

A quarta fase é a de granação dos frutos, quando os líquidos internos solidificam-se, dando formação aos grãos. Ocorre em pleno verão, de janeiro a março. A maturação dos frutos se dá na quinta fase, compreendendo normalmente os meses de abril, maio e junho. As duas primeiras fases correspondem ao período vegetativo e as três seguintes (terceira, quarta e quinta) correspondem ao período reprodutivo do cafeeiro.

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 e 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).

Além da hierarquia temporal e fisiológica da iniciação e diferenciação do botão de floral dentro de cada ramo, e entre diferentes nós do mesmo ramo plagiotrópico, a ocorrência de chuvas esporádicas durante as fases anteriores do desenvolvimento do botão floral são tidas como responsáveis por desencadear os vários períodos de floradas em *C. arabica* (Rena e Barros, 2004). Por essa razão, um fator importante e polêmico relacionado à irrigação do cafeeiro é a discutida necessidade de um déficit hídrico para a quebra de dormência do botão floral para indução da floração, proporcionando uma florada uniforme, sem causar danos à produção.

Entretanto, ainda não se tem conhecimento fisiológico suficiente para prover melhor integração entre a duração do déficit hídrico e a irrigação que

poderia maximizar o número de gemas florais maduras e, por extensão, uma floração e desenvolvimento de frutos mais uniformes (Soares et al., 2005).

A indução floral é um evento que ativa os processos que permitem que a planta esteja apta para o florescimento, por meio da transição e da expressão de genes, que ocorrem antes da iniciação (Bernier et al., 1981). Pode ser influenciada por um fator específico do ambiente como a duração do dia (fotoperíodo) ou por vários fatores atuando em conjunto (água, luz, nutrientes, temperatura).

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, ou seja, desencadeiam o desenvolvimento da fase vegetativa para reprodutiva (Guimarães & Mendes, 1998).

### 2.2.1 Fotoperíodo

As duas principais classes de respostas fotoperiódicas são plantas de dias curtos (PDC) e de dias longos (PDL). As plantas de dias curtos florescem no início da primavera ou do outono, especificamente em dias mais curtos (crisântemos, café, bico de papagaio, morangos, primulas) e as plantas de dias longos florescem principalmente no verão, onde temos os dias mais longos (espinafre, algumas batatas, certas variedades de trigo; alface, aveia, cravo, ervilha) (Taiz & Zeiger, 2006).

O primeiro grupo de plantas floresce quando comprimento do dia for menor ao seu fotoperíodo crítico. PDL somente florescem quando o comprimento do dia for maior ao seu fotoperíodo crítico. Uma consequência dessa definição é que PDL conseguem florescer em luz contínua. Fotoperíodo crítico é o valor em hora diária de iluminação capaz de provocar a floração, no

entanto, é o período de escuro que induz a floração. Por exemplo, PDL com fotoperíodo crítico igual 18 horas, deve florescer em períodos diários de iluminação superiores a 18 horas ou em períodos diários de escuro iguais ou inferiores a 6 horas. Existem ainda as plantas indiferentes, também chamadas neutras, que florescem sem nenhuma relação com o comprimento do dia, como por exemplo, o milho e o tomate.

O florescimento é, portanto, um ponto essencial para a sobrevivência de muitas plantas, que devem crescer em um ambiente que proporcione os estímulos necessários para a floração. Espécies que florescem em dia longo, de regiões temperadas, não florescem nos dias curtos dos trópicos. As exigências de resfriamento de plantas anuais de inverno, bianuais e perenes também não são satisfeitas em ambientes tropicais.

Entretanto, não somente o fotoperiodismo é um mecanismo importante responsável por essas respostas estacionais, a vernalização também constitui um dos mecanismos de destaque, pois consiste na indução ou aceleração da floração por temperatura baixa. Estes dois sistemas de controle permitem a sincronização da reprodução das plantas.

### 2.2.2 Vernalização

É um processo quantitativo e depende da extensão do tratamento de frio, o estágio da planta e da temperatura (Ciannamea, 2007).

A vernalização pode ser classificada em dois tipos de acordo com a idade em que planta responde a baixas temperaturas. Um tipo é a vernalização da semente, na qual plantas podem responder a baixas temperaturas durante a germinação da semente; o outro tipo é a vernalização da planta, na qual plantas precisam atingir um determinado estágio de desenvolvimento antes de tornarem-se sensíveis a baixas temperaturas. Plantas que crescem vegetativamente no

primeiro ano e florescem no ano seguinte depois do inverno geralmente fazem parte do tipo de plantas-responsivas a vernalização. Algumas espécies na família Brassicaceae, como *Arabidopsis* e *Brassica rapa*, são do tipo semente-responsiva, mas diversas variedades em *Brassica oleracea* são planta-responsiva (Friend, 1985; Lin et al., 2005). Lin et al., (2005) utilizaram uma construção *FLC::GUS* em *Brassica oleracea* variedade Capitata, que compreende a fusão da sequência genômica *BoFLC* com *GUS*, no ecotipo C24 de *Arabidopsis*, concluindo que há uma resposta variável para vernalização da planta e da semente.

Estudos genéticos e moleculares de *Arabidopsis* têm revelado que o tempo do florescimento é influenciado por muitos genes que fazem parte de vias de interação (Koornnef et al., 1994, 1998; Simpson & Dean, 2002; Blasquez, 2000). Em diferentes ecotipos responsivos à vernalização e em mutantes para tempo de florescimento em *Arabidopsis* os níveis de mRNA e proteína estão relacionados com o tempo de florescimento em resposta ao tratamento de frio (Michaels e Amasino, 1999).

### 2.3 Genética do florescimento

A regulação correta da transição para o florescimento é essencial para o desenvolvimento reprodutivo das plantas, dessa forma, elas desenvolveram mecanismos moleculares conservados para integrar tanto indícios ambientais como endógenos para regular precisamente o tempo de florescimento (Dornelas & Rodriguez, 2005 e Ciannamea, 2007). 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).

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).

Em um primeiro trabalho, hipóteses foram sugeridas para explicar como a especificação da identidade do órgão durante o desenvolvimento floral é controlada pelos genes seletores homeóticos (Haughn & Somerville, 1988; Schwarz-Sommer et al., 1990; Coen & Meyerowitz, 1991). Em um modelo clássico, três classes de atividades de genes homeóticos “função homeótica” foram propostas, chamadas A, B e C (Coen & Meyerowitz, 1991). Dentro de qualquer um dos quatro verticilos florais, a expressão de somente A especifica a formação de sépalas. A combinação AB especifica o desenvolvimento de pétalas, e a combinação BC especifica a formação de estames. A expressão somente de C determina o desenvolvimento de carpelos. O modelo também propõe que a função de genes A e B regulam negativamente um ao outro (significando que eles também exercem funções cadastrais) e que a função de B é limitada para o segundo e o terceiro verticilo independentemente das funções A e C. Em flores tipo selvagem, a função A é expressa no primeiro e segundo verticilo floral, a função B no segundo e terceiro verticilo, e a função C no terceiro e quarto verticilo. Por essa razão, sépalas, pétalas, estames e carpelos são específicos do verticilo um, dois, três e quatro, respectivamente.

Segundo Coen & Meyerowitz (1991), o modelo ABC foi largamente baseado em análises de mutantes de *Arabidopsis*, embora *Antirrhinum* também tenha sido considerada (Weigel & Meyerowitz, 1994).

Com base em estudos em petúnia (*Petunia hybrida*), o modelo ABC foi depois estendido para incluir a função D, criando-se um “modelo ABCD” (Colombo et al, 1995). Entretanto, uma nova função, nomeada E, foi adicionada, resultando no modelo revisado ABCDE (Pelaz et al., 2000, 2001; Pinyopich et al., 2003). No modelo revisado, a função D determina o desenvolvimento do óvulo, enquanto a função E é exigida para o desenvolvimento de todos os órgãos florais.

A rede de genes que controla a transição floral está envolvida na ativação de genes de identidade de meristemas, tais como *LEAFY* (*LFY*) e *APETALLA* (*AP*), os quais consequentemente regulam as três classes de genes homeóticos florais, A, B e C (Weigel & Meyerowitz, 1994).

Exceto o gene *AP2* (*APETALLA2*), todos os genes ABC tem uma parte altamente conservada, de aproximadamente 180 pares de bases de sequência de DNA, nomeados *MADS-box*. Essa sequência codifica um domínio de ligação de DNA do respectivo fator de transcrição do domínio MADS (Riechmann & Meyerowitz, 1997; Theissen et al., 2000).

#### 2.4 Fatores de Transcrição: Genes *MADS-box*

Nos anos mais recentes, tem sido descoberto que a chave do desenvolvimento do controle de genes são frequentemente membros de um número muito limitado de famílias multigênicas, as quais codificam fatores de transcrição. Fatores de transcrição são importantes reguladores de processos celulares, e a complexidade de organismos vivos exige um amplo número de fatores de transcrição (Theissen et al. 2000; Parenicová et al. 2003).

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.

MADS são as iniciais das quatro primeiras proteínas encontradas nessa família: MCM1 (da levedura fermentadora, *Saccharomyces cerevisiae*), AGAMOUS (de *Arabidopsis*), DEFICIENS (de *Antirrhinum*), e SRF (uma proteína de humano), nas quais a definição dessa família gênica é baseada (Schwarz-Sommer et al., 1990).

Theissen e colaboradores (2000) argumentaram que a flor é um modelo ideal para a evolução vegetal, e que a filogenia de genes *MADS-box* pode ter iniciado uma importante função durante a origem e evolução do desenvolvimento floral.

O domínio MADS é a região mais conservada de proteínas (em torno de 58 aminoácidos), e estas, como muitos fatores de transcrição eucarióticos, têm uma organização estrutural modular que ativam os processos de transcrição ligando-se a elementos de reconhecimento chamados CA<sub>n</sub>G boxes (CC(A/T)<sub>6</sub>GG), encontrados em promotores de genes alvo.

O único denominador comum de todos os genes *MADS-box* são os altamente conservados 180 pb, codificando o domínio DNA-*binding* desses fatores de transcrição. Segundo Alvarez-Buylla et al. (2000), a família de genes *MADS-box* pode ser subdividida em duas classes denominadas tipo I (*SRF-like* de *SERUM RESPONSE FACTOR-like*) e o tipo II (*MEF2-like* de *MYOCYTE ENHANCER FACTOR2-like*). Ambos os tipos de genes têm sido encontrados em animais, fungos e plantas.

Pouco é conhecido sobre a função de genes tipo I em plantas, e todos os genes MADS bem caracterizados contribuintes para o desenvolvimento vegetal são do tipo II (Alvarez- Buylla et al., 2000).

Genes MADS tipo II de plantas apresentam uma organização estrutural conservada, denominada domínio estrutural tipo-MIKC, que inclui quatro domínios característicos distribuídos no sentido N a C terminal da proteína: o domínio MADS-*box* (M), *intervening* (I), *keratin-like* (K), e C-terminal (C) (Hayes et al., 1988; Purugganan et al., 1995; Riechmann et al., 1996; Shore & Sharrocks, 1995). O domínio I é pouco conservado e contribui para a especificação da dimerização. O domínio K é caracterizado pela estrutura enrolada (“coiled-coil”), que facilita a dimerização das proteínas MADS-*box* (Davies et al., 1996; Fan et al., 1997). O domínio C é o menos conservado, em alguns casos, tem mostrado contribuir para a formação complexos multiméricos de proteínas MADS-*box* (Egea-Cortines et al., 1999; Honma & Goto, 2001).

Além de determinar funções homeóticas florais, genes MADS-*box* tem 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 (Theissen et al., 2000; Parenicová et al. 2003; Rijpkema et al. 2007). *Flowering Locus C*, por exemplo, é um gene mediador do tempo de florescimento, dependente de fatores ambientais como o frio, que marca o desenvolvimento da fase vegetativa para reprodutiva (Michaels and Amasino 1999).

## 2.5 Flowering Locus C (*FLC*)

O início do florescimento é induzido pelo cruzamento de distintas vias genéticas, como fotoperíodo, giberelina, autônoma e vernalização (FIGURA 2). *FLC* é o maior repressor floral da via de integração floral em *Arabidopsis* e quando ativo conduz a ativação de genes de identidade de meristemas. A expressão de *FLC* é positivamente regulada por *Frigida*, *Frigida Like1*, *Frigida Like2*, *Vernalization Independence4*, *Early Flowering7*, *Early Flowering8/VIP6*, e outros. Reciprocamente, a expressão de *FLC* é inibida pela vernalização e por genes como *VRN1*, *VRN2* e *VIN3*. Ainda, genes da via autônoma como *FCA*, *FY*, *FLD*, *FVE*, *FPA*, *LD*, etc reprimem a transcrição de *FLC* (Ciannamea, 2007).

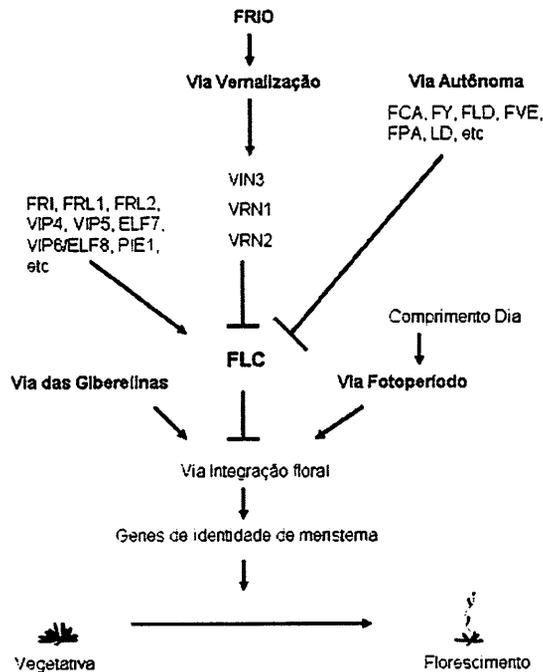


FIGURA 2. Vias de controle do tempo de florescimento em *Arabidopsis* (Ciannamea, 2007).

Em *Arabidopsis*, a vernalização é controlada por dois loci principais: *FLOWERING LOCUS C (FLC)* e *FRIGIDA (FRI)* (Shindo et al., 2005).

A promoção do florescimento pela vernalização envolve a modulação da expressão de *FLC*. O locus *FLC* foi identificado em *Arabidopsis* através de um fenótipo mutante. Michaels e Amasino (1999) determinaram o efeito do tratamento de frio nos níveis de mRNA *FLC*. Sementes de *FRI-SF2* linhagem Columbia (Col) foram submetidas ao tratamento de frio por 30 dias e então foram transferidas para o solo com sementes de plantas controle não expostas ao frio. As sementes foram coletadas após 10 dias de crescimento para análises moleculares. O nível de transcritos *FLC* foi reduzido depois da vernalização para níveis indetectáveis, o que se mostrou similar para com a diferença na expressão encontrada entre *FRI-SF2* em Col e mutantes *fri* nulo. Então, a perda do fenótipo de atraso do florescimento de *FRI-SF2* em Col pela vernalização é acompanhada por um correspondente declínio na expressão de *FLC*.

Através de um segundo modelo, Michaels e Amasino (2001) demonstraram que *FLC* tem uma função central na regulação do tempo de florescimento em *Arabidopsis* pela integração de sinais de *FRI*, via autônoma da promoção floral e da via de vernalização. Ainda, mostraram que os efeitos de *FRI* e vias autônomas mutantes no florescimento e a expressão de *SOC1* são dependentes de um alelo funcional de *FLC*. *FRI* e *FLC* interagem sinergicamente para atrasar o florescimento, e a perda da função de um dos dois genes elimina o fenótipo de florescimento tardio.

Durante a vernalização, diminuem os níveis de transcritos e proteínas *FLC* e a planta torna-se competente a florescer (Michaels & Amasino, 1999; 2000; Johanson et al., 2000). Ainda, a superexpressão de *FLC* pelo promotor CAMV 35S em acessos *Landsberg* de *Arabidopsis* (o qual não apresenta o alelo *FRI* ativo) é suficiente para modificar severamente o tempo de florescimento e

tornar a planta insensível a vernalização (Michaels & Amasino, 1999; Sheldon et al., 1999). Essas descobertas indicam que *FLC* é um potente repressor floral.

Desde que *FLC* seja reprimido pela vernalização, ele pode ser reativado na próxima geração. Os mecanismos de repressão de *FLC* pela vernalização têm sido amplamente estudados. A maioria dos efeitos moleculares precedentes a 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 transduzido a modificações na cromatina em locus *FLC*.

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 através do nível de repressão de *FLC*, e que a atividade de *VRN1* (*VERNALIZATION1*) e *VRN2* (*VERNALIZATION2*) 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). *FVE*, um repressor de *FLC* sob condições não-vernalizadas, tem sido identificado como um regulador negativo da expressão gênica responsiva ao frio (Kim et al., 2004). *FLC* regula negativamente a expressão de *SUPPRESSOR OF OVEREXPRESSION OF CO1* (*SOC1*), que codifica um fator de transcrição *MADS-box* e é quantitativamente induzido pela vernalização de forma recíproca a *FLC* (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.

Em seus estudos com vernalização, Finnegan et al., (2004) encontraram um pequeno grupo de genes, incluindo *FLC* e o gene *Upstream of FLC* (*UFC*),

que são coordenadamente regulados em resposta a diferentes modificações genéticas em mutantes do florescimento em *Arabidopsis*. O gene *UFC*, com função desconhecida, é localizado 4,7 Kb *upstream* de *FLOWERING LOCUS C*.

Sheldon e colaboradores (2007) demonstraram que a atividade de ambas as construções, *FLC::GUS* derivados paternalmente e maternalmente, contendo o gene repórter, são eliminadas após a repressão epigenética pela vernalização, mas o momento de sua expressão inicial é diferente.

*FLC* não é exclusivamente regulado pela vernalização. Um número adicional de reguladores positivos e negativos não relacionados tem sido descritos. Segundo Dennis & Peacock (2007), a redução da atividade de *FLC* e o florescimento precoce também podem ser induzidos através da ação de mecanismos epigenéticos, tal como uma redução no nível de metilação no DNA.

Quanto à organização genômica, em *Arabidopsis*, o gene *FLC* está organizado genomicamente em sete éxons, e o primeiro éxon contém o domínio conservado *MADS-box*, que constitui o primeiro domínio da estrutura MIKC (Michaels & Amasino, 1999). Em alguns acessos de *Arabidopsis*, alelos fracos de *FLC* possuem um transposon inserido no primeiro íntron (Gazzani et al., 2003).

A análise na sequência genômica de *Arabidopsis* revelou que, além de *FLC*, existem cinco genes *MADS-box* que codificam proteínas que são altamente relacionadas a ele e homólogas (53-87% de identidade), os *MADS AFFECTING FLOWERING 1-5 (MAF1-5)* (Bodt et al., 2003; Ratcliffe et al., 2003). O primeiro gene analisado, *MAF1* (também chamado *FLOWERING LOCUS M – FLM*), como *FLC*, é um importante repressor floral, apesar de sua expressão não ser tão fortemente influenciada pela vernalização. Os outros quatro loci foram nomeados *MAF2* a *MAF5*. A superexpressão de *MAF3* e *MAF4* produz alterações no tempo de florescimento, o que sugere que esses

também atuam como repressores florais e podem contribuir para necessidade da vernalização (Ratcliffe et al., 2001, 2003).

A repressão do desenvolvimento reprodutivo é realizada por *FLC* inibindo genes *downstream* a ele que promovem o florescimento, incluindo *SUPRESSOR OF OVEREXPRESSION OF CO1 (SOC1)* e *FLOWERING LOCUS T* (Borner et al., 2000; Onouchi et al., 2000; Michaels & Amasino, 2001). Assim, a promoção do florescimento ou pela via autônoma ou vernalização envolve a repressão de *FLC* e a subsequente depressão de alvos *FLC* (Ratcliffe et al., 2003).

Genes *FLC-like* vêm sendo descritos em várias espécies. Um estudo em *Brassica napus* demonstrou que seu *FLC* homólogo provoca atraso no florescimento quando introduzido em *Arabidopsis* (Tadege et al., 2001). O gene homólogo *FLC* de *Brassica oleracea* (repolho) foi clonado e sua expressão caracterizada em resposta a vernalização. Para isso, o gene *BoFLC* foi introduzido em *Arabidopsis* para investigar os diferentes mecanismos regulatórios envolvidos nas respostas vernalização da semente e da planta, tendo concluído que o florescimento em repolho pode ser promovido pela vernalização na planta adulta, mas não nas sementes (Lin et al., 2005). *B. oleracea*, como *Arabidopsis*, é uma planta dependente de vernalização na família Brassicaceae (Friend, 1985).

Genes *FLC-like* não haviam sido identificados em classes taxonômicas diferentes de *Arabidopsis*, *Brassica* e *Raphanus*, três gêneros dentro de Brassicaceae. Por esse motivo, inicialmente acreditava-se que genes *FLC-like* eram restritos a essa família (Schlappi & Patel 2001; Tadege et al. 2003; Lin et al., 2005; Hecht et al. 2005, Becker and Theissen, 2003). Contudo, em um recente trabalho, Reeves et al. (2007) identificaram um *FLC* homólogo *BvFLC* em beterraba (*Beta vulgaris ssp.*). *BvFLC* é reprimido pelo frio prolongado e atrasa o florescimento quando expresso em plantas transgênicas de *Arabidopsis*.

Recentemente, o gene *PtFLC* foi isolado de mutantes de florescimento precoce de *Poncirus trifoliata*, através do método RACE combinado com uma biblioteca de cDNA. A análise de PCR em Tempo Real mostrou a superexpressão de *PtFLC* durante o inverno, seguida por uma diminuição na primavera e verão (Zhang et al., 2009).

Tadege e colaboradores (2003) não conseguiram identificar a sequência *FLC-like* no genoma de arroz, porém, mostraram que a expressão ectópica de *FLC Arabidopsis* atrasa o florescimento nessa gramínea.

Mais pesquisas são necessárias para esclarecer como importantes genes *FLC-like* respondem à vernalização em espécies que não pertencem à família Brassicaceae.

### 3 REFERÊNCIAS BIBLIOGRÁFICAS

ALVAREZ-BUYLLA, E.R.; PELAZ, S.; LILJEGREN, S.J.; GOLD, S.E.; BURGEFF, C.; DITTA, G.S.; DE POUPLANA, L.R.; MARTINEZ-CASTILLA, L.; YANOFSKY, M.F. An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. **Proceedings of the National Academy of Sciences USA**, v.97, n.10, p.5328–5333, May 2000.

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**. v. 29, n.3, p.464-489, Dec. 2003.

BERNIER, G., KINET, J.M., SACHS, R.M. **The physiology of flowering**. Boca Raton: CRC, 1981. v.1-3.

BORNER, R.; KAMPMANN, G.; CHANDLER, J.; GLEISSNER, R.; WISMAN, E.; APEL, K.; MELZER, S. A MADS domain gene involved in the transition to flowering in Arabidopsis. **Plant Journal**, v.24, n.5, p.591–599, Sept. 2000.

BLASQUEZ, M.A. Flower development pathways. **Journal Cell Science**, Valencia, v.113, p.3547-3548, 2000.

BODT, S.D.; RAES, J.; PEER, Y.V.D.; THEISSEN, G. And then there were many: MADS goes genomic. **Trends in Plant Science**, v. 8, n.10, p. 475–483, Oct. 2003.

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, 1985.

CAMARGO, A.P.; FRANCO, C.F. Clima e fenologia do cafeeiro. In: INSTITUTO BRASILEIRO DO CAFÉ, MINISTÉRIO DA INDÚSTRIA E COMÉRCIO. **Cultura de café no Brasil**: manual de recomendações. 5.ed. Rio de Janeiro, 1985. p.19-50.

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, 2001.

- CIANNAMEA, S. **Molecular Characterization of the Vernalization Response in *Lolium Perenne***. 2007. 147p. Tese (doutorado) - Wageningen University, Wageningen.
- COEN, E.S.; MEYEROWITZ, E.M. The war of the worls: genetic interactions controlling flower development. **Nature**, v.353, p.31-37, Sept. 1991.
- COLOMBO, L.; FRANKEN, J.; KOETJE, E.; VAN WENT, J.; DONS, H.J.M.; ANGENENT, G.C.; VAN TUNEN, A.J. The petunia MADS box gene FBP11 determines ovule identity. **Plant Cell**, Baltimore, v.7, n.11, p.1859-1868, 1995.
- DAVIES, B.; EGEE-CORTINES, M.; DE ANDRADE SILVA, E.; SAEDLER, H.; SOMMER, H. Multiple interactions amongst floral homeotic MADS *box* proteins. **European Molecular Biology Organization Journal**, v. 15, p.4330–4343, 1996.
- DENNIS, E.S.; PEACOCK, W.J. Epigenetic regulation of flowering. **Current Opinion in Plant Biology**, v.10: 520-527 p. 2007.
- DORNELAS, M.C.; RODRIGUEZ, A.P.M. Identifying *Eucalyptus* expressed sequence tags related to *Arabidopsis* flowering time pathway genes. **Brazilian Journal Plant Physiology**, Piracicaba, v.17(2): 255-266 p. 2005.
- EGEE-CORTINES, M.; SAEDLER, H.; SOMMER, H. Ternary complex formation between the MADS-*box* proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in *Antirrhinum majus*. **European Molecular Biology Organization Journal**, v.18, 5370–5379, 1999.
- FAN, H.Y.; HU, Y.; TUDOR, M.; MA, H. Specific interactions between the K domains of AG and AGLs, members of the MADS domain family of DNA binding proteins. **Plant Journal**, v. 12, n.5, p.999–1010, Nov. 1997.
- FINNEGAN, E.J.; SHELDON, C.C.; JARDINAUD, F.; PEACOCK, W.J.; DENNIS, E.S. A Cluster of *Arabidopsis* genes with a coordinate response to an environmental signal. **Current Opinion in Plant Biology**, Canberra, v.14, p.911-916, 2004.
- FRIEND, D.J.C. Brassica. In AH Halevy, ed, Handbook of flowering, Vol II. CRC Press, Boca Raton, FL, p. 48-77. 1985.

GAZZANI, S.; GENDALL, A.R.; LISTER, C.; DEAN, C. Analysis of the Molecular Basis of Flowering Time Variation in *Arabidopsis* Accessions. **Plant Physiology**, Norwich, v. 132, p.1107-1114, June 2003.

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

HAUGHN, G.W.; SOMERVILLE, C.R. Genetic control of morphogenesis in *Arabidopsis*. **Developmental Genetics**, v. 9:73-89, (1988).

HAYES, T.E.; SENGUPTA, P.; COCHRAN, B.H. The human c-fos serum response factor and the yeast factor GRM/PRTF have related DNA-binding specificities. **Genes and Development**, v. 2, n. 12B, p.1713-1722, Dec. 1988.

HECHT, V.; FOUCHER, F.; FERRA'NDIZ, C.; MACKNIGHT, R.; NAVARRO, C.; MORIN, J.; VARDY, M.E.; ELLIS, N.; BELTRÁN, J.P.; RAINEAU, C.; WELLER, J.L. Conservation of *Arabidopsis* flowering genes in model legumes. **Plant Physiology**, v. 137: 1420–1434p. 2005.

HONMA, T.; GOTO, K. Complexes of MADS-*box* proteins are sufficient to convert leaves into floral organs. **Nature**, v. 409, p.525–529, Nov. 2001.

JOHANSON, U.; WEST, J.; LISTER, C.; MICHAELS, S.; AMASINO, R.; DEAN, C. 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.

KIM, H-J.; HYUN, Y.; PARK, J.Y.; PARK, M.J.; PARK, M.K.; KIM, M.D.; KIM, H.J.; LEE, M.H.; MOON, J. LEE, I.; KIM, J. A genetic link between cold responses and flowering time through FVE in *Arabidopsis thaliana*. **Nature Genetics**, korea, v.36, p.167-171, Jan. 2004.

KOORNNEEF, M.; VRIES, H.B.; HANHART, C.; SOPPE, W.; PEETERS, T. The phenotype of some late-flowering mutants is enhanced by a locus on chromosome 5 that is not effective in the Landsberg erecta wil-type. **Plant Journal**, Wageningen, The Netherlands, v.6: 911-919p. 1994.

KOORNNEEF, M.; ALONSO-BLANCO, C.; PEETERS, A.J.M.; SOPPE, W. Genetic control of flowering time in *Arabidopsis*. **Annual Review of Plant Physiology and Plant Molecular Biology**, Wageningen, v.49, p.345-370, June 1998.

LASHERMES, P.; COMBES, M.C.; ROBERT, J.; TROUSLOT, P.; D'HONT, A.; ANTONY, F.; CHARRIER, A. Molecular Characterization an origin of the *Coffea arabica* L. Genome. **Molecular and General Genetics**, v.261, p.259-266, 1999.

LEVY, Y.Y.; MESNAGE, S.; MYLNE, J.S.; GENDALL, A.R.; DEAN, C. Multiple roles of *Arabidopsis* VRN1 in vernalization and flowering time control. **Science**, Washington, v.297, n. 5579, p.243-246, July 2002.

LIN, S-I.; WANG, J-G.; POON, A-Y.; SU, C-L.; WANG, S-S.; CHIOU, T-J. Differential regulation of *FLOWERING LOCUS C* expression by vernalization in *Cabbage* and *Arabidopsis*. **Plant Physiology**, Taiwan, v.137, n.3, p.1037-1048, Mar. 2005.

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

MICHAELS, S.D.; AMASINO, R.M. Memories of winter: vernalization and the competence to flower. **Plant Cell, Environment**, v.23, n.11, p.1145-1153, Oct. 2000.

MICHAELS, S.D.; AMASINO, R.M. Loss of FLOWERING LOCUS C activity eliminates the late-flowering phenotype of *FRIGIDA* and autonomous pathway mutations but not responsiveness to vernalization. **The Plant Cell**, Madison, v. 13, n.4, p.935-941, Apr. 2001.

ONOUCHI, H.; IGENO, M.I.; PERILLEUX, C.; GRAVES, K.; COUPLAND, G. Mutagenesis of plants overexpressing CONSTANS demonstrates novel interactions among *Arabidopsis* flowering-time genes. **Plant Cell**, v. 12, 885-900. 2000.

PARENICOVÁ, L.; FOLTER, S.; KIEFFER, M.; HORNER, D.S.; FAVALLI, C.; BUSSCHER, J.; COOK, H.E.; INGRAM, R.M.; KATER, M.M.; DAVIES, B.; ANGENENT, G.C.; COLOMBO, L. Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: New openings to the MADS world. **The Plant Cell**, Madison, v. 15, n.7, p.1538-1551, July 2003.

PELAZ, S.; DITTA, G.S.; BAUMANN, E.; WISMAN, E.; YANOFSKY, M.F. B e C floral organ identity functions require SEPALLATA MADS-box genes. **Nature**, v.405:200-223, 2000.

PELAZ, S.; GUSTAFSON-BROWN, C.; KOHALMI, S.E.; CROSBY, W.L.; YANOFSKY, M.F. APETALA1 and SEPALLATA3 interact to promote flower development. **Plant Journal**, v. 26, n.4, p.385–394, May 2001.

PINYOPICH, A.; DITTA, D.S.; SAVIDGE, B.; LILJERGREN, S.J.; BAUMANN, E.; WISMAN, E.; YANOFSKY, M.F. Assessing the redundancy of MADS-box genes during carpel and ovule development. **Nature**, Washington, v.424, p.85–88, July 2003.

PURUGGANAN, M.D.; ROUNSLEY, S.D.; SCHMIDT, R.J.; YANOFSKY, M. Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. **Genetics**, 140:345-356 (1995).

RATCLIFFE, O.J.; NADZAN, G.C.; REUBER, T.L.; RIECHMANN, J.L. Regulation of flowering in *Arabidopsis* by an FLC Homologue. **Plant Physiology**, Hayward, California, v.126, 122-132 p. may 2001.

RATCLIFFE, O.J.; KUMIMOTO, R.W.; WONG, B.J.; REICHMANN, J.L. Analysis of the Arabidopsis MADS AFFECTING FLOWERING gene family: MAF2 prevents vernalization by short periods of cold. **Plant Cell**, Hayward, v. 15, n.5, p.1159–1169, 2003.

REEVES, P.A., Y, H.E., SCHIMITZ R.J., AMASINO, R.M., PANELLA L.M., RICHARDS C.M. Evolutionary conservation of the FLC-mediated vernalization response: evidence from the sugarbeet (*Beta vulgaris*). **Genetics**, v.176, n.1, p.295-307, May 2007.

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

RENA, A.B.; BARROS, R.S. Aspectos críticos no estudo da floração do café. In: ZAMBOLIM, L (Ed.). **Efeitos da irrigação sobre a qualidade e produtividade do café**. Viçosa, MG: UFV, 2004. p.149-172.

RIECHMANN, J.L.; KRIZEK, B.A.; MEYEROWITZ, E.M. Dimerization specificity of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. **Proceedings of the National Academy of Sciences USA**, 93:4793-4798, 1996.

RIECHMANN, J.L.; MEYEROWITZ, E.M. MADS domain proteins in plant development. **Biological Chemistry**, v.378:1079-1101, 1997.

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

SCHLÄPPI, M.; M. PATEL. Biennialism and vernalization promoted flowering in *Hyoscyamus niger*: a comparison with *Arabidopsis*. **Flowering Newsletter**. v.31, p.25–32, May 2001.

SCHWARS-SOMMER, Z.; HUIJSER, P.; NACKEN, W.; SAEDLER, H.; SOMMER, H. 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.; BURN, J.E.; PEREZ, P.P.; METZGER, J.; EDWARDS, J.A.; PEACOCK, W.J.; DENNIS, E.S. The FLM MADS box gene: a repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. **Plant Cell**, v.11, n.3, p.445–458, Mar. 1999.

SHELDON, F.; FINNEGAN, E.J.; DENNIS, E.S.; PEACOCK W.J. Quantitative effects of vernalization on *FLC* and *SOC1* expression. **Plant Journal**, Canberra, v. 45, p.871-883, 2006.

SHELDON, C.C.; HILLS, M.J.; LISTER, C.; DEAN, C.; DENNIS, E.S.; PEACOCK, W.J. Reverting of *FLOWERING LOCUS C* expression after epigenetic repression by vernalization. **Proceedings of the National Academy of Sciences USA**, Canberra, v. 105, n.6, p.2214-2219, Feb. 2007.

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

SHINDO, C.; ARANZANA, M.J.; LISTER, C.; BAXTER, C.; NICHOLLS, C.; NORDBORG, M.; DEAN, C. Role of *FRIGIDA* and *FLOWERING LOCUS C* in determining variation in flowering time of *Arabidopsis*. **Plant Physiology**, Rockville, v.138, p.1163- 1173, Mar. 2005.

SHORE, P.; SHARROCKS, A.D. The MADS-box family of transcription factors. **European Journal of Biochemistry**, v.229, n.1, p.1-13, Apr. 1995.

SOARES, A.R.; MANTOVANI, E.C.; RENA, A.B.; SOARES, A.A. Irrigação e fisiologia da floração em cafeeiros adultos na região da zona da mata de Minas Gerais. **Acta Scientiarum**, Maringá, v. 27, n. 1, p. 117-125, 2005.

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

TADEGE, M.; SHELDON, C.C.; HELLIWELL, C.A.; STOUTJESDIJK, P.; DENNIS, E.S.; PEACOCK, W.J. Control of flowering time by FLC orthologues in *Brassica napus*. **Plant Journal**, v.28, n.5, p.545-553, Dec. 2001.

TADEGE, M.; SHELDON, C. C.; HELLIWELL, C. A.; UPADHYAYA, N. M.; DENNIS, E. S.; PEACOCK, W.J. Reciprocal control of flowering time by OsSOC1 in transgenic *Arabidopsis* and by FLC in transgenic rice. **Plant Biotechnology Journal**, v. 1, n.5, p.361–369, Sept. 2003.

TAIZ, L.; ZEIGER, E. O controle do florescimento. In: \_\_\_\_\_. **Fisiologia Vegetal**. 3. ed. Porto Alegre: Artmed, 2006. p.581-611.

THEISSEN, G.; BECKER, A.; DI ROSA, A.; KANNO, A.; KIM, J.T.; MÜNSTER, T.; WINTER, K.-U.; SAEDLER, H. A short history of MADS-box genes in plants. **Plant Molecular Biology**, v. 42, n.1, p.115–149, Jan. 2000.

WEIGEL, D.; MEYEROWITZ, E.M. The ABCs of floral homeotic genes. **Cell**, v. 78, n.2, p.203-209, July 1994.

VIEIRA L.G.E.; ANDRADE A.C.; COLOMBO C.A.; PEREIRA, L.F.P.; SANTOS, S.N.; MORAES, A.H.A.; METHA, A.; OLIVEIRA, A.C.; LABATE, C.A.; MARINO, C.L. 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.; LI, Z-M.; MEI, L.; YAO, J-L.; HU, C-G. *PtFLC* homolog from trifoliolate orange (*Poncirus trifoliata*) is regulated by alternative splicing and experiences seasonal fluctuation in expression level. **Planta Journal**, Jan. 2009.

## **CAPÍTULO 2**

### **IDENTIFICAÇÃO *IN SILICO* DO GENE *FLOWERING LOCUS C* NO BANCO CAFEST**

## RESUMO

LAZZARI, Fabiane. Identificação *in silico* do gene *Flowering Locus C* no banco CAFEST. In: \_\_\_\_\_. **Identificação e expressão do gene *Flowering Locus C* (FLC) em café arábica**. 2009. Cap.2, p. 27-47. Dissertação (Mestrado em Agronomia/ Biotecnologia Vegetal) – Universidade Federal de Lavras.\*

*FLOWERING LOCUS C* (FLC) é um gene MADS-box que codifica um fator de transcrição com função de repressão da transição floral e cuja família gênica apresenta um motivo conservado. Esse trabalho teve como objetivo identificar o gene *FLC* presente na Base de Dados do Transcriptoma Café (CAFEST). Por meio da interface *Gene Project*, foi possível a procura por *reads*, montagem de *clusters* e análise dos prováveis genes *FLC*. Mais de mil *reads* encontrados foram agrupados, formando 39 EST-*contigs* que foram anotados e corrigidos quanto aos problemas de montagens errôneas e bases incorretas. Sequências MADS MIKC de *Arabidopsis* e as prováveis sequências *FLC* de café foram utilizadas para análises filogenéticas. O único EST-*contig* encontrado no CAFEST pertencente à família FLC foi o C15, que mostrou ser homólogo aos genes *FLC* (*AGL25*) e *MAF1-5*, de *Arabidopsis*. Além disso, pela análise filogenética entre *FLC* de outras espécies com o *contig* C15 foi possível agrupá-lo em um mesmo grupo monofilético que o gene *BvFLC* de *B. vulgaris*, e ambos afastados daqueles pertencentes à família Brassicaceae, indicando que possivelmente os genes fora desta família podem ter evoluído de um gene ancestral diferente. Através do perfil de expressão observado no *Northern* eletrônico, *FLC/C15* está expresso em ramos infectados com *Xylella sp* e em tecidos de plantas sob estresse hídrico no campo.

---

\* Orientador: Antonio Chalfun Júnior, PhD - UFPA

## ABSTRACT

LAZZARI, Fabiane. *In silico* identification of the *Flowering Locus C* gene in the bank CAFEST. In : \_\_\_\_\_. **Identification and expression of the gene *Flowering Locus C (FLC)* in arabica coffee.** 2009. Cap.2, p 27-47. Dissertation (Master in Agronomy/ Plant Biotechnology) - Federal University of Lavras.\*

*Flowering Locus C (FLC)* is a MADS-box gene that encodes a transcription factor that represses floral transition and belongs to a gene family having a conserved motif. This work aimed at to identify the *FLC* gene in the Coffee Transcriptome Database (CAFEST). By using the Gene Project interface it was possible to search for reads, clustering and analysis of genes *FLC* like genes. Over a thousand reads were found and clustered, generating the 39 EST-contigs, which were annotated and corrected for erroneous assembling. *Arabidopsis* MADS MIKC sequences and the coffee *FLC* putative one were used for the phylogenetic analysis. The only EST-contig found belonging to the FLC family was the C15 showing to be homologue to *FLC* gene (*AGL25*) and *MAF1-5* in *Arabidopsis*. In addition, phylogenetic analysis of the *FLC* with other species showed C15 in the same monophyletic group that the *BvFLC* gene of *B. vulgaris*, and both far from those belonging to the Brassicaceae family, indicating that, most likely, genes outside of the family may have evolved from a different ancestral gene. Based on the expression profile observed in the electronic Northern results, *FLC/C15* is expressed in infected areas with *Xylella sp* and plant tissues under drought stress in the field.

---

\* Adviser: Antonio Chalfun Júnior, PhD - UFLA

# 1 INTRODUÇÃO

Estudos realizados em *Arabidopsis* demonstram que os genes *MADS-box* não somente são responsáveis por determinar funções homeóticas florais, tendo também funções em outros aspectos do crescimento da planta, 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 (Theissen et al., 2000; Parenicová et al. 2003; Rijpkema et al. 2007).

Todos os genes *MADS* bem caracterizados contribuintes para o desenvolvimento vegetal são do tipo II (Alvarez- Buylla et al., 2000). Essa classe de proteínas apresenta uma organização estrutural conservada, denominada domínio estrutural tipo-MIKC, que inclui quatro domínios funcionais diferentes, distribuídos no sentido N a C terminal da proteína: o domínio *MADS-box* (M), a região *Intervening* (I), *keratin-like* (K), e C-terminal (C). (Hayes et al., 1988; Purugganan et al., 1995; Riechmann et al., 1996; Shore & Sharrocks, 1995).

O domínio *MADS* é a região mais altamente conservada de proteínas, e estas, como muitos fatores de transcrição eucarióticos, tem uma organização estrutural modular que ativam os processos de transcrição ligando-se a elementos de reconhecimento chamados *CArG boxes* (CC (A/T)<sub>6</sub> GG), encontrados em promotores de genes alvo.

*Flowering Locus C (FLC)* é um gene *MADS-box* mediador do tempo de florescimento, dependente de fatores ambientais como o frio, que marca o desenvolvimento da fase vegetativa para reprodutiva em *Arabidopsis* (Michaels & Amasino 1999).

Após o sequenciamento praticamente completo do genoma da *Arabidopsis* (*Arabidopsis* Genome Initiative, 2000), foi permitida uma descrição

compreensiva da complexidade e diversidade da família de genes MADS em espécies de plantas, que vêm contribuindo amplamente em estudos relacionados ao florescimento.

Genes *FLC-like* não haviam sido identificados em classes taxonômicas diferentes de *Arabidopsis*, *Brassica* e *Raphanus*, três gêneros dentro de Brassicaceae. Por esse motivo, inicialmente acreditava-se que genes *FLC-like* eram restritos dessa família (Schläppi & Patel 2001; Tadege et al. 2003; Hecht et al. 2005, Becker and Theissen, 2003). Contudo, em seu trabalho, Reeves et al. (2007) identificaram um *FLC* homólogo *BvFLC* em beterraba (*Beta vulgaris ssp.*). *BvFLC* é reprimido pelo frio prolongado e atrasa o florescimento quando expresso em plantas transgênicas de *Arabidopsis*.

Recentemente, o gene *PtFLC* foi isolado de mutantes de florescimento precoce de *Poncirus trifoliata*, através do método RACE combinado com uma biblioteca de cDNA. A análise de PCR em Tempo Real demonstrou a superexpressão de *PtFLC* durante o inverno, seguida por uma diminuição na primavera e verão (Zhang et al., 2009).

Com o sequenciamento do transcriptoma do cafeeiro (Vieira et al., 2006) e a criação do banco CAFEST de ESTs (*Expressed Sequence Tags*) tornou-se possível a realização de buscas por sequências formadoras dos prováveis genes relacionados às características de interesse.

O objetivo deste estudo foi encontrar prováveis proteínas FLC de *Coffea arabica* no CAFEST, com o objetivo de identificar *in silico* o(s) gene(s), seus domínios conservados, as bibliotecas em que se expressa e a relação dele com *FLC* de outras espécies.

## 2 MATERIAIS E MÉTODOS

### 2.1 Identificação de sequências *FLC*: análise *in silico*

Para a análise *in silico* no banco de dados, foram seguidos os procedimentos descritos por Oliveira (2007). As sequências de *FLC* 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* (disponível em [www.lge.ibi.unicamp.br/cafe](http://www.lge.ibi.unicamp.br/cafe)), foi possível a procura por *reads*, montagem de *clusters* e análise dos prováveis genes *FLC*.

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

Um segundo tipo de busca foi realizado por meio da ferramenta BLAST (*Basic Local Alignment Search*), utilizando-se as sequências *FLC* disponíveis no GenBank. Assim, por meio da ferramenta tBLASTN (Altschul et al., 1997), foram comparados os aminoácidos da fita *query* com os nucleotídeos traduzidos do banco de dados, e selecionados todos os *reads* que apresentaram alinhamento significativo ( $e\text{-value} > 10^{-5}$ ).

Posteriormente os reads encontrados foram agrupados (*clustering*), formando os chamados EST-*contigs*. Todos os EST-*contigs* formados foram utilizados para uma nova busca no banco CAFEST visando-se encontrar novas sequências prováveis de genes *FLC* bem como corrigir clusters incompletos, processo também chamado de saturação. Esse processo foi repetido até que não fosse 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 através do algoritmo BLASTp (Altschul et al., 1997).

## 2.2 Filogenia

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* pôde ser medida pelo teste probabilístico de bootstraps, originado de 100 replicatas (Sitnikova et al., 1995).

## 2.3 Identificação de Motivos Comuns de Agrupamento

Para descobrir motivos de agrupamento entre as sequências *FLC* selecionadas no CAFEST, utilizou-se 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 motivos 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 apareciam expressos. As bibliotecas foram as seguintes: AR1, folhas tratadas com ácido araquidônico; CA1/IC1, calos não embriogênicos; CS1, células em suspensão tratadas com NaCl; FB1/FB2/FB3/FB4, botões florais em diferentes estágios do florescimento; FR1/FR2, botões florais e frutos de diferentes estágios; FR4, frutos (*C. racemosa*); LV4/LV5, folhas jovens de ramos ortotrópicos; LV8, folhas maduras de ramos plagiotrópicos; RT8, células em suspensão com estresse ao alumínio; RX1, talos infectados com *Xylella spp*; SH2, pool de tecidos de plantas sob estresse hídrico no campo; SI3, sementes no início da germinação. As bibliotecas utilizadas foram descritas por Vieira e colaboradores (2006) e Lin e colaboradores (2005).

Esse procedimento exigiu que os dados de expressão dos *reads* nas bibliotecas fossem previamente 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, onde expressão zero ou negativa foi representada por coloração mais clara sendo aumentada gradativamente até atingir o preto, que representava o grau máximo de expressão positiva.

### 3 RESULTADOS E DISCUSSÃO

A primeira clusterização (agrupamento) realizada com os mais de 500 *reads* encontrados no banco de dados CAFEST formou 26 EST-*contigs*. A saturação (nova busca e agrupamento com o objetivo de corrigir *clusters* incompletos) foi muito importante, houve um aumento significativo com total de 1157 *reads*, sendo que após a última clusterização foram formados 39 EST-*contigs*.

Após a análise das sequências protéicas, da parte funcional de cada uma delas e seus domínios conservados, foram realizados um alinhamento e a comparação com as sequências disponíveis no NCBI que possibilitaram identificar qual, ou quais, EST-*contigs* encontrados poderiam ser verdadeiros.

Algumas prováveis proteínas FLC encontradas no CAFEST que apresentavam somente o domínio conservado *K-box*, ou que não apresentavam domínio conservado, provavelmente estavam incompletas, pois não foi possível a identificação de seus ESTs complementares. Esse fato pode ser explicado por erros na metodologia aplicada aos processos de sequenciamento e isolamento dos genes expressos, ou ineficiência dos métodos de seleção de *reads*. Esses *contigs* incompletos foram excluídos da análise e omitidos nos resultados. Dos 39 EST-*contigs*, apenas 25 foram mantidos na análise.

A análise dos *motifs* de agrupamento entre os prováveis EST-*contigs* FLC encontrados no CAFEST pode ser observada na figura 1, onde o domínio

MADS muito conservado (*motif 1*) aparece em todos os EST-*contigs*, sendo seguido pelos demais *motifs* característicos da família gênica.

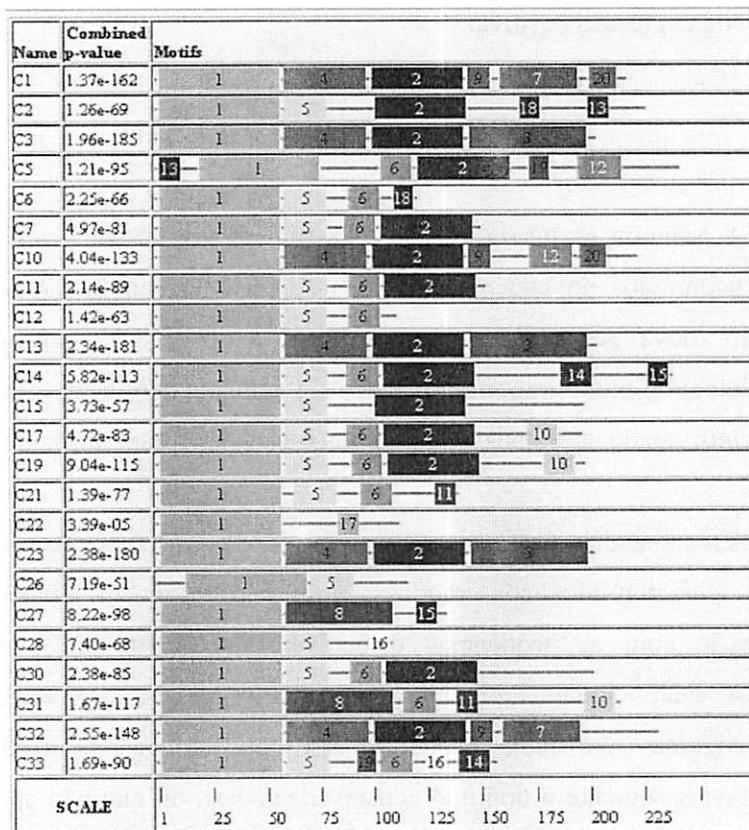


FIGURA 1. Representação gráfica mostrando os motivos de agrupamento entre os prováveis EST-*contigs FLC* encontrados no CAFEST, utilizando-se o programa MEME (<http://meme.sdsc.edu/meme/meme.html>).

Análises filogenéticas entre genes *FLC* têm sido relatadas em vários estudos (Becker e Theissen, 2003; Reeves et al., 2007). As análises filogenéticas (FIGURA 2) mostraram que o único EST-*contig* encontrado no CAFEST pertencente à subfamília *FLC* foi o C15 (621 nucleotídeos), que demonstrou ser homólogo aos genes *FLC (AGL25)* e *MAF1-5*, de *Arabidopsis*.

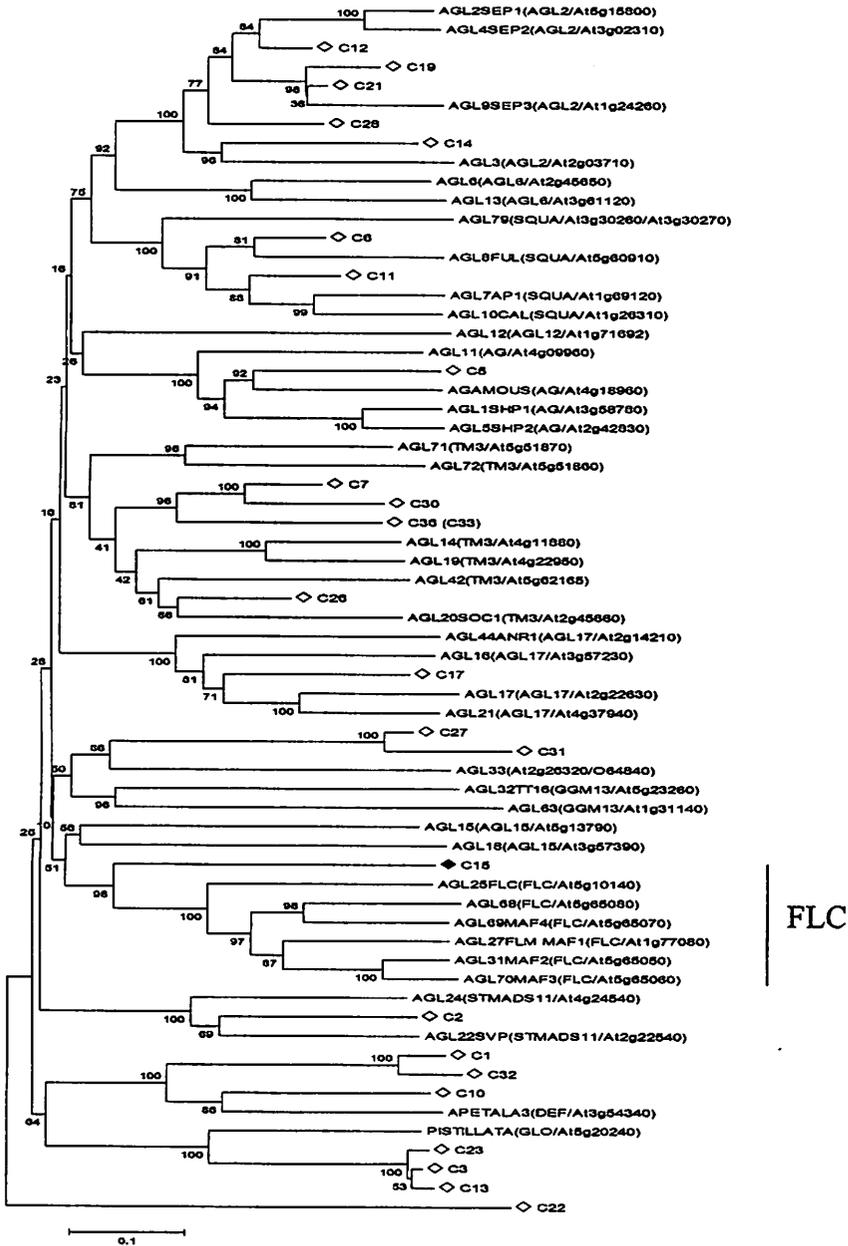


FIGURA 2. Árvore filogenética relacionando as 24 prováveis sequências FLC de *C. arabica* encontradas no CAFEST, e as proteínas de genes MADS tipo MIKC de *Arabidopsis thaliana* publicados e depositados no NCBI (<http://www.ncbi.nlm.nih.gov>).

Os motivos de agrupamento entre o C15 (*contig* C15) encontrado no CAFEST e as sequências *FLC* de *Arabidopsis* estão representados na figura 3.

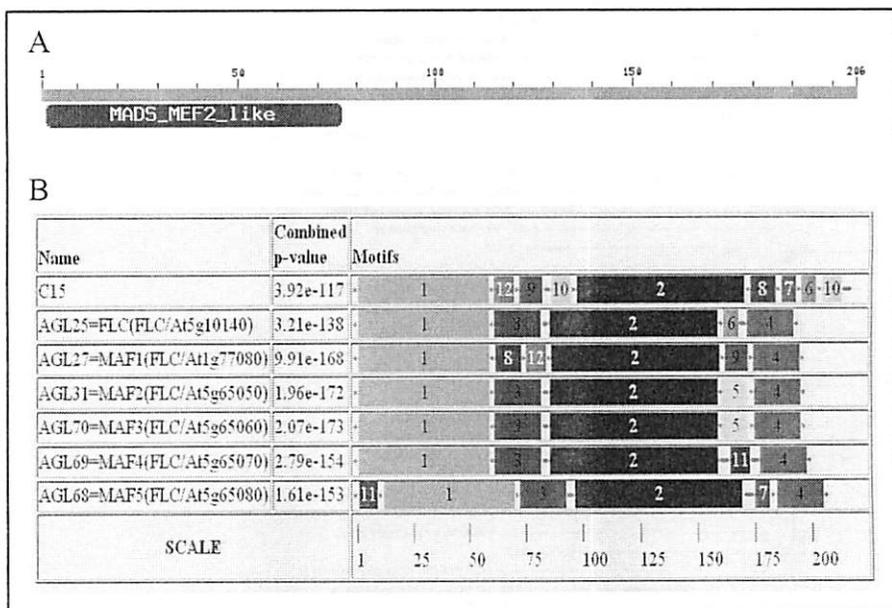


FIGURA 3. (A). Domínio MADS conservado da proteína *FLC Coffea arabica*. (B). Representação gráfica mostrando os motivos de agrupamento entre o C15 encontrado no CAFEST e as sequências *FLC* de *Arabidopsis*, utilizando-se o programa MEME (<http://meme.sdsc.edu/meme/meme.html>).

A relação filogenética entre proteínas *FLC* de outras espécies e a proteína *FLC* (C15) de *Coffea arabica* encontrada no CAFEST (FIGURA 4) demonstra os grupos de genes *FLC Brassica* e *FLC Arabidopsis* juntos e bem separados dos genes *MAF*, indicando que todas as cópias *FLC Brassica* são homólogas de *FLC Arabidopsis*. Esses três grupos de genes também se encontram bem separados do grupo que inclui genes que não pertencem a família Brassicaceae (*BvFLC* de *B. vulgaris* e o gene *FLC* (C15) de *C. arabica*).



Foi demonstrado pela análise filogenética a formação de três grupos monofiléticos, o de FLC de *Brassica*, o das proteínas MAF de *Arabidopsis* e o das proteínas FLC que não fazem parte da família Brassicaceae. O primeiro grupo monofilético inclui todas as proteínas FLC *Brassica* e a proteína FLC de *Arabidopsis* com alto valor *bootstrap* (100). Proteínas FLC *Brassica* dividem-se dentro de quatro grupos. O primeiro grupo de *Brassica* forma dois subgrupos, um contém BnFLC4 e BoFLC4 e o outro, BnFLC2 e BrFLC2, indicando que cópias FLC2 e FLC4 são muito similares entre si.

A análise *neighbour-joining* (FIGURA 4) também demonstra que os quatro grupos contendo as proteínas BoFLC1, BoFLC3, BoFLC4 e BoFLC5 foram derivados de um gene ancestral comum, o que é suportado pelo *bootstrap* relativamente alto (89), mas com fraca proximidade interna no grupo. Três das cinco proteínas BnFLC identificados (BnFLC1, BnFLC3, BnFLC5) são homólogas para cópias FLC *B.rapa* (BrFLC1, BrFLC3, BrFLC5).

O segundo grupo monofilético inclui as proteínas MAF1-5 de *Arabidopsis*.

Análises realizadas sobre a sequência genômica de *Arabidopsis* revelaram que, além de FLC, existem cinco genes MADS-*box* que codificam proteínas que são altamente relacionadas a ele e são homólogas (53-87% de identidade), os MADS AFFECTING FLOWERING 1-5 (MAF1-5) (Bodt et al., 2003; Ratcliffe et al., 2003). O primeiro gene analisado, MAF1 (também chamado FLOWERING LOCUS M – FLM), como FLC, é um importante repressor floral, apesar de sua expressão não ser tão fortemente influenciada pela vernalização. Os outros quatro loci foram nomeados MAF2 a MAF5. A superexpressão de MAF3 e MAF4 produz alterações no tempo de florescimento, o que sugere que esses também atuam como repressores florais e podem contribuir para necessidade da vernalização (Ratcliffe et al., 2001, 2003).

O terceiro grupo monofilético formado inclui a proteína BvFLC e a proteína FLC (C15) de *C. arabica*, com alto valor *bootstrap* (100). Isso indica que os genes fora da família Brassicaceae podem ter evoluído de um gene ancestral diferente. Muitos estudos precisam ser feitos para confirmar a relação interespecífica do gene *FLC*, bem como a conservação de sua sequência e função entre as espécies.

Genes *FLC-like* não haviam sido identificados em classes taxonômicas diferentes de *Arabidopsis*, *Brassica* e *Raphanus*, três gêneros dentro de Brassicaceae. Por esse motivo, inicialmente acreditava-se que genes *FLC-like* eram restritos dessa família (Schläppi & Patel 2001; Tadege et al. 2003; Hecht et al. 2005, Becker and Theissen, 2003). Contudo, genes *FLC* vem sendo identificados, em trabalhos recentes, em espécies fora da família Brassicaceae. Reeves et al. (2007) identificou o gene *BvFLC* homólogo em beterraba (*Beta vulgaris ssp.*). *BvFLC* é reprimido pelo frio prolongado e atrasa o florescimento quando expresso em plantas transgênicas de *Arabidopsis*. Ainda, *PtFLC* homólogo foi isolado de citrus, *Pocirus trifoliata*, e sua expressão é regulada por *splicing* alternativo (Zhang et al., 2009).

Muitas outras plantas respondem à vernalização, sugerindo que genes da classe FLC são respostas evolucionárias relativamente recentes para o modo de como as plantas respondem ao término do inverno e o início de uma época mais favorável para o desenvolvimento e reprodução (Becker and Theissen, 2003).

Para a análise espacial de expressão, foram calculadas as frequências de *reads* formadores de cada EST-*contig* nas bibliotecas em que apareciam expressos (Vieira et al., 2006). Para isso, os dados de expressão dos *reads* nas bibliotecas foram normalizados, para dar uma ideia exata do grau de expressão dos prováveis genes em cada tratamento e local da planta. Contudo, o resultado no cálculo da normalização do *contig* C5 na biblioteca FB3 (botões florais em diferentes estágios do florescimento) apresentou um valor muito elevado,

divergente dos demais. Esse valor foi alterado e desconsiderado, já que estava mascarando os resultados da análise de *Northern* eletrônico (FIGURA 5).

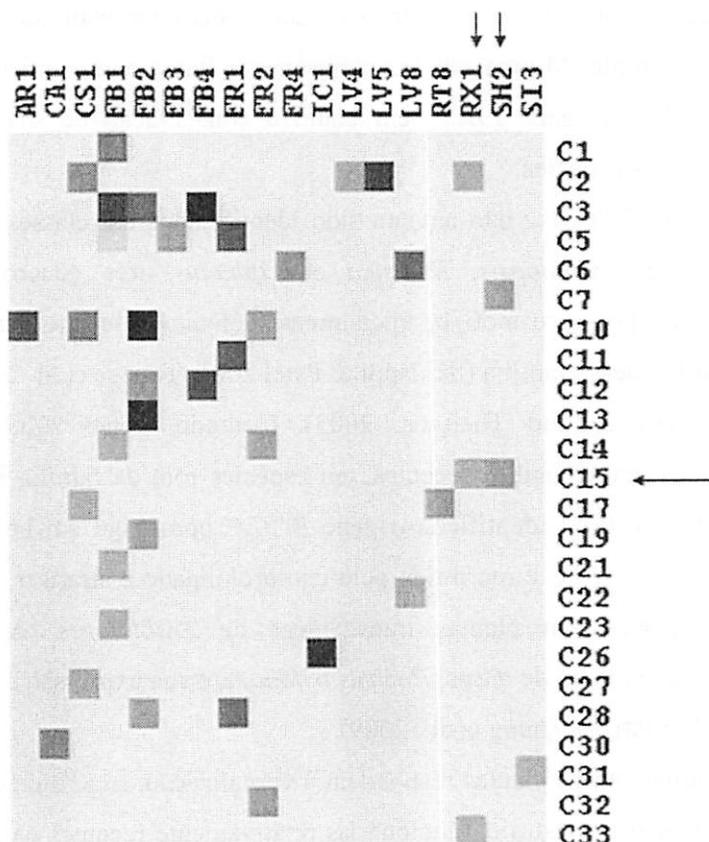


FIGURA 5. Northern Eletrônico representando os níveis de expressão dos EST-contigs nas diferentes bibliotecas por meio de uma escala cinza. Bibliotecas: AR1, folhas tratadas com ácido araquidônico; CA1/IC1, calos não embriogênicos; CS1, células em suspensão tratadas com NaCl; FB1/FB2/FB3/FB4, botões florais em diferentes estágios do florescimento; FR1/FR2, botões florais e frutos de diferentes estágios; FR4, frutos (*C. racemosa*); LV4/LV5, folhas jovens de ramos ortotrópicos; LV8, folhas maduras de ramos plagiotrópicos; RT8, células em suspensão com estresse ao alumínio; RX1, talos infectados com *Xylella spp*; SH2, pool de tecidos de plantas sob estresse hídrico no campo; SI3, sementes no início da germinação. As bibliotecas utilizadas foram descritas por Vieira et al., 2006 e Lin et al., 2005.

Os resultados de expressão foram apresentados em uma escala cinza, onde expressão zero ou negativa foi representada por coloração mais clara sendo aumentada gradativamente até atingir o preto, que representava o grau máximo de expressão positiva (FIGURA 5). O EST-*contig* C15 apresentou padrão de expressão nas bibliotecas de talos infectados com *Xylella spp.* (RX1) e de tecidos sob estresse hídrico (SH2), ambas de tecidos vegetativos submetidos a estresses. Esse fato é muito relevante, já que é uma indicativa de que a expressão de transcritos *FLC* pode ser detectável quando a planta sofre estresse, ou por ataque de um patógeno, ou devido a um déficit hídrico.

Os resultados observados na análise espacial pelo *Northern* eletrônico geram uma expectativa para a confirmação de expressão tecido específica, na qual não se espera observar perfil de expressão em tecidos de inflorescências, sementes ou raiz.

Devido ao fato de não aparecer expresso em bibliotecas de sementes ou de raízes e, sobretudo pela expressão aparentemente induzida por fatores externos relativos ao tratamento, é provável que esteja filogeneticamente mais próximo ao gene *FLC* de *Arabidopsis*, sendo este seu homólogo (FIGURA 2).

Ainda, não podem ser desconsiderados os possíveis erros no sequenciamento dessas bibliotecas, o que poderia estar mascarando os resultados de expressão dos *reads* nas demais bibliotecas. Contudo, são necessários estudos de confirmação da expressão tecido específica *in vivo*.

## 4 CONCLUSÕES

O elevado número de *reads* e a eficiência na clusterização contribuíram para uma abordagem eficiente *in silico* para identificação do gene *FLC* da família *MADS-box* em *Coffea arabica*. O gene *FLC* presente no banco de dados CAFEST apresentou alta homologia com o *FLC* de *Arabidopsis* e de *B. vulgaris*.

Esses resultados ampliam o interesse em estudos de confirmação da expressão do gene *FLC* para entender a regulação do florescimento em plantas de café.

## 5 REFERÊNCIAS BIBLIOGRÁFICAS

ALTSCHUL, S. F.; MADDEN, T. L.; SCHAFFER, A. A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. J. Gapped BLAST and PSI BLAST: A new generation of protein database search programs. **Nucleic Acids Research**, v.25, n.17, p.3389-3402, Sept. 1997.

ALVAREZ-BUYLLA, E.R.; PELAZ, S.; LILJEGREN, S.J.; GOLD, S.E.; BURGEFF, C.; DITTA, G.S.; DE POUPLANA, L.R.; MARTINEZ-CASTILLA, L.; YANOFSKY, M.F. An ancestral *MADS-box* gene duplication occurred before the divergence of plants and animals. **Proceedings of the National Academy of Sciences USA**, v.97, n.10, p.5328–5333, May 2000.

*Arabidopsis* Genome Initiative. Analysis of the genome sequence of the flowering plants *Arabidopsis thaliana*. **Nature**, v. 408:796-815, 2000.

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, 2., 1994, Menlo Park. **Proceeding...** Menlo Park: AAAI Press, 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**, v. 29, n.3, p.464-489, Dec. 2003.
- BODT, S.D.; RAES, J.; PEER, Y.V.D.; THEISSEN, G. And then there were many: MADS goes genomic. **Trends in Plant Science**, v. 8, n.10, p. 475–483, Oct. 2003.
- EISEN, M.B.; SPELLMAN, P.T.; BROWN, P.O.; BOTSTEIN, D. Cluster analysis and display of genome-wide expression patterns. **Proceedings of the National Academy of Sciences USA**, v. 95, p.14863-8, 1998.
- HAYES, T.E.; SENGUPTA, P.; COCHRAN, B.H. The human c-fos serum response factor and the yeast factor GRM/PRTF have related DNA-binding specificities. **Genes and Development**, v. 2, n. 12B, p.1713-1722, Dec. 1988.
- HECHT, V.; FOUCHER, F.; FERRA NDIZ, C.; MACKNIGHT, R.; NAVARRO, C.; MORIN, J.; VARDY, M.E.; ELLIS, N.; BELTRÁN, J.P.; RAINEAU, C.; WELLER, J.L. Conservation of Arabidopsis flowering genes in model legumes. **Plant Physiology**, v. 137: 1420–1434p. 2005.
- MICHAELS, S.D.; AMASINO, R.M. FLOWERING LOCUS C a novel MADS domain that acts as a Repressor of Flowering. **The Plant Cell**, Madison, v.11, p.949-956, 1999.
- OLIVEIRA, R. R. Identificação e análise da expressão *in silico* de genes MADS-box no banco de ESTs CAFEST (*Coffea arabica*, *C. Canephora* e *C. racemosa*). 2007. 30p. Monografia de graduação em Ciências Biológicas– Universidade Federal de Lavras, Lavras.
- PARENICOVÁ, L.; FOLTER, S.; KIEFFER, M.; HORNER, D.S.; FAVALLI, C.; BUSSCHER, J.; COOK, H.E.; INGRAM, R.M.; KATER, M.M.; DAVIES, B.; ANGENENT, G.C.; COLOMBO, L. Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: New openings to the MADS world. **The Plant Cell**, Madison, v. 15, n.7, p.1538-1551, July 2003.
- PURUGGANAN, M.D.; ROUNSLEY, S.D.; SCHMIDT, R.J.; YANOFSK, M. Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. **Genetics**, 140:345-356, 1995.

- RATCLIFFE, O.J.; NADZAN, G.C.; REUBER, T.L.; RIECHMANN, J.L. Regulation of flowering in *Arabidopsis* by an FLC Homologue. **Plant Physiology**, Hayward, California, v.126, 122-132 p. may 2001.
- RATCLIFFE, O.J.; KUMIMOTO, R.W.; WONG, B.J.; REICHMANN, J.L. Analysis of the Arabidopsis MADS AFFECTING FLOWERING gene family: MAF2 prevents vernalization by short periods of cold. **Plant Cell**, Hayward, v. 15, n.5, p.1159–1169, 2003.
- REEVES, P.A.; Y, H.E.; SCHIMITZ R.J.; AMASINO, R.M.; PANELLA L.M.; RICHARDS C.M. Evolutionary conservation of the FLC-mediated vernalization response: evidence from the sugarbeet (*Beta vulgaris*). **Genetics**, v.176, n.1, p.295-307, May 2007.
- RIECHMANN, J.L.; KRIZEK, B.A.; MEYEROWITZ, E.M. Dimerization specificity of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. **Proceedings of the National Academy of Sciences USA**, 93:4793-4798, 1996.
- RIJPKEMA, A.S.; GERATS T.; VANDENBUSSCHE, M. Evolutionary complexity of MADS complexes. **Current Opinion in Plant Biology**, Netherlands, v.10, n.1, p.32–38, May 2007.
- SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, v. 4, p.406–425, 1987.
- SCHLÄPPI, M.; M. PATEL. Biennialism and vernalization promoted flowering in *Hyoscyamus niger*: a comparison with *Arabidopsis*. **Flowering Newsletter**. v.31, p.25–32, 2001.
- SHORE, P.; SHARROCKS, A.D. The MADS-box family of transcription factors. **European Journal of Biochemistry**, v.229, n.1, p.1-13, Apr. 1995.
- SITNIKOVA, T.; RZHETSKY, A.; NEI, M. Interior-branch and bootstrap tests of phylogenetics trees. **Molecular Biology and Evolution**, v.12, p.319-33, 1995.
- TADEGE, M.; SHELDON, C. C.; HELLIWELL, C. A.; UPADHYAYA, N. M.; DENNIS, E. S.; PEACOCK, W.J. Reciprocal control of flowering time by OsSOC1 in transgenic *Arabidopsis* and by FLC in transgenic rice. **Plant Biotechnology Journal**, v. 1, n.5, p.361–369, Sept. 2003.

TAMURA, K.; DUDLEY, J.; NEI, M.; KUMAR, S. H. MEGA4: Molecular Evolutionary Genetics Analysis: software: version 4.0. **Molecular Biology and Evolution**, v. 24, n. 8, p. 1596-1599, May 2007.

THEISSEN, G.; BECKER, A.; DI ROSA, A.; KANNO, A.; KIM, J.T.; MÜNSTER, T.; WINTER, K.-U.; SAEDLER, H. A short history of MADS-box genes in plants. **Plant Molecular Biology**, v. 42, p.115–149, 2000.

THOMPSON, J.D.; HIGGINS, D.G.; GIBSON, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Research**, v. 22, p.4673–4680, 1994.

VIEIRA L.G.E.; ANDRADE A.C.; COLOMBO C.A.; PEREIRA, L.F.P.; SANTOS, S.N.; MORAES, A.H.A.; METHA, A.; OLIVEIRA, A.C.; LABATE, C.A.; MARINO, C.L. 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.; LI, Z-M.; MEI, L.; YAO, J-L.; HU, C-G. *PtFLC* homolog from trifoliolate orange ( *Poncirus trifoliata* ) is regulated by alternative splicing and experiences seasonal fluctuation in expression level. **Planta Journal**, Jan. 2009.

## **CAPÍTULO 3**

### **ISOLAMENTO E CLONAGEM DO GENE *FLOWERING LOCUS C* EM CAFÉ**

## RESUMO

LAZZARI, Fabiane. Isolamento e clonagem do gene *Flowering Locus C* em café. In: \_\_\_\_\_. **Identificação e expressão do gene *Flowering Locus C* (FLC) em café arábica**. 2009. Cap.3, p. 48-72. Dissertação (Mestrado em Agronomia/Biotecnologia Vegetal) – Universidade Federal de Lavras.\*

O locus *Flowering Locus C* (FLC) foi identificado em *A. thaliana* através de um fenótipo mutante nulo, demonstrando que a função de FLC é reprimir o florescimento. O produto do gene e função de FLC são conhecidos em *Arabidopsis*, *Brassica sp.*, *B. vulgaris* (beterraba) e *Poncirus trifoliata*. O objetivo deste trabalho foi isolar e clonar o gene FLC de *C. arabica*, para obtenção de sua sequência, e para confirmação da expressão tecido específica. Primers específicos foram desenhados para clonar o cDNA de *CaFLC* obtido a partir do RNA total extraído de tecidos de folha. Além disso, o RNA total de tecidos de raiz, fruto e flor de plantas de café foi utilizado para estudos de expressão. O sequenciamento foi realizado e a homologia foi confirmada comparando-se o resultado com a sequência obtida previamente *in silico*. Similaridade entre as sequências foi observada pelo alinhamento, demonstrando que a análise preliminar utilizando bioinformática é uma ferramenta muito poderosa em estudos moleculares. PCR quantitativo em Tempo Real foi realizado para confirmar expressão de *CaFLC* tecido específica. Avaliando-se os níveis de expressão nos diferentes tecidos, foi possível observar que a expressão quantitativa relativa do gene FLC se dá diferencialmente entre os tecidos, sendo que em folhas foi observado o mais elevado nível de expressão.

---

\* Orientador: Antonio Chalfun Júnior, PhD - UFLA

## ABSTRACT

LAZZARI, Fabiane. Isolation and cloning of the *Flowering Locus C* gene in coffee. In : \_\_\_\_\_. **Identification and expression of the gene *Flowering Locus C (FLC)* in arabica coffee.** 2009. Cap.3, p. 48-72. Dissertation (Master in Agronomy/ Plant Biotechnology) - Federal University of Lavras.\*

The *Flowering Locus C (FLC)* locus was identified in *Arabidopsis* by a null mutant phenotype, demonstrating that the function of *FLC* is to repress flowering. *FLC* gene product and function are known in *Arabidopsis*, *Brassica sp.*, *B. vulgaris* (sugar beet), and *Poncirus trifoliata*. The objective of this study was to isolate and clone the *C. arabica Flowering Locus C* gene in which *FLC* tissue specific expression might be confirmed. Specific primers were designed to clone cDNA *CaFLC* from total RNA obtained from leaf tissues. Moreover, tissues from root, flower and fruit of coffee plants were used to extract RNA to be used in expression analysis. Sequencing was performed and the homology could be confirmed by comparing the result with the *in silico* sequence previously obtained. Similarity between sequences could be observed showing that the preliminary analysis using bioinformatics is a very powerful tool in molecular studies. Real-time quantitative PCR was carried out to confirm *CaFLC* tissue specific expression. Based on the expression levels in different tissues, it was possible to observe that the relative quantitative expression of the *FLC* gene occurs differentially among tissues and in leaves it was observed the highest level of expression.

---

\* Adviser: Antonio Chalfun Júnior, PhD - UFLA

# 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.

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 (Koorneef 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 através da 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 envolve a modulação da expressão do gene *Flowering Locus C (FLC)*, que é um gene *MADS-box* com função de repressão da transição floral (Michaels & Amasino, 1999).

Genes *FLC-like* vêm sendo descritos em várias espécies. O produto do gene e função de *FLC* são conhecidos em *Arabidopsis*, *Brassica sp.*, *B. vulgaris* (beterraba) e *Poncirus trifoliata* (citrus) (Michaels e Amasino, 1999; Tadege et al. 2001; Lin et al., 2005; Reeves et al., 2007; Zhang et al., 2009).

O objetivo deste trabalho foi isolar e clonar o gene *FLC* de *C. arabica*, com o objetivo de obter sua sequência e compará-la com a sequência obtida *in*

*silico*. Ainda, verificar através da análise qRT-PCR em que tecidos da planta o gene está sendo expresso.

## 2 MATERIAIS E MÉTODOS

Amostras de tecidos (folha, fruto, flor e raiz) de três plantas diferentes da espécie *C. arabica* cultivar Rubi foram coletadas no campo experimental da UFLA, em Lavras/MG. As amostras de cDNA de folha de *C. arabica* cultivar Rubi e IAPAR, e de cDNA de raiz de *C. canephora* utilizadas na análise de qRT-PCR foram cedidas pela EMBRAPA-Cenargen, de Brasília/DF.

Todos os procedimentos, exceto qRT-PCR, foram realizados no Laboratório Central de Biologia Molecular da UFLA – LCBM. As análises de qRT-PCR foram realizadas no laboratório de Biotecnologia, na EMBRAPA Recursos Genéticos e Biotecnologia – Cenargen.

### 2.1 Extração e quantificação do RNA

Todos os materiais utilizados para extração (cadinhos e pistilos de porcelana, microtubos e ponteiras, água destilada) foram tratados com DEPC (diethylpyrocarbonate) para inativar RNAses. Os materiais foram submersos em DEPC 0,5% *overnight* e autoclavados por 30 minutos, já que o DEPC inativa outros componentes além de RNase.

Para extração do RNA, foi utilizado o método Borato Quente, proposto por Birtic & Kranner (2006) com modificações. Tetraborato decahidrato de sódio 0,2M, contendo 30mM de etileno glycol bis (2-aminoethyl ether)-N,N,N<sub>2</sub>,N<sub>2</sub>-ácido tetraacético, 1% (w/v) de sódio dodecyl sulfato e 1% (w/v) de sódio dexycolate, com pH ajustado a 9,0 com hidróxido de sódio 1M, foi usado como tampão de extração.

1g de tecido de cada amostra foi macerado em N<sub>2</sub> líquido, até a obtenção de um pó fino.

A cada microtubo foram adicionados 1mL de tampão de extração, suplementados com 6% (w/v) de Polyvinylpyrrolidone (PVP) e 4μL de DTT 1M, homogeneizados em vórtex e aquecidos a 80°C. Em seguida, a amostra macerada (100mg) foi transferida para a solução tampão aquecida. À solução foram adicionados 1mg de Proteinase K e 2μL de β-mercaptoetanol, que foram homogeneizados em vortex e mantidos a 42°C por 1,5 horas. Como primeiro passo para eliminar elementos indesejáveis, foram adicionados 80μL de Cloreto de Potássio 2M a cada amostra e incubadas no gelo durante 45 minutos. Após esse período, o material foi submetido à centrifugação por 20 minutos a 4°C com velocidade máxima (16.000g) e a fase aquosa foi transferida para um novo tubo. Os ácidos nucleicos foram ressuspensos pela adição de 300μL de Cloreto de Lítio 8M gelado. As amostras foram incubadas no gelo no mínimo 1 hora ou *overnight* a -20°C, em seguida foram centrifugadas por 20 minutos a 4°C e foi descartado o sobrenadante. O pellet foi lavado duas vezes com 750μL de Cloreto de Lítio 2M gelado e submetido a centrifugação por 10 minutos a 4°C e o sobrenadante foi descartado. O pellet contendo os ácidos nucleicos foi então ressuspensionado em 20μL de água destilada tratada com DEPC por vortex.

Depois de centrifugação por 10 minutos, a 18.000g, a integridade do RNA foi visualizada em gel de agarose 1,2%, corado com brometo de etídio, sob luz ultravioleta. As amostras foram quantificadas em espectrofotômetro (Nanodrop<sup>®</sup> Espectrophotometer ND-1000) a A<sub>260</sub> nm com o objetivo de determinar a quantidade de RNA total necessária para o tratamento com DNaseI.

As amostras de RNA total foram tratadas com DNase I (New England BioLabs), segundo especificações do fabricante, para eliminação de eventual DNA presente. Uma reação contendo 10μg de RNA total, 2u de enzima DNase

I, 10 $\mu$ L de tampão de reação 10X e água destilada DEPC, para completar um volume final de 100 $\mu$ L, foi incubada a 37°C, temperatura ótima para atividade da enzima, por 30 minutos. Em seguida, o RNA total foi purificado da reação pela adição de Acetato de Potássio 2M para uma concentração final de 10% do volume, incubado no gelo durante 15 minutos e submetido à centrifugação a 18.000g, a 4°C, por 10 minutos. Para precipitar e facilitar a separação do RNA dos carboidratos e proteínas, o sobrenadante (transferido para um novo tubo) foi misturado com etanol 100% e mantido a -80°C durante 2 horas. Após centrifugação a 18.000g, a 4°C, por 30 minutos, o sobrenadante foi descartado. O pellet foi lavado com 400 $\mu$ L de etanol 70%, centrifugado a 18.000g, a 4°C, por 10 minutos e então seco em câmara de fluxo a temperatura ambiente, durante, aproximadamente, 10 minutos, para evaporação total do álcool, o qual poderia interferir em futuras reações. O pellet foi finalmente ressuspendido em 20 $\mu$ L de água destilada DEPC e estocado a -80°C.

## 2.2 Desenho de primers

Os primers utilizados neste estudo foram desenhados com o programa Oligo Perfect<sup>TM</sup> Designer (Invitrogen) e estão descritos na Tabela 1. Primers específicos para o gene *FLC* foram desenhados a partir da sequência obtida *in silico* neste trabalho, através da busca no banco CAFEST. O gene *Ubiquitina (UBQ)* foi utilizado como controle endógeno constitutivo da expressão e primers específicos para esse gene foram desenhados a partir das sequências disponíveis no GenBank (<http://www.ncbi.nlm.nih.gov/>).

Os primers CAflcT (*forward e reverse*) foram desenhados para isolar a sequência cDNA *FLC* de *Coffea arabica*. Os primers qCAFLC e UBQ foram desenhados para a análise de qRT-PCR.

TABELA 1. Primers utilizados para clonagem e análise qRT-PCR.

Gene	Sequência do Primer
UBQ-F	5' CCAAGATCCAGGACAAGGAA 3'
UBQ-R	5' TGGATGTTGTAGTCGGCAAG 3'
qCAFLC-F	5' GCTGAGGGAAGAAAACAAGC 3'
qCAFLC-R	5' CGTGTGTAAGGTCCCAAAT 3'
CAflcT-F	5' <u>CACCGAAGGAATTAAGGATGG</u> 3'
CAflcT-R	5' TTAAAGTAGTTCAAGCATAGCCTGTC 3'

### 2.3 Obtenção da sequência cDNA

O fragmento cDNA total foi obtido a partir do RNA total extraído de folhas de *C. arabica*, através do kit ImProm-II™ Reverse Transcription System - A3800 da PROMEGA. 2µg de RNA total foram combinados com 0,5µg do primer Oligo-(dT)<sub>15</sub> em microtubos e o mix RNA template/oligo-dT foi desnaturado a 70°C por 5 minutos e imediatamente transferidos para o gelo por 5 minutos. O mix da reação de Transcrição Reversa (RT) foi montado em microtubo de 1,5µL em gelo, onde foram adicionados: água RNase Free (para um volume final de 15µL), 4,0µL tampão de reação 5x ImProm-II™, 3,0µL MgCl<sub>2</sub> (concentração final 3,0mM), 1,0µL Mix dNTP (concentração final 0.5mM cada dNTP), 0,5µL Recombinant RNasin® Ribonuclease Inhibitor (20u) e 1µL da enzima Transcriptase Reversa (RT) ImProm-II™. O Mix RT (15µL) foi combinado ao RNA/oligo-dT (5µL) com volume final da reação de 20µL. As condições da reação RT foram as seguintes: anelamento a 25°C por 5 minutos e extensão a 42°C por 1 hora.

O cDNA foi subsequentemente usado como *template* para amplificação PCR. Para isso foi necessária a inativação da enzima RT, através da incubação dos tubos a 70°C por 15 minutos.

O cDNA *FLC* foi amplificado usando os primers forward e reverse: CAflcT-F: 5'-CACCGAAGGAATTAAGGATGG-3'; e CAflcT-R: 5'-TTAAAGTAGTTCAAGCATAGCCTGTC-3' (TABELA 1). A reação de amplificação foi realizada em um volume final de 50 µL contendo 10 µL da reação RT, Tampão - 10X *thermophilic polymerase reaction buffer* (sem MgCl<sub>2</sub>), 25mM MgCl<sub>2</sub> (2mM concentração final), 10mM dNTP mix (0.2mM final), primer CAflcT-F (1µM final), primer CAflcT-R (1µM final), Taq DNA polymerase EasyPath comum (5.0u).

A reação foi submetida ao termociclador Eppendorf Mastercycler gradient, programado com temperatura de desnaturação inicial de 94°C, por 1 minuto, seguidos de 34 ciclos de amplificação compostos de 3 etapas: 94°C, por 1 minuto; 60°C, por 1 minuto; 72°C, por 1 minuto. Após os 34 ciclos, as amostras foram submetidas a uma etapa final de 5 minutos a 72°C.

As amostras foram armazenadas a -20°C.

#### 2.4 Eletroforese e Purificação do produto da amplificação (cDNA) por PCR

O produto da amplificação foi submetido a eletroforese em gel de agarose 1,2% (m/v) corado com 2µL de Brometo de Etídio (0,5µg/mL), sob corrente elétrica de 110V em Tampão TAE (0,001M EDTA pH 8,0; 0,04M TRIS pH 8,0; 0,02M ácido acético), por um período de 45 minutos, sendo visualizado sob luz ultravioleta e a imagem captada pelo fotodocumentador EDAS 290 (Kodak<sup>®</sup>).

O fragmento obtido foi recortado do gel com bisturi e transferido para um microtubo, onde foi submetido à eluição por meio do Kit MinElute (QUIAGEN<sup>®</sup>), de acordo com as informações do fabricante.

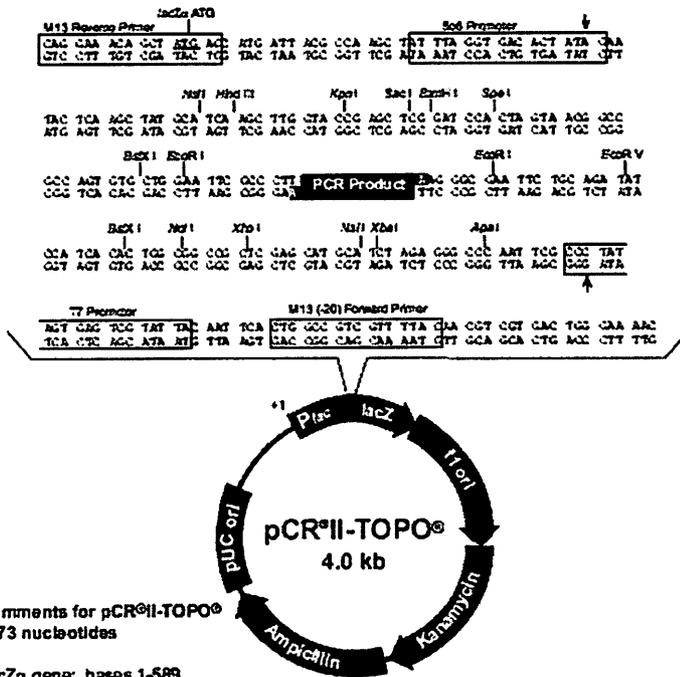
A fatia do gel foi pesada e foram acrescentados 3 volumes do tampão QG a cada volume do gel. O microtubo foi incubado, a 50°C, por 10 minutos ou até a diluição completa do gel, agitando-se no vortex a cada 3 minutos, durante a incubação.

Foi adicionado 1 volume de isopropanol à amostra, invertendo-se várias vezes o microtubo. A amostra foi aplicada em uma coluna MinElute inserida dentro de um microtubo de 2mL e centrifugada, a 16.000g por 1 minuto, sendo o filtrado posteriormente descartado. Foram adicionados 500µL do tampão QG e em seguida centrifugado por 1 minuto. Para lavar o cDNA, foram adicionados 750µL do tampão PE à coluna que foi centrifugada por 1 minuto. O filtrado foi descartado e a coluna foi centrifugada novamente, a 16.000g, por 1 minuto.

A coluna foi colocada em um tubo limpo de 1,5mL e, para eluir o cDNA, foram adicionados 10µL do tampão EB (10mM Tris-HCl, pH 8,5) no centro da membrana. A coluna foi deixada em repouso por 1 minuto e então centrifugada, por 1 minuto. O produto da eluição foi armazenado em a -20°C.

## 2.5 Ligação com vetor

O produto da eluição foi submetido à clonagem no vetor pCR<sup>®</sup>II usando TOPO TA Cloning<sup>®</sup> Kit Dual Promoter (Invitrogen) (Figura 1). A reação de ligação do fragmento ao vetor foi realizada a temperatura ambiente (22-23°C), por 20 minutos, misturando-se 2µL do produto da PCR purificado, 2µL de água deionizada, 1µL de Vetor TOPO TA Cloning e 1µL de solução salina.



Comments for pCR<sup>II</sup>-TOPO<sup>®</sup>  
 3973 nucleotides

- LacZ $\alpha$  gene: bases 1-589
- M13 Reverse priming site: bases 205-221
- Sp6 promoter: bases 239-258
- Multiple Cloning Site: bases 289-383
- T7 promoter: bases 408-425
- M13 (-20) Forward priming site: bases 433-448
- fl ori: bases 550-1027
- Kanamycin resistance ORF: bases 1381-2155
- Ampicillin resistance ORF: bases 2173-3033
- pUC origin: bases 3178-3851

FIGURA 1. Detalhes do vetor pCR<sup>II</sup> TOPO TA Cloning (Invitrogen), tais como a origem de replicação, os marcadores selecionáveis de resistência à Canamicina e à Ampicilina, o operon LacZ interrompido por sítios de enzimas de restrição. (Instruction Manual Invitrogen TA Cloning<sup>®</sup> Kit Dual Promoter, disponível em <http://tools.invitrogen.com/content/sfs/manuals>)

## 2.6 Transformação e clonagem bacteriana

O produto da reação de ligação foi utilizado para transformar células da bactéria *Escherichia coli* DH5 $\alpha$ <sup>TM</sup>-T1<sup>R</sup> quimicamente competente (Invitrogen),

em que 2 $\mu$ L do produto da ligação foram misturados com 50 $\mu$ L de células competentes DH5 $\alpha$ <sup>TM</sup>-T1<sup>R</sup> e colocados em gelo, por 30 minutos. Logo em seguida, foram submetidos a choque térmico por 30 segundos a 42°C e transferidos para o gelo. Foram adicionados 250 $\mu$ L de meio SOC (Invitrogen) e incubados em TermoMixer Eppendorf, por uma hora, a 37°C sob agitação de 300 rpm. Logo após a incubação, o plaqueamento foi feito utilizando-se 250 $\mu$ L da solução de transformação em placa de Petri, contendo 20mL de meio de cultura Luria-Bertani (LB) + ágar, 50 $\mu$ g/mL ampicilina e 40mg/mL de X-Gal (5-bromo-4-cloro-3-indolil-D-galactosídeo).

Após o plaqueamento, as culturas foram mantidas a 37°C, durante 16 horas, para permitir o crescimento das colônias bacterianas. Após esse período, 10 colônias brancas transformadas foram selecionadas aleatoriamente, transferidas a outra placa para subcultivo com o mesmo meio, e submetidas a PCR de colônia, para comprovar a presença do fragmento de interesse. Para isso, com auxílio de ponteiros esterilizados, as colônias foram coletadas para outra placa, identificadas, e a ponteira de cada uma foi depositada no tubo contendo o mix da PCR. O mix da reação de amplificação foi realizada em um volume final de 20  $\mu$ L contendo Tampão da enzima 10X, 25mM MgCl<sub>2</sub> (2mM concentração final), 10mM dNTP mix (0.2mM final), primer CAflcT-F (1 $\mu$ M final), primer CAflcT-R (1 $\mu$ M final), Taq DNA polymerase EasyPath comum (5.0u) e água destilada. A reação foi submetida ao termociclador Eppendorf Mastercycler gradient, programado com temperatura de desnaturação inicial de 94°C, por 1 minuto, seguidos de 34 ciclos de amplificação compostos de 3 etapas: 94°C, por 1 minuto; 60°C, por 1 minuto; 72°C, por 1 minuto. Após os 34 ciclos, as amostras foram submetidas a uma etapa final de 5 minutos a 72°C. O produto da PCR foi analisado através de eletroforese em gel de agarose 1,2% e as colônias que apresentaram o fragmento de interesse foram selecionadas para o processo de extração de DNA plasmidial.

A placa contendo as 10 colônias selecionadas foi incubada a 37°C por 16 horas, para o crescimento das colônias. Das colônias brancas que apresentaram o inserto, de acordo com o resultado do gel, 3 foram transferidas, com o auxílio de palitos esterilizados, para tubos Falcon de 15 mL contendo 3 mL de meio LB (Lauria Bertani) suplementados com 50µg/mL ampicilina. Em seguida, os tubos foram mantidos, a 37°C, sob agitação, durante 18 horas, em agitador orbital ajustado para 300 rpm, para possibilitar o crescimento das células de *E. coli*. Após o período, o meio contido no tubo foi utilizado para a extração de DNA dos plasmídeos.

## 2.7 Extração do DNA plasmidial - Miniprep

Os plasmídeos foram extraídos das células bacterianas por lise alcalina. O conteúdo de cada tubo foi transferido para microtubos e submetido à centrifugação por 1 minuto, 16.000 g, para sedimentação das células bacterianas. O sobrenadante foi descartado e o pellet foi ressuspensionado em 300µL de solução P1 (Tris-HCl 50 mM pH 8,0; EDTA 10 mM pH 8,0; RNase 100 µg/mL) e homogeneizado no vortex. Para a lise celular, foram adicionados, para cada tubo, 300µL de solução P2 (SDS 1%, NaOH 200mM), misturados por inversão e incubados a temperatura ambiente durante 5 minutos. Para a precipitação, foram adicionados 300µL de solução P3 (Acetato de Potássio 3M pH 5,5 com ácido acético glacial) a cada tubo e misturados por inversão, formando um precipitado branco. Em seguida, os tubos foram submetidos à centrifugação, a 16.000 g, por 10 minutos, a 14°C. Após centrifugação, o sobrenadante contendo o DNA plasmidial foi transferido para um novo tubo, onde foram adicionados 400 µL de isopropanol. Os tubos foram invertidos e centrifugados, a 16.000 g, por 10 minutos, à temperatura ambiente. O sobrenadante foi descartado. Foram adicionados 700 µL de etanol 70% para lavagem do pellet que, após agitação por

inversão, foi centrifugado, a 16.000g, por 5 minutos, a 14°C. O sobrenadante foi descartado e o pellet contendo o DNA plasmidial foi seco em capela de fluxo laminar. Após secagem completa, o pellet foi ressuspensionado em 20 µL de Tris-HCl 10mM pH 8,0. A qualidade do DNA plasmidial foi avaliada por eletroforese em gel de agarose 1,2%/EtBr, e a quantificação realizada em espectrofotômetro (Nanodrop® Spectrophotometer ND-1000) a A<sub>280</sub> nm e, posteriormente, armazenado a -20°C, para serem posteriormente submetidos ao sequenciamento.

## 2.8 Digestão com enzimas de restrição

Para confirmação da transformação bacteriana, o DNA foi digerido com enzima de restrição *EcoRI* (Invitrogen®). Foram misturados 4µL de DNA plasmidial (100 ng/µL), 2µL de tampão de reação (10x REact 3 – Invitrogen ®), 2µL de enzima *EcoRI* e 12µL de água ultrapura. A reação foi mantida em incubação, a 37°C, por 12 horas. O resultado da digestão foi visualizado por eletroforese em gel de agarose 1,2%.

## 2.9 Sequenciamento (*forward e reverse*)

Os plasmídeos contendo os insertos de DNA de interesse foram sequenciados no sequenciador automático MegaBace 1000 (Amersham Biosciences). Para o sequenciamento, foi utilizado o método enzimático baseado na síntese de DNA *in vitro*, na presença de nucleosídeos trifosfatados terminadores de cadeia (Sanger et al., 1977).

Os plasmídeos purificados foram amplificados por PCR, na presença de DNA polimerase, de um primer universal (M13) que se anela à sequência do vetor utilizado para a clonagem, de dNTPs em excesso e de uma pequena

quantidade de didesoxirribonucleosídeos trifosfatos terminadores de cadeia (ddNTPs), marcados com substância fluorescente (diclororhodamina ligada à fluoresceína). Quando ddNTPs (ddATP, ddTTP, ddCTP e ddGTP), os quais não apresentam o grupo 3'-OH da desoxirribose presente em nucleotídeos normais, são aleatoriamente incorporados na cadeia de DNA que está sendo sintetizada, a adição do nucleosídeo seguinte é bloqueada. Dessa forma, a reação enzimática gera fragmentos de DNA marcados com substância fluorescente, os quais são posteriormente separados por eletroforese e detectados, permitindo a determinação da sequência completa da fita molde de DNA. O sequenciamento foi realizado com o kit "Big Dye Terminator" (Applied Biosystems). Para a reação de sequenciamento, foram adicionados a uma microplaca para PCR contendo 96 cavidades, 1µL de cada plasmídeo e 4µL de "Big Dye Terminator" (DNA polimerase/dNTPs/ddNTPs) diluído [4X] no tampão de sequenciamento (Tris-HCl 50 mmol.L-1 e cloreto de magnésio 1,25 mmol.L-1), acrescido de 5 picomoles de primer *forward* e *reverse* M13. A reação foi realizada em termociclador Eppendorf Mastercycler gradient, utilizando-se o seguinte programa: 30 ciclos de 95°C, por 20 segundos; 50°C, por 15 segundos e 60°C, durante 1 minuto. Os produtos amplificados da reação de sequenciamento foram precipitados adicionando-se 1µL de Acetato de Amônio (7,5 mol.L-1) às amostras em cada cavidade da microplaca de PCR. Em seguida, adicionou-se duas vezes o volume com etanol 95%. Posteriormente, as placas foram vedadas com adesivo, agitadas manualmente por inversão, permanecendo à temperatura ambiente por 15 minutos. Após centrifugação, por 45 minutos, a 4.000g, 20°C, os sobrenadantes foram descartados e os sedimentos foram lavados com 100µL de etanol 70%. As placas foram centrifugadas, por 15 minutos, a 4.000g, 20°C. Os sobrenadantes foram descartados e as microplacas foram novamente centrifugadas de forma invertida sobre papel absorvente (1.000g, 1 minuto) para

remover o excesso de etanol, sendo, em seguida, mantidas a 37°C, por 10 minutos, para secagem.

Depois, foram adicionados 10µL de loading solution em cada poço da microplaca, a qual foi injetada no Mega Bace 1000. Os parâmetros de eletroforese utilizados foram: voltagem de injeção: 2 Kv; tempo de injeção: 80 segundos; voltagem de corrida: 6 Kv e tempo de corrida: 240 minutos. As sequências foram determinadas utilizando-se o programa Sequence Analyzer.

## 2.10 Análises das sequências em banco de dados

Para pesquisa de similaridade, a sequência de cDNA obtida do fragmento isolado de café foi comparada com a previamente obtida *in silico* e com sequências de bancos públicos pelo GenBank. Essa comparação com o banco de dados foi realizada utilizando-se o programa BlastX (Altschul et al., 1997).

## 2.10 Análise de expressão tecido específico do gene *FLC* por qRT-PCR

A análise da expressão gênica quantitativa por qRT-PCR foi realizada em sistema de detecção de sequências ABI PRISM 7500 Real-Time PCR (Applied Biosystems), usando SYBR Green, o qual se baseia na capacidade do fluoróforo de intercalar as fitas duplas de cDNA. O cDNA foi obtido a partir de RNA extraído de tecidos de folha, flor, fruto e raiz de café, pelos métodos descritos nos itens 2.1 e 2.3, neste capítulo.

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  $\mu$ M de cada primer e 5,0  $\mu$ L de Master Mix SYBR green UDG com ROX (Invitrogen) para um volume final 10,0  $\mu$ L/amostra. As amostras foram processadas em trélicas. Os resultados foram normalizados usando  $C_T$ s (Ciclo *Threshold*) obtidos para controles endógenos presentes na mesma reação. O  $C_T$  foi determinado pelo número de ciclos no qual a fluorescência gerada dentro de uma reação cruza a linha *Threshold*. O método usado foi o  $C_T$  comparativo.

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  $2^{-\Delta\Delta C_T}$ .

Primers foram desenhados para produzir pequenos *amplicons* (TABELA 1), geralmente entre 80 a 100pb, para maximizar a eficiência da amplificação.

As análises desse estudo foram realizadas a partir de um ensaio de quantificação relativa. A quantificação relativa é sempre baseada no transcrito referência. A normalização de um gene alvo com um controle endógeno foi feita pela expressão de um gene referência, para compensar variações internas na PCR.

A abundância de transcritos gênicos foi normalizada pela expressão do gene referência *Ubiquitina (UBI)*. O nível de expressão no tecido de raiz de *C. canephora* foi utilizado como base na comparação, já que se apresentou menor em relação aos outros tecidos. Os resultados foram normalizados usando  $C_T$ s (Ciclo *Threshold*) obtidos para controles endógenos presentes na mesma reação.

### 3 RESULTADOS E DISCUSSÃO

Para confirmar se o gene *FLC* está presente em *C. arabica*, primers foram desenhados baseando-se na sequência obtida *in silico* (capítulo 2), para isolar o gene. A partir de um RNA de qualidade, o cDNA foi sintetizado e amplificado por PCR, gerando um produto que foi clonado, e teve sua sequência confirmada através do sequenciamento, nomeada de *CaFLC*.

A sequência do gene *FLC* obtida *in silico* contém 621 nucleotídeos. Durante a amplificação por PCR, os primers CAflcT (TABELA 1) devem se anelar a essa sequência formando um fragmento com tamanho esperado de 638pb. A eletroforese em gel de agarose demonstra o tamanho do fragmento dentro do padrão esperado, ou seja, o fragmento está localizado entre as bandas 500 e 1000pb, estando mais próximo da banda menor, baseando-se no marcador Lader 1Kb (FIGURA 2).

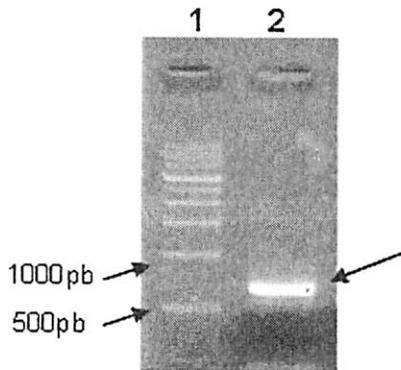


FIGURA 2. Produto da amplificação por PCR do cDNA de folhas de café (*C. arabica*). (1): marcador Lader 1kb; (2): *amplicon*

A confirmação do tamanho da sequência não foi possível porque o sequenciamento não foi completo, entretanto, por meio do sequenciamento

*reverse* obteve-se uma sequência de 233 nucleotídeos, que foi alinhada com a sequência obtida *in silico* (C15) neste trabalho (FIGURA 3), através do programa ClustalW (Thompson et al., 1994).

```

C15 -----ATGGGGGGAG-GAAGGTGGAGATTAAGAAAATCGAGGACAAG 42
CaFLC CACCGAAGSAATTAGTGATGGGGCGNGTGAAGGTGGAGATTAAGAAAATCGAGGACAAG 60
      ***** * *****

C15 AACAGCAGGCAA-GTCACGTTTTCCAAGCGGAGAAG-CGGACTGATGAAGAAAGCCAAG 100
CaFLC AACAGCAGGCAAAGTCNCGTTTTCCNAGCGGAGAAGACGGACTGATGAAGAAAGCCAAG 120
      ***** * *****

C15 AACTTTCC-GIT-CTCTGGCAGTGGATGTTGC-TGTCC-TCA-TCTTCTCTGGT----C 151
CaFLC AACTTMCCTGTINACTTCTGGACGTGGATGTTGCATGTCGGTCACTCTTCTCTGGTACGCT 180
      ***** * * *****

C15 GCGGCAAGCTCTACGACTTCTGCAGCACCAACAGTTTGGCCAAGATCCTACAACGATATC 211
CaFLC GGAANAAGCTCTACGACTTCTGCAGCACCC----CAAACANGATATAGGAACAAA--- 233
      ***** * * * * *

C15 GCAACTACGCAGAACGAGAAGACGGGTCTGCAAGAATTAGCGGGGTAGAGAAACGTAACC 271
CaFLC -----

C15 CTGAAGGCAGAAATGTCGTGACAATCAGAAAGCTGCTGGA AAAAGTTGAAAGGGATCTAG 331
CaFLC -----

C15 AGGAGCCAGATGTTGACCACCTTAACCTGAGTGAAGTACTAGTGAATGGGAAGCAACTTG 391
CaFLC -----

C15 AAGATGCACCTCATTCAAACAAGATCTAGGAAGACAGATTACTGATGGAATCAATAACCA 451
CaFLC -----

C15 GTCTAAGTGAAGTGGAAAAGATGCTGAGGGAAAGAAAACAAGCTTCTGCAAAAATAAGGTAG 511
CaFLC -----

C15 CTGCAGGTACATCCAATGAGAAGGAATGACTTGCATCCTTGAATTTGGGGACCTTACAC 571
CaFLC -----

C15 ACGTTGGAATGATTTCTGGGCAGCGACAGGCTATGCTTGAAGTACTTTAA 621
CaFLC -----

```

FIGURA 3. Alinhamento das sequências *Contig* C15 e *CaFLC* pelo programa ClustalW (Thompson et al., 1994).

O alinhamento apresentou alguns gaps (espaços vazios) para proporcionar um alinhamento mais perfeito, apesar disso, o fragmento *CaFLC* alinhado apresentou identidade com o *Contig* C15.

A comparação com o banco de dados (NCBI) foi realizada utilizando-se o programa BLASTx (Altschul et al., 1997). A sequência *CaFLC* apresentou homologia com genes MADS de várias espécies e principalmente com o gene *BvFLC* (*B. vulgaris*), cujos alinhamentos demonstraram 50% de identidade. Esse

resultado é muito relevante, já que a sequência *CaFLC* submetida ao banco de dados não estava completa.

*BvFLC* foi identificado em beterraba por Reeves et al., (2007). Em análises filogenéticas realizadas neste trabalho, ele apresentou alta homologia com o *Contig C15*, com valor *Bootstrap* 100.

Esses dados demonstram a eficiência da análise genômica feita previamente usando bioinformática, através de bancos de dados, que se constitui numa ferramenta muito poderosa para estudos moleculares. Num futuro próximo, deve ser realizado um novo sequenciamento, a fim de obter a sequência completa do gene.

Para análise de expressão tecido-específica do gene *CaFLC* em café, qRT-PCR foi realizada para comparar os níveis de expressão nos tecidos de folha (cultivares Rubi e IAPAR), flor, fruto e raiz (Rubi), todas de *C. arabica*, e ainda, com uma amostra de raiz de *C. canephora*.

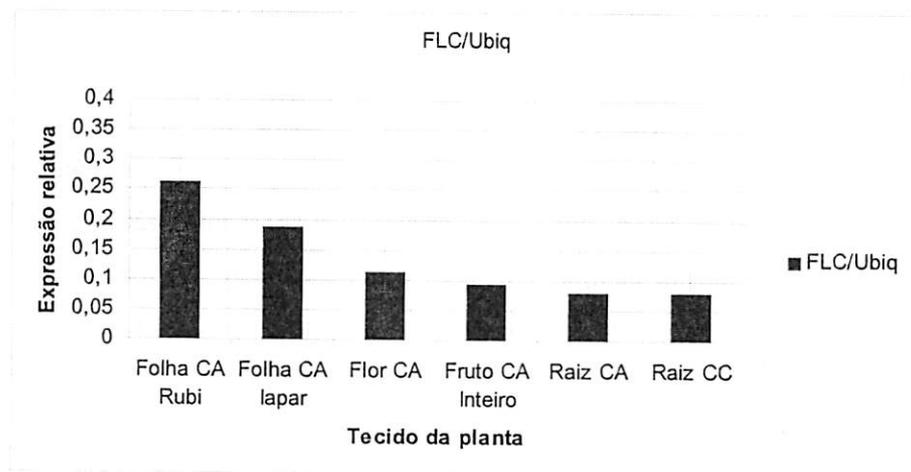


FIGURA 4. Perfil da expressão quantitativa relativa por qRT-PCR do gene *FLC*. Colunas representam a expressão de transcritos gênicos em diferentes tecidos de plantas de café, com valores dados a partir de "pool" de RNA de réplicas biológicas. A abundância de transcritos gênicos foi normalizada pela expressão dos genes *Ubiquitina*.

Avaliando-se os níveis de expressão nos diferentes tecidos, foi possível observar que a expressão quantitativa relativa do gene *CaFLC* em plantas de cafeeiro se dá diferencialmente (FIGURA 4). Os níveis de expressão observados foram significativamente maiores em tecidos de folha e flor, que em fruto e raiz.

Zhang e colaboradores (2009) analisaram o nível de expressão tecido-específica do gene *PtFLC* em citrus e observaram um perfil muito semelhante ao do café, confirmando a expressão do gene em tecidos de folha, flor, raiz e fruto, sendo os maiores níveis de expressão observados também em tecidos de folha e flor.

Um fato interessante é que em *Arabidopsis* as análises de expressão tecido-específica revelaram um padrão muito diferente, já que a expressão mRNA *FLC* não foi detectável em inflorescências jovens, em contraste com os tecidos de ápice vegetativo, raiz e haste. Com isso, Michaels & Amasino (1999) observaram que a diminuição da expressão de *FLC* está associada com a transição para o florescimento, e é possível que essa regulação seja necessária para que a transição ocorra.

Baseado nas conclusões de Michaels & Amasino (1999) sobre *Arabidopsis*, nos resultados de expressão *FLC* mostrados em café neste estudo e ainda naqueles mostrados em citrus por Zhang e colaboradores (2009), fica evidente que a regulação da transição para o florescimento, que envolve a regulação de *FLC*, se dá de forma diferente em plantas fora da família Brassicaceae. Isso indica que genes *FLC-like* fora da família Brassicaceae podem ter evoluído de um gene ancestral diferente.

Essas observações são importantes e devem ser levadas em consideração, já que *Arabidopsis* vem sendo amplamente usada como modelo para estudos relacionados ao florescimento.

Em *Arabidopsis*, a regulação do tempo de florescimento é dependente de fatores ambientais como o frio, que marca o desenvolvimento da fase vegetativa para reprodutiva (Finnegan et al., 2004; Ciannamea, 2007; Levy et al., 2002). Isso também ocorre em cafeeiros, onde a indução floral e o florescimento são regulados pela temperatura em regiões subtropicais, e a indução floral aparece com a diminuição da temperatura e se desencadeia no inverno (Camargo e Camargo, 2001).

Essa inibição da expressão de *FLC* pela vernalização é estável para a manutenção do ciclo de vida vegetal, mas a expressão *FLC* retorna em altos níveis na próxima geração em *Arabidopsis* (Michaels & Amasino, 1999).

Estudos com o gene *FLC* geram uma oportunidade para controlar o tempo de florescimento em muitas espécies, utilizando ferramentas moleculares. Em café esses estudos são de grande importância econômica, já que envolvem o problema do florescimento desuniforme, gerando grandes perdas de produção. Contudo, o tempo e regulação do florescimento são características que envolvem aspectos ecológicos e uma grande significância evolucionária.

## 4 CONCLUSÕES

O fragmento do gene *CaFLC* apresentou identidade com a sequência previamente encontrada no banco CAFEST.

*CaFLC* foi identificado nos tecidos de folha, flor, raiz e fruto de café, apresentando um nível de expressão diferencial entre os tecidos, sendo o maior perfil detectado em folhas.

Os resultados apresentados são um indício de que a regulação da transição para o florescimento, que envolve a regulação de *FLC*, pode ocorrer de forma diferente em plantas fora da família Brassicaceae, e que esses genes podem ter evoluído de um ancestral diferente.

Espera-se que a identificação de um gene chave envolvido no controle da transição do florescimento em *C. arabica* presente neste trabalho, seja utilizado em futuros trabalhos de confirmação de expressão, bem como possa marcar a importância de comparações de mecanismos moleculares que regulam a transição floral em cafeeiros.

## 5 REFERÊNCIAS BIBLIOGRÁFICAS

- ALTSCHUL, S. F.; MADDEN, T. L.; SCHAFFER, A. A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. J. Gapped BLAST and PSI BLAST: A new generation of protein database search programs. **Nucleic Acids Research**, v.25, n.17, p.3389-3402, Sept. 1997.
- BIRTIC, S.; KRANNER, I. Isolation of high-quality from polyphenol-, polysaccharide- and lipid-rich seeds. **Phytochemical Analysis**, Royal Botanic Gardens, Kew, v.17: 144–148. 2006.
- CAMARGO, A.P.; CAMARGO, M.B.P. Definição e esquematização das fases fenológicas do café arábica nas condições tropicais do Brasil. **Bragantia**, Campinas, v.60, n.1, p.65-68, 2001.
- CIANNAMEA, S. **Molecular Characterization of the Vernalization Response in *Lolium Perenne***. 2007. 147p. Tese (Doutorado) - Wageningen University, Wageningen.
- CZECHOWSKI, T.; BARI, R.P.; STITT, M.; SCHEIBLE, W.R.; UDVARDI, M.K. Real-time RT-PCR profiling of over 1400 *Arabidopsis* transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. **The Plant Journal**, v.38, p.366-379, 2004.
- FINNEGAN, E.J.; SHELDON, C.C.; JARDINAUD, F.; PEACOCK, W.J.; DENNIS, E.S. A Cluster of *Arabidopsis* genes with a coordinate response to an environmental signal. **Current Opinion in Plant Biology**, Canberra, v.14, p.911-916, 2004.
- GUIMARÃES, R.J.; MENDES, A.N.G. **Morfologia e Fisiologia do café**. Lavras: UFLA, 1998. 38p.
- HIGUCHI, R.; FOCKLER, C.; DOLLINGER, G.; WATSON, R. Kinet PCR analyses: real-time monitoring of DNA amplification reactions. **Biotechnology**, New York, v. 11: 1026-1030p. 1993.
- KOORNNEEF, M.; ALONSO-BLANCO, C.; PEETERS, A.J.M.; SOPPE, W. Genetic control of flowering time in *Arabidopsis*. **Annu Rev Plant Physiology Plant Molecular Biology**, Wageningen, v.49, p.345-370, June 1998.

- LEVY, Y.Y.; MESNAGE, S.; MYLNE, J.S.; GENDALL, A.R.; DEAN, C. Multiple roles of *Arabidopsis* VRN1 in vernalization and flowering time control. **Science**, Washington, v.297, n. 5579, p.243-246, July 2002.
- LIN, S-I.; WANG, J-G.; POON, A-Y.; SU, C-L.; WANG, S-S.; CHIOU, T-J. Differential regulation of *FLOWERING LOCUS C* expression by vernalization in *Cabbage* and *Arabidopsis*. **Plant Physiology**, Taiwan, v.137, n.3, p.1037-1048, Mar. 2005.
- MICHAELS, S.D.; AMASINO, R.M. FLOWERING LOCUS C a novel MADS domain that acts as a Repressor of Flowering. **The Plant Cell**, Madison, v.11, p. 949-956, May 1999.
- RATCLIFFE, O.J.; NADZAN, G.C.; REUBER, T.L.; RIECHMANN, J.L. Regulation of flowering in *Arabidopsis* by an FLC Homologue. **Plant Physiology**, Hayward, California, v.126, 122-132 p. may, 2001.
- REEVES, P.A., Y, H.E., SCHIMITZ R.J., AMASINO, R.M., PANELLA L.M., RICHARDS C.M. Evolutionary conservation of the FLC-mediated vernalization response: evidence from the sugarbeet (*Beta vulgaris*). **Genetics**, v.176, n.1,p.295-307, May 2007.
- SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chainterminating inhibitors. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 74, n. 12, p. 5463-5467, June 1977.
- SIMPSON, G.G.; DEAN, C. *Arabidopsis*, the rosetta stone of flowering time. **Science**, Norwich, UK, v. 296, p.285-289. 2002.
- THOMPSON, J.D.; HIGGINS, D.G.; GIBSON, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Research**, v. 22, p.4673-4680, 1994.
- ZHANG, J-Z.; LI, Z-M.; MEI, L.; YAO, J-L.; HU, C-G. *PtFLC* homolog from trifoliolate orange (*Poncirus trifoliata* ) is regulated by alternative splicing and experiences seasonal fluctuation in expression level. **Planta Journal**, Jan. 2009.