

**IDENTIFICAÇÃO E CARACTERIZAÇÃO
PARCIAL DOS GENES *SEPALLATA* EM
CAFÉ ARÁBICA**

MÁRCIA FABIANA BARBOSA DE PAULA

2009

MÁRCIA FABIANA BARBOSA DE PAULA

**IDENTIFICAÇÃO E CARACTERIZAÇÃO PARCIAL DOS GENES
SEPALLATA EM CAFÉ ARÁBICA**

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

Orientador
Prof. Antonio Chalfun Júnior

LAVRAS
MINAS GERAIS – BRASIL
2009

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

Paula, Márcia Fabiana Barbosa de.

Identificação e caracterização parcial de genes SEPALLATA em
café arábica / Márcia Fabiana Barbosa de Paula. – Lavras : UFLA,
2009.

102 p. : il.

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

Orientador: Antonio Chalfun Júnior.

Bibliografia.

1. Florescimento. 2. Coffea arábica. 3. Extração RNA. 4.
Clonagem. 5. Bioinformática. I. Universidade Federal de Lavras. II.
Título.

CDD – 633.733

MÁRCIA FABIANA BARBOSA DE PAULA

**IDENTIFICAÇÃO E CARACTERIZAÇÃO PARCIAL DOS GENES
SEPALLATA EM CAFÉ ARÁBICA**

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

APROVADA em 31 de julho de 2009.

Prof. Luciano Vilela Paiva UFLA

Prof. Marcio Alves Ferreira UFRJ

Prof. Antonio Chalfun Júnior
UFLA
(Orientador)

LAVRAS
MINAS GERAIS – BRASIL

DEDICO

Aos meus pais que sempre me apoiaram nessa caminhada

AGRADECIMENTOS

A Deus por me iluminar em todos os momentos da minha vida.

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

Ao Laboratório Central de Biologia Molecular, onde foi realizado todo o trabalho, pela estrutura disponibilizada, por esforços do prof. Luciano Vilela Paiva.

Ao prof. Marcio Alves Ferreira pela participação na banca examinadora.

Ao prof. Antonio Chalfun Júnior pela orientação, dedicação, confiança, amizade, respeito e conselhos em todos os momentos. Por todas as oportunidades oferecidas.

Aos professores do curso de pós-graduação em Biotecnologia Vegetal que muito contribuíram com aulas teóricas e práticas para o desenvolvimento do raciocínio na utilização da Biologia Molecular.

Ao pesquisador Marcelo Murrad pelas dicas.

A todos funcionários. As secretárias do curso, Ana Carolina e Magda. A Eula pelo carisma e alegria.

A todos do LCBM, em especial ao Leo, Kalynka, Douglas, Evânia, Carol, Humberto, Guilherme, pelo convívio e amizade.

A todas as turmas do curso de Biotecnologia, em especial aos meus amigos Danielle, Glacy, Rosane, Felipe, Suzana, Luciana, Cássia, Luís Gustavo, Humberto, Horllys. A Fabiane pela amizade e dedicação aos trabalhos realizados. Ao Anderson pela auxílio.

Ao grupo de trabalho do prof. Antonio Chalfun: Christiane, Carla, André, Samuel, Geraldo, Ariadne.

As minhas grandes amigas, Nádia, Brenda, Elizangela, Lilian e Solange pela paciência, amizade, carinho e companheirismo.

Aos meus amigos da graduação, que mesmo distantes torcem por mim.

A minha família, em especial meus pais, que me dão força em todos os momentos. As minhas avós, aos avôs (*in memorium*), madrinha e padrinho, tios e primos, que sempre me apoiaram.

A todos que de alguma forma contribuíram para a conclusão desse trabalho.

MUITO OBRIGADA!

SUMÁRIO

	Página
LISTA DE ABREVIATURAS	i
RESUMO GERAL	iii
GENERAL ABSTRACT	iv
CAPÍTULO 1 Café e genética do florescimento	1
1 Introdução geral	1
2 Referencial teórico	3
2.1 Café	3
2.2 Fenologia do cafeeiro	4
2.3 Florescimento	7
2.4 Genética do florescimento do modelo ABCE	10
2.4.1 MADS-box	13
2.4.2 Genes <i>SEPALLATA</i> (classe E)	15
3 Referências bibliográficas	18
CAPÍTULO 2 Identificação <i>in silico</i> de genes <i>SEPALLATA</i> em <i>Coffea arabica</i>	
Resumo	26
Abstract	27
1 Introdução	28
2 Materiais e métodos	30
2.1 Identificação de sequências <i>SEPALLATA</i>	30
2.2 Análise Filogenética	31
2.3 Identificação de motivos comuns entre as sequências	31
2.4 <i>Northern</i> eletrônico	31
3 Resultados e discussão	33
4 Conclusões	47
5 Referências bibliográficas	48
CAPÍTULO 3 Eficiência de protocolos de extração de RNA em diferentes tecidos do cafeeiro	
Resumo	53
Abstract	54
1 Introdução	55
2 Materiais e métodos	57
2.1 Material vegetal	57
2.2 Isolamento do RNA	57

2.2.1 Reagente CONCERT™	57
2.2.2 Tri Reagente®	58
2.2.3 Método Borato quente	59
2.2.4 Método CTAB	60
2.3 Tratamento com DNase	60
2.4 Síntese de cDNA e qRT-PCR	61
3 Resultados e discussão	63
4 Conclusões	70
5 Referências bibliográficas	71
CAPÍTULO 4 Isolamento e caracterização parcial dos genes <i>SEPALLATA</i> em café	
Resumo	74
Abstract	75
1 Introdução	76
2 Materiais e métodos	78
2.1 Material vegetal	78
2.2 Extração de DNA	78
2.3 Extração de RNA	79
2.4 Síntese de cDNA	80
2.5 Desenho dos <i>primers</i>	81
2.6 Amplificação dos fragmentos e eluição das bandas do gel	82
2.7 Clonagem e transformação bacteriana	83
2.8 Miniprep	86
2.9 Sequenciamento	87
2.10 Análise das sequências	88
2.11 Análise quantitativa da expressão dos genes	89
3 Resultados e discussão	91
4 Conclusões	99
5 Referências bibliográficas	100

LISTA DE ABREVIATURAS

μg	micrograma
μL	microlitro
<i>AG</i>	<i>AGAMOUS</i>
<i>AP1/2/3</i>	<i>APETALA1/2/3</i>
<i>C1/ C1SEP1/2</i>	<i>contig 1</i>
<i>C8/ C8SEP4</i>	<i>contig 8</i>
<i>C14/ C14SEP3</i>	<i>contig14</i>
CTAB	brometo de cetil trimetilamônio
<i>DEF</i>	<i>DEFICIENS</i>
DEPC	Dimetil pirocarbonato
DNA	ácido desoxirribonucléico
DNase/ RNase	endonucleases
dNTP	deoxirribonucléico trifosfato
DTT	Ditiotreitol
EDTA	ácido etilenodiamonotetracético
EGTA	ácido etileno glicol tetrácido
EST	Expressed Sequence Tags
FB1/FB2/FB3/FB4	fases de desenvolvimento dos botões florais
FL	abertura das flores
FR2	fruto
g	grama (unidade de peso) / gravidade (centrífuga)
G1 – G6	fases de desenvolvimento das gemas
KOAc	Acetato de potássio
LB	Meio Luria Bertani
LiCl	Cloreto de Lítio
M	Molar (unidade de molaridade)

mM	micro molar
<i>MADS box</i>	família gênica de fatores de transcrição
min	minuto
MgCl ₂	Cloreto de Magnésio
MIKC	domínios dos MADS box
mL	mililitro
mRNA	RNA mensageiro
NaCl	Cloreto de sódio
NaOH	Hidróxido de sódio
NCBI	Centro Nacional de Informações em Biotecnologia
pb	pares de bases
PCR	Reação em Cadeia da Polimerase
pH	logaritmo do inverso da concentração de H ⁺
<i>PI</i>	<i>PISTILLATA</i>
PVP	polivinilpirrolidona
PVPP	polivinilpolipirrolidona
qRT PCR	PCR em tempo real
RNA	ácido ribonucléico
rRNA	RNA ribossômico
RT PCR	reação da polimerase em cadeia a partir da transcrição reversa
seg	segundo
<i>SEP/ SEP</i>	<i>genes SEPALLATA/ proteínas SEPALLATA</i>
<i>SEP1/2/3/4</i>	<i>SEPALLATA1/2/3/4</i>
<i>SHP1/2</i>	<i>SHATERPROOF1/2</i>
<i>STK</i>	<i>SEEDSTICK</i>
u	Unidade

RESUMO GERAL

PAULA, Márcia Fabiana Barbosa de. **Identificação e caracterização parcial de genes *SEPALLATA* em café arábica**. 2009. 102p. Dissertação (Mestrado em Agronomia/Biotecnologia Vegetal)-Universidade Federal de Lavras, Lavras, MG.*

O café é um dos mais importantes *commodity* internacionais, produzido em mais de sessenta países e classificado entre as cinco exportações agrícolas mais valorizadas de países em desenvolvimento. Uma das maiores dificuldades dessa cultura consiste na heterogeneidade do florescimento, conseqüentemente, desuniformidade na maturação dos frutos, o que dificulta a colheita e prejudica a qualidade dos grãos. Tendo em vista a importância econômica do café e a dificuldade apresentada, este trabalho surgiu a partir da necessidade de um melhor entendimento dos fatores genéticos que compõe essa rota, já que o processo de florescimento depende da expressão equilibrada de uma rede complexa de genes. Os genes *SEPALLATA* (*SEP*) fazem parte dessa rota e são responsáveis pela determinação dos quatro verticilos florais, indicando com a ajuda de outros genes a identidade do meristema floral. Desta forma, esse trabalho teve como objetivo estudar os genes *SEP* em *Coffea arabica*. Primeiramente foram realizadas análises *in silico* para identificar esses genes no genoma café, por meio do banco de dados CAFEST. Três prováveis genes *SEP*, foram identificados e submetidos a análises de motivos, domínios, agrupados em árvores filogenéticas e análise da expressão por *Northern* eletrônico. Essas seqüências apresentaram os domínios característicos da família gênica, agrupando com genes *SEP* de outras espécies, e como esperado, a expressão foi observada em bibliotecas de flores e frutos em desenvolvimento. Os experimentos para caracterização parcial foram realizados no Laboratório de Biologia Molecular da Universidade Federal de Lavras, UFLA, onde, após serem desenhados primers específicos, essas *amplicons* foram submetidas à clonagem, sequenciamento e comparação com genes *SEP* obtidas na análise *in silico*. Por meio de comparação, somente duas seqüências foram consideradas similares e caracterizadas como possíveis genes *SEP1/2* e *SEP3*. Além disso, foi realizada a análise de expressão por PCR em tempo real, onde também foram observadas as expressões dos fragmentos em tecidos de flores e frutos. Com todas essas análises foi possível a identificação e caracterização parcial de pelo menos dois genes *SEP* em café, porém, se faz necessário, a partir dos resultados obtidos, de trabalhos de expressão *in situ* e estudos de mutantes para uma melhor caracterização desses genes.

* Orientador: Antonio Chalfun Júnior - UFLA

GENERAL ABSTRACT

PAULA, Márcia Fabiana Barbosa de. **Identification and partial characterization of *SEPALLATA* genes in coffee arabica**. 2009. 102p. (Master in Agronomy/Plant Biotechnology)-Universidade Federal de Lavras, Lavras, MG.*

Coffee is one of the most important international commodity, produced in more than sixty countries and it is classified as one of the five agricultural exportation most valued in development countries. One of the biggest difficulties in this crop is the heterogeneity of the flowering process, therefore, imbalance on the fruit maturation, which makes the harvesting more difficult to be done, damaging grains quality. Since coffee has the economic importance and its difficulties presented, this research aims at understanding some genetic factors composing this pathway, since flowering process depends on a balanced expression of a complex network of genes. *SEPALLATA* genes constitute part of this pathway and are responsible for determination of the four floral whorls. *SEPALLATA* together with other genes determinate the floral meristem identity. Therefore the aim of this work was to study *SEPALLATA* genes in *Coffea arabica*. Firstly, *in silico* analyses were done to identify these genes in coffee genome, by *CAFEST* database. Three putative *SEP* genes were identified and submitted to analysis of the motifs, domains, grouped in phylogenetic tree and expression analysis by electronic *Northern*. The obtained sequences presented all the characteristic domains of this gene family, when grouped with *SEP* genes of other species, and as expected, the expression was observed in flower and developing fruits libraries. The parcial characterization was performed at the Central Laboratory of Molecular Biology of the Federal University of Lavras, UFLA. Specific primers were designated to clone the genes and the amplicons were submitted to cloning, sequencing, and compared with *SEP* genes obtained from the *in silico* analyses. By this comparison, two sequences were considered similar and characterized as putative *SEPI/2* and *SEP3* genes. In addition to that, an expression analysis by qRT-PCR was performed, where it could be observed the expression in flower and fruit tissues. These analyses provided the identification and characterization of, at least, two coffee *SEP* genes, however, it is necessary more investigation such as *in situ* expression and mutants to shed light on the *SEP* genes function in coffee.

* Adviser: Antonio Chalfun Júnior - UFLA

CAPÍTULO 1

CAFÉ E GENÉTICA DO FLORESCIMENTO

1 INTRODUÇÃO GERAL

O cafeeiro (*Coffea sp.*) pertence a família *Rubiacea* e abriga cerca de 100 espécies. Destas, apenas duas têm importância econômica: *Coffea arabica* e *Coffea canephora*, que respondem por cerca de 70% e 30%, respectivamente, da produção mundial de café e produzem grãos conhecidos respectivamente, como arábica e robusta. No Brasil é considerado um dos principais produtos agrícolas, sendo o segundo item em importância do comércio internacional de *commodities*. Porém, uma das maiores dificuldades na produtividade e qualidade dos grãos, é o florescimento sequencial que causa uma maturação desuniforme dos frutos, dificultando a colheita e causando perdas econômicas de produção.

O florescimento é um evento transitório da fase vegetativa para a fase reprodutiva, onde as plantas desenvolvem mecanismos para garantir a perpetuação da espécie, e evoluem para integrar sinais endógenos e ambientais, no controle preciso do florescimento. Em *Arabidopsis thaliana*, estudos genéticos e moleculares estabeleceram a presença de quatro rotas distintas que controlam e sincronizam o desenvolvimento floral: o comprimento do dia (fotoperíodo), tratamento de frio (vernalização), via autônoma e via das giberelinas (Blasquez, 2000). Fatores ambientais como, suprimento de água, temperatura do ar e níveis de irradiância influenciam no florescimento, desenvolvimento dos frutos e produtividade. 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.

A compreensão e identificação do mecanismo de florescimento, assim como os genes envolvidos nos processos de regulação floral, são de fundamental importância, para elucidar como esses mecanismos ocorrem em café.

Vários genes já foram identificados como peças-chaves nessa regulação molecular, e nesse complexo, além de alguns serem responsáveis pelas vias de interação entre o ambiente e a planta, outros são responsáveis diretamente na formação da flor. O modelo ABCE representa quatro classes de genes que são responsáveis pela identidade floral.

Os genes da classe E, *SEPALLATA*, são pertencentes à família de fatores de transcrição *MADS-box*, e possuem um papel fundamental na determinação do meristema floral e desenvolvimento de frutos (Pelaz et al., 2000; Honma & Goto, 2001; Ditta et al., 2004). Em *Arabidopsis* são identificados quatro genes *SEP* que atuam de forma redundante, mas o número de genes, seu desenvolvimento e funções bioquímicas, variam de acordo com a espécie.

Neste contexto, o principal objetivo deste trabalho foi estudar os genes envolvidos nesse complexo de regulação gênica do florescimento, por meio da análise de genes essenciais na determinação floral, como os genes *SEP* do cafeeiro.

2 REFERENCIAL TEÓRICO

2.1 Café

O Cafeeiro é pertencente ao gênero *Coffea* da família *Rubiaceae*, e conta com aproximadamente 100 espécies identificadas (Bridson & Verdcourt, 1988). Todas as espécies conhecidas são diplóides ($2n=2x=22$ cromossomos) e obrigatoriamente munidas de sistemas de autoincompatibilidade reprodutiva. Com exceção da espécie *C. arabica* ($2n=4x=44$), única espécie poliploide no gênero *Coffea* cuja origem vem de duas diferentes espécies diploides, *C. canephora* e *C. eugenioides* (Lashermes et al., 1999), e possui taxa de autofecundação em torno de 90% (Conagin & Mendes, 1961; Charrier & Berthaud, 1985).

Dentre as mais de 100 espécies no gênero *Coffea* (Cros, 1994), somente *C. arabica* e *C. canephora* (Gonçalves & Silvarolla, 2001) são cultivadas comercialmente, correspondendo com 70% e 30%, respectivamente, da produção mundial.

O café é considerado um dos principais produtos agrícolas, sendo o segundo item em importância do comércio internacional de *commodities*, e é um dos mais importantes produtos agrícolas comercializados no mundo e no Brasil. Atualmente o Brasil é responsável por 30% do mercado internacional, volume equivalente à soma da produção dos outros seis maiores países produtores. É também o segundo mercado consumidor, atrás somente dos Estados Unidos. As áreas cafeeiras estão concentradas no centro-sul do país, onde se destacam quatro estados produtores: Minas Gerais, São Paulo, Espírito Santo e Paraná (ABIC). A Companhia Nacional de Abastecimento-CONAB (2009) divulgou a segunda estimativa de safra de café para 2009, propondo uma safra de 39 milhões de sacas. A produção do café arábica representa 72,49% (28,3 milhões

de sacas de café) da produção do País, e tem como maior produtor o estado de Minas Gerais, com 66% (18,97 milhões de sacas).

2.2 Fenologia do cafeeiro

A fenologia compreende o estudo de eventos periódicos da vida da planta em função da sua reação às condições do ambiente. O ciclo fenológico do cafeeiro apresenta duas fases, as quais ocorrem simultaneamente: vegetativa e reprodutiva. As plantas de café são classificadas como sendo bienais, ou seja, demoram dois anos para completar o ciclo, diferente da maior parte das plantas, que florescem e frutificam no mesmo ano fenológico.

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

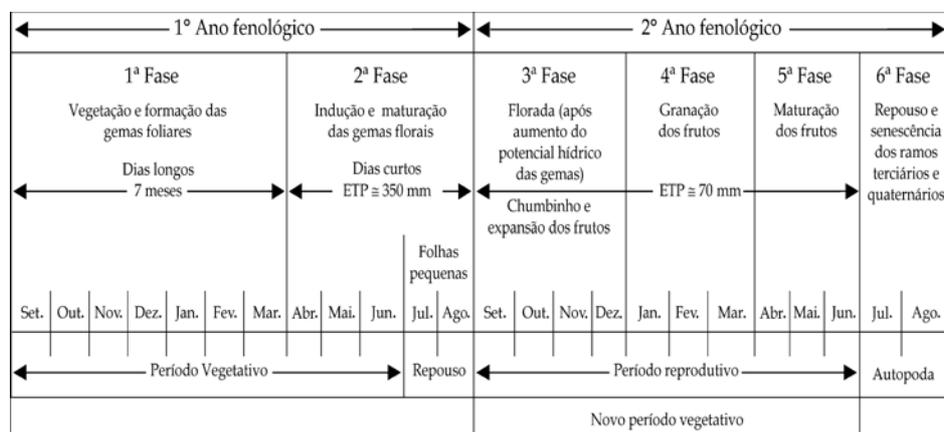


FIGURA 1 Esquematização das duas fases fenológicas do cafeeiro arábica (fase vegetativa e reprodutiva), subdivididas em seis, durante 24 meses, nas condições climáticas tropicais do Brasil (Camargo & Camargo, 2001).

A primeira fase, vegetação e formação das gemas vegetativas, ocorre normalmente de setembro a março. São sete meses de dias longos, com fotoperíodo acima de 13 a 14 horas de luz efetiva ou acima de 12 horas de brilho solar (Camargo, 1985).

A segunda fase, indução e maturação das gemas florais, é caracterizada por dias curtos, indo normalmente de abril a agosto, onde é observado um crescimento das gemas florais existentes. Após o completo desenvolvimento, entram em dormência e ficam prontas para a antese, que acontecerá quando ocorrer um aumento substancial de seu potencial hídrico, causado pela chuva ou irrigação. Nos dois meses finais dessa etapa, julho a agosto, as gemas dormentes produzem um par de folhas pequenas, separando o primeiro ano fenológico do segundo (Mes, 1957; Gouveia, 1984).

A terceira fase, de setembro a dezembro, inicia-se com a florada, após um aumento do potencial hídrico nas gemas florais maduras (choque hídrico), e vai até a expansão dos frutos.

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.

Para uma melhor visualização de algumas etapas do desenvolvimento floral do cafeeiro, Moraes et al. (2008), elaboraram uma escala que representa desde a formação da gema até o desenvolvimento dos frutos (Figura 2). A passagem da gema à fase reprodutiva (indução) compreende uma sequência de eventos de natureza bioquímica e morfológica que começa com o estímulo indutivo. Em *Coffea arabica* um leve achatamento do meristema e sua elevação acima do nível original são os primeiros indícios visuais de transição floral (Rena & Maestri, 2004). Na sequência são lançadas pequenas gemas que vão crescendo e ficando visivelmente diferenciadas (G1 – G6). Na fase representada

como FL ocorre a antese e as fases caracterizadas como F são representações do desenvolvimento dos frutos.

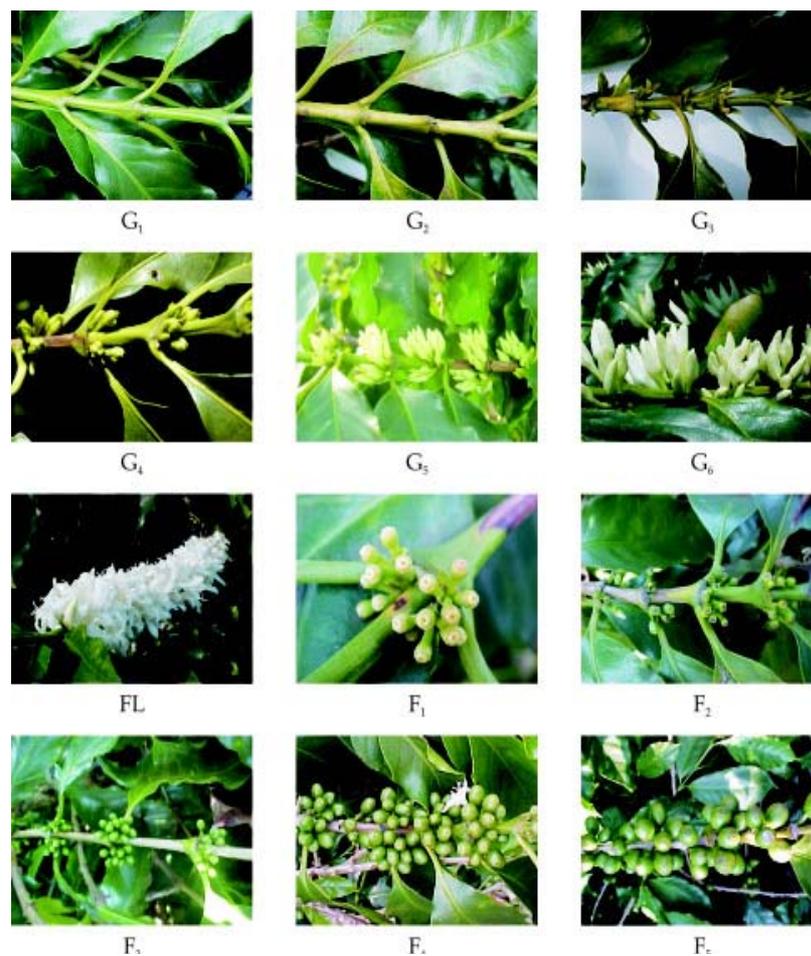


FIGURA 2 Escala ilustrativa de algumas fases reprodutivas do cafeeiro. G1 - refere-se aos nós com gemas indiferenciadas; G2 - nós com gemas intumescidas; G3 - gemas com até 3 mm de comprimento; G4 - gemas medindo 3,1 a 6 mm de comprimento; G5 - gemas de 6,1 a 10 mm (coloração verde claro); G6 - gema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

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

2.3 Florescimento

A fase reprodutiva é marcada pela capacidade da planta em produzir flores e é o resultado das mudanças que ocorrem no meristema das gemas (Larcher, 2000). O florescimento compreende uma sequência de eventos morfofisiológicos, que vai desde a indução floral até a antese, passando pelas fases intermediárias da evocação floral, diferenciação ou iniciação dos primórdios florais e crescimento da flor (Rena & Maestri, 1987). Em café essas etapas envolvem uma complexa sequência de eventos genéticos, bioquímicos, fisiológicos e morfológicos, os quais são afetados por vários fatores como temperatura, luz, solo, ar e disponibilidade de água (Rena & Barros, 2004). Verificando a distribuição sazonal das chuvas, Maestri & Barros (1997) observaram que as oscilações de temperatura do ar e a variação fotoperiódica são os fatores climáticos mais importantes no controle do crescimento e florescimento do *C. arabica*.

A floração do café é conhecida como floração gregária, em que todas as plantas individuais florescem simultaneamente numa certa extensão geográfica, entretanto o número de floradas varia desde umas poucas até várias ao longo do ano nas regiões equatoriais (Soares et al., 2005). Em regiões cafeeiras de latitudes médias, como no Sudeste e algumas regiões do Sul do Brasil, que apresentam época seca e fria bem definida, podem aparecer de 2 a 3 floradas de intensidades decrescentes. Em regiões equatoriais chuvosas, como Costa Rica,

pode ser observado até 15 floradas, isso pode dificultar a colheita, o controle de pragas e doenças e até interferir na qualidade final dos grãos (Rena & Maestri, 1985; Rena & Barros, 2004; Silva et al., 2005).

Renna & Barros (2004) dividiram a floração do café em quatro fases: iniciação, diferenciação floral, repouso ou quiescência e abertura floral (antese). Cada fase é afetada por vários fatores, sendo difícil isolar o efeito de um fator específico, pois um mesmo fator pode afetar mais de uma fase, o que torna difícil o enfoque experimental e a interpretação do fenômeno (Kumar, 1979).

O processo de iniciação floral culmina com a formação dos primórdios florais claramente reconhecíveis ao microscópio, sendo precedido de reações fisiológicas que resultam na produção do “estímulo floral” e da “evocação” do meristema, e após essa fase, irreversivelmente irá transformar-se em flor ou inflorescência (Rena & Barros, 2004). Não há como delimitar a indução floral no cafeeiro, pois são processos quantitativos e graduais que se sobrepõem sem uma nítida separação. Dessa forma, fatores como luz, temperatura, água e déficit de pressão de vapor (DPV) da atmosfera têm seus efeitos confundidos durante essa fase e até na diferenciação das partes florais.

A diferenciação do botão floral é amplamente controlada por hormônios vegetais, os quais são ativados principalmente pelo fotoperíodismo (comprimento do dia) e por uma queda de temperatura do ar (estação fria) (Wintgens, 2004). Observando externamente, só se pode assegurar que as gemas estão floralmente determinadas quando os botões, ainda verdes, se tornam individualmente visíveis e prontos para se abrirem sob a ação de um estímulo. Assim, uma gema com forma achatada e triangular poderá ou não estar floralmente diferenciada, mas se ela estiver inchada e envolvida pelas brácteas rudimentares espessadas, coberta por uma mucilagem, já estará diferenciada num botão floral (Rena & Maestri, 1985; Rena & Barros, 2004).

O período de dormência ocorre durante a estação seca, coincidindo com um baixo crescimento vegetativo, podendo durar semanas ou meses, dependendo das condições ambientais. Existem duas etapas durante a paralisação do crescimento das gemas florais. Primeiramente há uma dormência endógena causada pela presença de substâncias inibidoras, a qual seria revertida durante a seca. A partir de um dado momento, passaria a existir uma quiescência ou dormência imposta, que se estenderia enquanto durasse o estresse dos cafeeiros por deficiência hídrica. Em experimentos realizados no Peru, região costeira onde raramente chove, verificou-se que nos cafeeiros constantemente irrigados os botões florais não se abriam, ao contrário da abertura floral regular quando os cafeeiros eram submetidos a déficit de água. Desta forma, alguns autores sugerem que um período de seca é necessário para quebrar a dormência dos botões e que a subsequente adição de água é necessária apenas para a indução do crescimento das gemas quiescentes, fenômeno denominado “hidroperiodismo” (Alvim, 1973; Rena & Maestri, 1985).

Segundo Rena & Barros (2004), outra característica que pode ser a causa primária desse fenômeno da dormência, é que durante a época seca, os botões florais acumulam grandes quantidades de inibidores como o ácido abscísico, responsável por cerca de 75% da dormência, estabelecendo-se um balanço desfavorável na relação entre promotores e inibidores do crescimento.

A última fase, a antese, segundo Barros et al. (1978), é promovida pelas seguintes condições: (i) uma queda rápida da temperatura, isoladamente, quando o balanço hídrico das plantas não for crítico; (ii) chuvas abundantes ou irrigação após um longo período de seca; (iii) quedas bruscas na temperatura e suprimento de água (chuva ou irrigação) atuando complementarmente ou sinergicamente. Após a dormência ser quebrada, os botões florais reiniciam imediatamente seu crescimento, levando à abertura das flores 8 a 15 dias mais tarde (Camargo & Camargo, 2001; Wintgens, 2004). No campo, esse estímulo ocorre quando há

uma “chuva de florada” no final da estação seca, sendo necessário no mínimo 10 mm de chuva, segundo Camargo & Camargo (2001), Rena & Barros (2004) e Wintgens (2004).

As floradas, na maioria das regiões cafeeiras do Brasil, tanto coincidem com o início do rápido crescimento vegetativo (agosto/setembro) como têm continuação durante a estação de crescimento máximo (novembro/dezembro) (Rena & Barros, 2004). A abertura das flores se dá nas primeiras horas da manhã e, no segundo dia, a corola começa a murchar, para cair no dia seguinte (Thomaziello et al., 2000).

A uniformidade da floração tem implicações diretas na uniformidade de maturação dos frutos de café, a qual, por sua vez, terá grande influência na qualidade final do produto. Desta forma, o conhecimento de técnicas viáveis de uniformização da floração é de fundamental importância para a organização e redução dos custos da colheita e para a obtenção de um produto de melhor qualidade e maior valor agregado.

2.4 Genética do modelo de florescimento ABCE

Como as plantas não possuem a capacidade de mover-se para ambientes favoráveis, elas desenvolveram mecanismos que permitem determinar o melhor momento de produzir suas estruturas reprodutivas e assim aumentar as chances de perpetuação das espécies. Desta forma a regulação correta da transição para o florescimento é essencial para o desenvolvimento reprodutivo das plantas, por isso apresentam controles moleculares conservados para integrar tanto indícios ambientais como endógenos na regulação precisa do tempo de florescimento (Dornelas & Rodriguez, 2005; Ciannamea, 2007).

Os indícios endógenos são representados por três classes de genes que regulam o desenvolvimento floral e 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).

A especificação dos quatro verticilos é controlada por quatro classes de genes seletores homeóticos (Haughn & Somerville, 1988; Schwarz-Sommer et al., 1990; Coen & Meyerowitz, 1991), em um modelo definido como modelo de florescimento ABCE.

Esse modelo foi descoberto na década de 90, e à medida que se conheciam novos mutantes homeóticos, o modelo foi gradativamente aperfeiçoado. Em 1991, os pesquisadores Coen & Meyerowitz (1991) descreveram o desenvolvimento floral a partir de um modelo, denominado modelo ABC. Nesse modelo eles classificaram alguns genes que determinavam a formação dos verticilos florais em três diferentes classes e tentaram explicar como a ação combinada dessas três classes determinava a identidade dos quatro órgãos florais. Desta forma, determinaram em *Arabidopsis* que os genes *APETALA1* (*AP1*) e *APETALA2* (*AP2*) são pertencentes à classe A; *APETALA3* (*AP3*) e *PISTILLATA* (*PI*) à classe B, e o gene *AGAMOUS* (*AG*) à classe C. A interação entre esses genes promove a identidade de diferentes verticilos florais. A expressão de genes do tipo A, induzem a formação das sépalas. Quando expressos genes do tipo A e B, pétalas são formadas. A expressão dos genes tipo B e C leva à formação do estame, e quando somente os genes do tipo C são expressos, desenvolvem capelos. Os genes classificados como A e C atuam de maneira antagônica, e de acordo com o modelo, flores mutantes da

função A transformam sépalas em carpelos, e pétalas em estames quando mutante da função C.

Ao produzirem um triplo mutante para função ABC, onde os órgãos florais transformaram-se em folhas, descobriu-se a formação de um complexo ternário com genes chamados *SEPALLATA* (*SEP*), adicionando ao modelo a função E (*SEPI/2/3*). Ao verificar que os três genes *SEP* são altamente redundantes e os mutantes para cada um dos genes apresentam apenas uma leve alteração no fenótipo, foi construído um triplo mutante *SEPI, 2, 3* que apresentou o fenótipo igual ao da perda de função dos genes das classes B e C, ou seja, as pétalas, estames e carpelos foram substituídos por órgãos sepalóides. Este fato indicou que os genes *SEP* eram necessários na formação dos três verticilos florais (Pelaz et al., 2000; Honma & Goto, 2001). Ditta et al. (2004) descreveu outro gene pertencente à classe E, que possuía a função de determinar além dos outros verticilos, também as sépalas, sendo denominado *SEP4*. Desta forma foi determinado que os quatro genes *SEP* possuem um papel central na identidade de todos os verticilos florais.

Atualmente o modelo ABCE é representado pela formação de um complexo de genes. Determinando aos genes das classes A e E, a formação das sépalas, o primeiro verticilo floral; genes da classe A, B e E especificando pétalas, o segundo verticilo; genes da classe B, C e E especificando estames, o terceiro verticilo; genes da classe C e E especificando carpelos, o quarto verticilo; e genes da classe D e E especificando a formação do óvulo (Shitsukawa et al., 2007) (Figura 3).

Os genes pertencentes ao modelo ABCE, exceto o *AP2*, são pertencentes a uma família de fatores de transcrição chamados de *MADS-box*, que além das funções especificadas no modelo de florescimento, estão envolvidas também em várias etapas do desenvolvimento de plantas, animais e fungos.

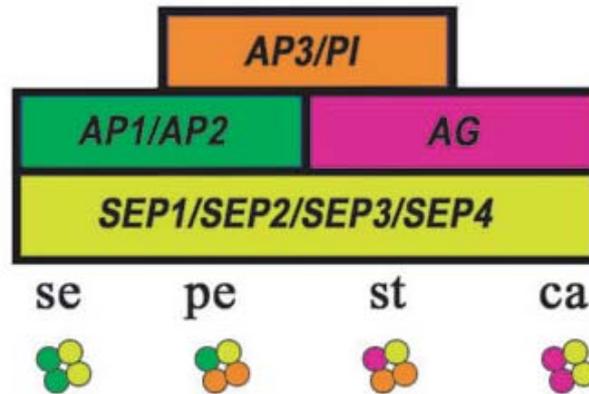


FIGURA 3 Modelo “ABCE” de identidade dos órgãos florais. Representação em laranja – genes classificados como B; verde – genes A; rosa – genes C; Amarelo – genes E; e as combinações para a formação dos verticilos. Adaptado de Castillejo et al. (2005)

2.4.1 *MADS-box*

A família de genes *MADS-box* é composta de fatores de transcrição que possuem papel fundamental no controle do desenvolvimento e transdução de sinais em provavelmente todos eucariotos (Becker & Theißen, 2003). São envolvidos em muitos aspectos do desenvolvimento da planta, incluindo a identidade do meristema, tempo de florescimento, determinação da flor, identidade do órgão floral, fertilidade do pólen, desenvolvimento do óvulo, identidade e desenvolvimento do fruto (Causier et al., 2002), e alongação lateral da raiz (Zhang & Forde, 1998). Também foram encontrados genes *MADS-box* sendo expressos em folhas, nas células-guarda e em tricomas (Alvarez-Buylla et al., 2000).

O nome da família gênica *MADS-box* é derivado das iniciais dos primeiros genes descritos, M de *MINICHROMOSOME MAINTENANCE1*

(*MCM1*) de leveduras, A de *AGAMOUS (AG)* de *A. thaliana* (Yanofsky et al., 1990), D de *DEFICIENS (DEF)* de *Antirrhinum majus* (Sommer et al., 1990; Schwarz-Sommer et al., 1992) e S de *SERUM RESPONSIVE FACTOR (SRF)* de humanos (Norman et al., 1988). As proteínas *MADS-box* foram primeiro caracterizadas em leveduras e mamíferos (Lamb & McKnight, 1991; Shore & Sharrocks, 1995).

Todos os membros compartilham uma sequência altamente conservada de aproximadamente 180 nucleotídeos, que codifica um domínio de ligação ao DNA, denominado de *MADS*. O domínio *MADS* reconhece a sequência específica [CC(A/T)₆GG], denominada *CArG-box*, na região regulatória de seu gene alvo (Hayes et al., 1988; Shore & Sharrocks, 1995; Riechmann et al., 1996).

Existem outros domínios importantes na interação entre as proteínas. Estudos filogenéticos demonstraram que a duplicação do gene *MADS-box* ancestral ocorreu antes da divergência de plantas, animais e fungos, dando origem a duas linhagens, que foram caracterizadas como: tipo I (SRF-like) e tipo II (MEF2-like) (Alvarez-Buylla et al., 2000). As proteínas *MADS-box* de plantas melhor caracterizadas fazem parte da categoria tipo II e apresentam organização estrutural modular e conservada, chamada domínio MIKC, devido à presença de quatro domínios característicos da região amino terminal (N) até a região carboxy terminal (C) da proteína: *MADS-box* (M); intermediário (I), queratina-like (K) e carboxi terminal (C). O M é um domínio altamente conservado e liga-se a região do DNA formando homodímeros. A região I também participa da interação homodímeros. O domínio K é moderadamente conservado e esta presente somente em proteínas de plantas, envolvido na dimerização entre proteínas e na especificidade funcional (Davies et al., 1996; Fan et al., 1997). O domínio I, pouco conservado, ajuda determinar a formação dos dímeros que se ligam ao DNA (Riechmann & Meyerowitz, 1997). Por fim, o domínio C-

terminal, pouco conservado, pode funcionar como um domínio de trans-ativação (Egea-Cortines et al., 1999; Honma & Goto, 2001; Robles & Pelaz, 2005).

São encontrados 39 genes *MADS-box* do tipo II em *Arabidopsis* (Parenicová et al., 2003), que possuem várias funções no desenvolvimento da planta, principalmente na determinação das funções homeóticas florais, representada pelo modelo ABCE. E dos genes responsáveis por esse modelo, os pertencentes à classe E são de fundamental importância na determinação de cada verticilo, formando complexo com outros genes.

2.4.2 Genes *SEPALLATA* (classe E)

Em *Arabidopsis*, são encontrados quatro membros da classe E de genes *MADS-box* para identificação dos órgãos florais que exibem funções redundantes, *SEP1*, *SEP2*, *SEP3* e *SEP4*. Esses genes *SEP* podem formar complexos de heterodímeros protéicos com os produtos das classes A, B, C regulando a identidade das sépalas, pétalas, estames, carpelos (Honma & Goto, 2001; Pelaz et al., 2001; Ferrario et al., 2003) (Figura 4). A estrutura desses complexos protéicos ainda é desconhecida, mas sabe-se que as proteínas *SEP* apareceram para contribuir com cada complexo formado.

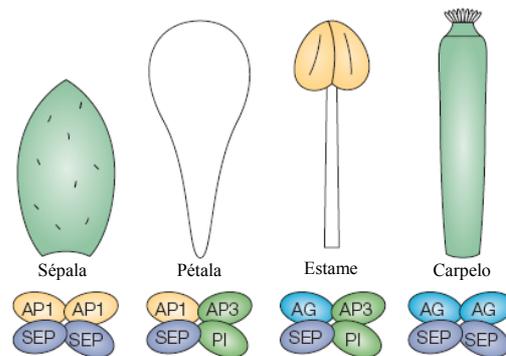


FIGURA 4 Complexos protéicos na determinação dos verticilos. Adaptado de Krizek & Fletcher (2005).

As funções desses genes foram apresentadas analisando mutantes. Mutantes simples para os genes *SEP* possuem somente fenótipos sutis, ao contrário, mutantes triplos para os genes *SEP1*, *SEP2* e *SEP3* possuem flores transformadas em órgãos sepalóide, e mutantes quádruplos para genes *SEP1*, *SEP2*, *SEP3* e *SEP4* possuem um fenótipo com variações mais drásticas, transformando as flores em estruturas foliares (Pelaz et al., 2000; Honma & Goto, 2001; Ditta et al., 2004), dando suporte à hipótese de que as flores são consideradas folhas modificadas (Goethe, 1790).

Ditta et al. (2004) demonstraram efeitos de dosagens e redundância para os genes *SEP*. Isso pode ser observado no desenvolvimento do óvulo, onde os genes *SEP* agem em doses dependentes dos genes do tipo C (AG) e genes D (STK (SEEDSTICK); SHP1 (SHATERPROOF1); SHP2 (SHATERPROOF2) (Favaro, 2003).

Os quatro genes *SEP* em *Arabidopsis* possuem funções redundantes, mas geralmente em muitas espécies essas funções e o número de genes presentes, ainda são desconhecidos. Entretanto, diversos papéis nos órgãos

florais têm sido elucidados. A expressão de genes *SEP3* em petúnia, chamados de PhFBP2 (Angenent et al., 1992), e em tomate LeTM5 (Pnueli et al., 1994), demonstraram um rompimento das estruturas florais, enquanto que a baixa expressão em gérbera (GhGRCD1) levou somente à esterilidade dos estames (Kotilainen et al., 2000). A baixa expressão do gene *SEP1/2* em tomate (LeTM29) (Ampomah-Dwamena et al., 2002), gérbera (GhGRCD2) (Uimari et al., 2004) e em pêssego (PrpMADS7) (Xu et al., 2008) produziram plantas com alterações no fenótipo, entretanto em *Populus tremula* o fenótipo formado foi indistinguível da planta do tipo selvagem (Cseke et al., 2005).

Os efeitos dos genes *SEP* contribuindo na regulação do florescimento também são variáveis. Em *Arabidopsis*, a superexpressão do gene *SEP3* estimula um florescimento precoce (Pelaz et al., 2001) e efeitos similares foram observados em outras espécies, como em lírio (Tzeng et al., 2003), petúnia (Ferrario et al., 2003), arroz (Kang & An, 1997), *Betula pendula* (Lemmetynen et al., 2004), trigo (Zhao et al., 2006) e em pêssego (Xu et al., 2008). A expressão dos genes *SEP* de *Capsicum* (CaMADS1) de *Nicotiana* (NsMADS4) produzem plantas com florescimento precoce e curto (Sung et al., 2001), mas diferentemente, em *Tabacco* (Jang et al., 2002) não modificam a forma das plantas transgênicas. Entretanto, as plantas de *Arabidopsis* que expressam *SEP* de *Populus tremula* (Cseke et al., 2005), de lírio (Tzeng et al., 2003) e de pêssego (Xu et al., 2008), não demonstraram mudanças visíveis comparadas com a selvagem.

Os genes *SEP* também possuem um importante papel no desenvolvimento de frutos. Estudo com o gene *SEP4* (LeMADS-RIN) em tomate, demonstrou que é necessário para o desenvolvimento do fruto porque não houve o amadurecimento em mutantes *rin* (Vrebalov et al., 2002).

3 REFERÊNCIAS BIBLIOGRÁFICAS

- ALVAREZ-BUYLLA, E.R.; LILJEGREN, S.J.; PELAZ, S.; GOLD, S.E.; BURGEFF, C.; DITTA, G.S.; VERGARA-SILVA, F.; YANOFSKY, M.F. MADS-box gene evolution beyond flowers: expression in pollen, endosperm, guard cells, roots and trichomes. **Plant Journal**, Oxford, v.24, n.4, p.457-466, Nov. 2000.
- ALVIM, P.T. Factors affecting flowering of coffee. In: SBR, A.M. (Ed.). **Genes, enzymes and population**. New York: Plenum, 1973. p.193-202.
- AMPOMAH-DWAMENA, C.; MORRIS, B.A.; SUTHERLAND, P.; VEIT, B.; YAO, J.L. Down-regulation of TM29, a tomato SEPALLATA homolog, causes parthenocarpic fruit development and floral reversion. **Plant Physiology**, Washington, v.130, n.2, p.605-617, Sept. 2002.
- ANGENENT, G.C.; BUSSCHER, M.; FRANKEN, J.; MOL, J.N.M.; TUNEN, A.J. van. Differential expression of two MADS box genes in wild-type and mutant petunia flowers. **Plant Cell**, Rockville, v.4, n.8, p.983-993, Aug. 1992.
- BARROS, R.S.; MAESTRI, M.; COONS, M.P. The physiology of flowering in coffee: a review. **Journal of Coffee Research**, Balehonnur, v.8, n.2/3, p.29-73, 1978.
- BECKER, A.; THEISSEN, G. The major clades of MADS-box genes and their role in the development and evolution of flowering plants. **Molecular Phylogenetics and Evolution**, Orlando, v.29, n.3, p.464-489, Dec. 2003.
- BLAZQUEZ, M.A. Flower development pathways. **Journal of Cell Science**, London, v.113, n.20, p.3547-3548, Oct. 2000.
- BRIDSON, D.M.; VERDCOURT, B. **Flora of tropical East Africa: rubiaceae**. Town: Museums of Cape, 1988. 747p.
- CAMARGO, A.P. Florescimento e frutificação do café arábica nas diferentes regiões cafeeiras do Brasil. **Pesquisa Agropecuária Brasileira**, Brasília, v.20, n.7, p.831-839, jul. 1985.
- CAMARGO, A.P.; CAMARGO, M.B.P. Definição e esquematização das fases fenológicas do cafeeiro arábica nas condições tropicais do Brasil. **Bragantia**, Campinas, v.60, n.1, p.65-68, jan. 2001.

CASTILLEJO, C.; ROMERA-BRANCHA, M.; PELAZ, S. A new role of the *Arabidopsis SEPALLATA3* gene revealed by its constitutive expression. **Plant Journal**, Oxford, v.43, n.4, p.586-596, Aug. 2005.

CAUSIER, B.; KIEFFER, M.; DAVIES, B. MADS-box genes reach maturity. **Science**, New York, v.296, n.5566, p.275-276, Apr. 2002.

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

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

COEN, E.S.; MEYEROWITZ, E.M. The war of the whorls: genetic interactions controlling flower development. **Nature**, London, v.353, n.5, p.31-37, Sept. 1991.

COMPANHIA NACIONAL DE ABASTECIMENTO. **Segunda estimativa: maio 2009**. Disponível em: <http://www.conab.gov.br/conabweb/download/safra/2cafe_09.pdf>. Acesso em: 3 jun. 2009.

CONAGIN, C.H.T.M.; MENDES, A.J.T. Pesquisas citológicas e genéticas em três espécies de *Coffea*: auto-incompatibilidade em *Coffea canephora*. **Bragantia**, Campinas, v.20, n.2, p.787-804, set. 1961.

CROSS, J. **Implications phylogénétiques des variations de l'ADN chloroplastique chez les caféiers (genres *Coffea* L. Et *Psilanthus* Hook. F.)**. 1994. 60f. Dissertação (Mestrado em Melhoramento Genético e Vegetal)-Universite Montpellier II, Montpellier.

CSEKE, L.J.; CSEKE, S.B.; RAVINDER, N.; TAYLOR, L.C.; SHANKAR, A.; SEN, B.; THAKUR, R.; KARNOSKY, D.F.; PODILA, G.K. SEP-class genes in *Populus tremuloides* and their likely role in reproductive survival of poplar trees. **Gene**, Amsterdam, v.358, n.26, p.1-16, Sept. 2005.

DAVIES, B.; EGEE-CORTINES, M.; SILVA, E.A. de; SAEDLER, H.; SOMMER, H. Multiple interactions amongst floral homeotic MADS-box proteins. **EMBO Journal**, Oxford, v.15, n.16, p.4330-4343, Aug. 1996.

DITTA, G.; PINYOPICH, A.; ROBLES, P.; PELAZ, S.; YANOFSKY, M.F. The SEP4 gene of *Arabidopsis thaliana* functions in floral organ and meristem identity. **Current Biology**, London, v.14, n.21, p.1935-1940, Nov. 2004.

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, n.2, p.255-266, Apr. 2005.

EGEA-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*. **EMBO Journal**, Oxford, v.18, n.19, p.5370-5379, Oct. 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**, Oxford, v.12, n.5, p.999-1010, Nov. 1997.

FAVARO, R. MADS-box protein complexes control carpel and ovule development in *Arabidopsis*. **The Plant Cell**, Rockville, v.15, n.11, p.2603-2611, Oct. 2003.

FERRARIO, S.; IMMINK, R.G.H.; SHCHENNIKOVA, A.; BUSSCHER-LANGE, J.; ANGENENT, G.C. The MADS box gene FBP2 is required for SEPALLATA function in petunia. **The Plant Cell**, Rockville, v.15, n.14, p.914-925, Mar. 2003.

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

GOETHE, J.W. **Versuch die metamorphose der pflanzen zu erklären**. Gotha: C.W. Ettinger, 1984. No page. (Acta humaniosa der Verlag Chemie GmbH).

GONÇALVES, W.; SILVAROLLA, M.B. Nematóides parasitos do cafeeiro. In: ZAMBOLIM, L. (Ed.). **Tecnologias de produção de café com qualidade**. Viçosa, MG: UFV, 2001. p.199-267.

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

HAUGHN, G.W.; SOMERVILLE, C.R. Genetic control of morphogenesis in *Arabidopsis*. **Developmental Genetics**, New York, v.9, n.2, p.73-89, Feb. 1988.

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

HONMA, T.; GOTO, K. Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. **Nature**, London, v.409, n.6819, p.525-529, Jan. 2001.

JANG, S.; AN, K.; LEE, S.; AN, G. Characterization of tobacco MADS-box genes involved in floral initiation. **Plant and Cell Physiology**, Kyoto, v.43, n.2, p.230-238, Feb. 2002.

KANG, H.G.; AN, G. Isolation and characterization of a rice MADS box gene belonging to the *AGL2* gene family. **Molecular Cell**, Cambridge, v.7, n.1, p.559-566, Feb. 1997.

KRIZEK, B.A.; FLETCHER, J.C. Molecular mechanisms of flower development: an armchair guide. **Nature**, London, v.6, n.9, p.688-698, Sept. 2005.

KOTILAINEN, M.; ELOMAA, P.; UIMARI, A.; ALBERT, V.A.; YU, D.; TEERI, T.H. *GRCD1*, an *AGL2*-like MADS box gene, participates in the C function during stamen development in *Gerbera hybrida*. **Plant Cell**, Rockville, v.12, n.10, p.1893-1902, Oct. 2000.

KUMAR, D. Some aspects of the hysiology of *Coffea arabica* L. **Kenya Coffee**, Nairobi, v.44, n.519, p.9-47, 1979.

LAMB, P.; MCKNIGHT, S.L. Diversity and specificity in transcriptional regulation: the benefits of heterotypic dimerization. **Trends in Biochemical Science**, Amsterdam, v.16, n.11, p.417-422, Nov. 1991.

LARCHER, W. **Ecofisiologia vegetal**. São Paulo: Rima Artes e Textos, 2000. 531p.

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

LEMMETYINEN, J.; HASSINEN, M.; ELO, A.; PORALI, I.; KEINONEN, K.; MAKELA, H.; SOPANEN, T. Functional characterization of SEPALLATA3 and AGAMOUS orthologues in silver birch. **Physiologia Plantarum**, Copenhagen, v.121, n.1, p.149-162, May 2004.

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

MAESTRI, M.; BARROS, R.S. **Coffee**. New York: Academic, 1997. 278p.

MALCOMBER, S.T.; KELLOGG, E.A. SEPALLATA gene diversification: brave new whorls. **Trends in Plant Science**, Oxford, v.10, n.9, p.427-435, Sept. 2005.

MES, M.G. Studies on the flowering of *Coffea arabica* L. Various phenomena associated with the dormancy of the coffee flower buds. **Portugaliae Acta Biologica**, Lisboa, v.5, n.1, p.25-44, 1957.

MORAIS, H.; CARAMORI, P.H.; KOGUISHI, M.S.; RIBEIRO, A.M.A. Escala fenológica detalhada da fase reprodutiva de *coffea arabica*. **Bragantia**, Campinas, v.67, n.1, p.257-260, jan. 2008.

NORMAN, C.; RUNSWICK, M.; POLLOCK, R.; TREISMAN, R. Isolation and properties of cDNA clones encoding SRF, a transcription factor that binds to the c-fos serum response element. **Cell**, Cambridge, v.55, n.6, p.989-1003, Dec. 1988.

PARENICOVA, L.; FOLTER, S. de; 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. **Plant Cell**, Rockville, v.15, n.7, p.1538-1551, June 2003.

PELAZ, S.; DITTA, G.S.; BAUMANN, E.; WISMAN, E.; YANOFSKY, M.F. B and C floral organ identity functions require SEPALLATA MADS-box genes. **Nature**, London, v.405, n.6783, p.200-203, May 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**, Oxford, v.26, n.4, p.385-394, May 2001.

PNUELI, L.; HAREVEN, D.; BRODAY, L.; LIFSCHITZ, E. The *TM5* [*tdr5*] MADS-box gene mediates organ differentiation in the three inner whorls of tomato flowers. **Plant Cell**, Rockville, v.6, n.2, p.175-186, Feb. 1994.

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.

RENA, A.B.; MAESTRI, M. **Ecofisiologia do cafeeiro**: ecofisiologia da produção agrícola. Piracicaba: Associação Brasileira para Pesquisa da Potassa e do Fosfato, 1987. 147p.

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

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

RIECHMANN, J.; KRIZEK, B.; MEYEROWITZ, E. Dimerization specificity of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v.93, n.10, p.4793-4798, May 1996.

RIECHMANN, J.L.; MEYEROWITZ, E.M. MADS domain proteins in plant development. **Biological Chemistry**, Baltimore, v.378, n.9, p.1079-1101, Sept. 1997.

ROBLES, P.; PELAZ, S. Flower and fruit development in *Arabidopsis thaliana*. **The International Journal of Developmental Biology**, Washington, v.49, n.5, p.633-643, July 2005.

SCHWARZ-SOMMER, Z.; HUE, I.; HUIJSER, R.; FLOR, R.J.; HANSEN, R.; TETENS, E.; LÖNNIG, W.; SAEDLER, H.; SOMMER, H. Characterisation of the *Antirrhinum* floral homeotic MADS-box gene *deficiens*: evidence for DNA binding and autoregulation of its persistent expression throughout flower development. **EMBO Journal**, Oxford, v.11, n.1, p.251-263, Jan. 1992.

SCHWARZ-SOMMER, Z.; HUIJER, P.; NACKEN, W.; SAEDLER, H.; SOMMER, H. Genetic control of flower development by homeotic genes in *Antirrhinum majus*. **Science**, New York, v.250, n.4983, p.931-936, Nov. 1990.

SHITSUKAWA, N.; TAHIRA, C.; KASSAI, K.I.; HIRABAYASHI, C.; SHIMIZU, T.; TAKUMI, S.; MOCHIDA, K.; KAWAURA, K.; OGIHARA, Y.; MURAI, K. Genetic and epigenetic alteration among three homoeologous genes of a class E MADS box gene in hexaploid wheat. **Plant Cell**, Rockville, v.19, n.6, p.1723-1737, June 2007.

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

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

SOARES, A.R.; MANTOVANI, E.C.; RENA, A.B.; SOARES, A.A. Irrigation and physiology of mature coffee blooming in the Zona da Mata region (Viçosa) - state of Minas Gerais - Brazil. **Acta Scientiarum Agronomy**, Maringá, v.27, n.1, p.117-125, jan./fev. 2005.

SOMMER, H.; BELTÁN, J.P.; HUIJSER, P.; PAPE, H.; LONNIG, S.H. Deficiens, a homeotic gene involved in the control of flower morphogenesis in *Antirrhinum majus*: the protein shows homology to transcription factors. **EMBO Journal**, Oxford, v.9, n.3, p.605-613, Mar. 1990.

SUNG, S.K.; MOON, Y.H.; CHUNG, J.E.; LEE, S.Y.; PARK, H.G.; AN, G. Characterization of MADS box genes from hot pepper. **Molecular Cell**, Cambridge, v.11, n.3, p.352-359, June 2001.

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

THOMAZIELLO, R.A.; FAZUOLI, L.C.; PEZZOPANE, J.R.M.; FAHL, J.I.; CARELLI, M.L.C. **Café arábica**: cultura e técnicas de produção. Campinas: Instituto Agrônômico, 2000. 82p. (Boletim Técnico, 187).

TZENG, T.Y.; HSIAO, C.C.; CHI, P.J.; YANG, C.H. Two lily SEPALLATA-like genes cause different effects on floral formation and floral transition in *Arabidopsis*. **Plant Physiology**, Washington, v.133, n.3, p.1091-1101, Oct. 2003.

UIMARI, A.; KOTILAINEN, M.; ELOMAA, P.; YU, D.Y.; ALBERT, V.A.; TEERI, T.H. Integration of reproductive meristem fates by a SEPALLATA-like MADS-box gene. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v.101, n.44, p.15817-15822, Oct. 2004.

VREBALOV, J.; RUEZINSKY, D.; PADMANABHAN, V.; WHITE, R.; MEDRANO, D.; DRAKE, R.; SCHUCH, W.; GIOVANNONI, J. A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (*rin*) locus. **Science**, New York, v.296, n.5566, p.343-346, Apr. 2002.

WINTGENS, J.N. **Coffee**: growing, processing, sustainable production. Weinheim: Verlag, 2004. 24p.

XU, Y.; ZHANG, L.; XIE, H.; ZHANG, Y.Q.; OLIVEIRA, M.M.; MA, R.C. Expression analysis and genetic mapping of three SEPALLATA-like genes from peach (*Prunus persica* (L.) Batsch). **Tree Genetics & Genomes**, Heidelberg, v.4, n.4, p.693-703, Oct. 2008.

YANOFSKY, M.; MA, H.; BOWMAN, I.; DREWS, G.; FELDMANN, K.; MEYEROWITZ, E. The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors. **Nature**, London, v.346, n.6279, p.35-40, July 1990.

ZHANG, H.; FORDE, B.G. An *Arabidopsis* MADS box gene that controls nutrient-induced changes in root architecture. **Science**, New York, v.279, n.5349, p.407-409, Jan. 1998.

ZHAO, X.Y.; CHENG, Z.J.; ZHANG, X.S. Overexpression of TaMADS1, a SEPALLATA-like gene in wheat, causes early flowering and the abnormal development of floral organs in *Arabidopsis*. **Planta**, Berlin, v.223, n.4, p.698-707, Apr. 2006.

CAPÍTULO 2

IDENTIFICAÇÃO *IN SILICO* DE GENES *SEPALLATA* EM *Coffea arabica*

RESUMO

Os genes da subfamília *SEPALLATA* (*SEP*) são fatores de transcrição pertencentes à família *MADS-box*, que contribuem para a determinação do meristema e formação dos quatro verticilos florais. Esse trabalho teve como objetivo identificar os genes *SEP* presente em *Coffea arabica* através da 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 *SEP*. Mais de 1000 *reads* foram encontrados formando 41 *EST-contigs* e desses apenas os que apresentaram similaridade com *SEP* de outras espécies e pelo menos presença do domínio *MADS* foram utilizados no trabalho, resultando em somente três prováveis genes *SEP* de café. Duas dessas sequências apresentaram o domínio *MIKC*, com a parte *MIK* bastante conservada e o domínio C variável, característica comum entre as proteínas *SEP* de outras espécies. A análise filogenética apresentou uma separação clara entre as prováveis proteínas *SEP* de café e de outras espécies com outras proteínas *MADS-box*. Quando as proteínas *SEP* foram analisadas separadamente, as prováveis sequências de café foram agrupadas em três grupos distintos, formados de proteínas caracterizadas como *SEP1/2*, *SEP3* e *SEP4*. Através do perfil de expressão dado pelo *Northern* eletrônico, foi possível visualizar a expressão dessas sequências em tecidos de flores e frutos em diferentes estádios de desenvolvimento, fornecendo mais uma evidência que as sequências identificadas *in silico* são prováveis *SEPALLATA* de café.

ABSTRACT

The genes of the subfamily *SEPALLATA* (*SEP*) are transcription factors that belong to the MADS-box family, that contribute to the meristem determination and the formation of four floral whorls. The aim of this work was to identify the *SEP* gene in *Coffea arabica* by the Coffee transcriptome database (CAFEST). By the Gene Project interface, it was possible to search for reads, to assemble clusters and analyze the putative *SEP* genes. More than 1000 *reads* were found generating 41 *EST-contigs*, from which only those presenting similarity with other *SEP* genes from different species and at least the presence of one MADS domain were used in this work, resulting in only three putative coffee *SEP*. Two of these sequences presented the MIKC domain, with the MIK region conserved and the C domain variable, a very common characteristic between other *SEP* proteins. The phylogenetic analysis showed a clear division between the putative coffee *SEP* proteins and from other species with other *MADS-box* proteins. When the *SEP* proteins were analyzed separately, coffee sequences were grouped in three different clusters, formed by proteins characterized as *SEP1/2*, *SEP3* and *SEP4*. By the expression profile using the electronic *Northern*, it was possible to check the expression of these sequences in flower and fruits tissues in different stages of development, providing another evidence that the sequences obtained by the *in silico* analysis are putative *SEPALLATA* coffee genes.

1 INTRODUÇÃO

A família de genes *MADS-box* de fatores de transcrição está envolvida no desenvolvimento celular de leveduras, plantas e animais. É dividida em dois tipos, sendo os genes *MADS* que participam do desenvolvimento vegetal, do tipo II (Alvarez-Buylla et al., 2000). As proteínas originadas desses genes apresentam uma organização estrutural conservada, denominada domínio estrutural tipo-*MIKC*, que inclui quatro domínios funcionais diferentes, distribuídos no sentido N a C terminal: o domínio *MADS-box* (M), a região intermediária (I), queratina-*like* (K), e carboxi terminal (C) (Hayes et al., 1988; Purugganan et al., 1995; Shore & Sharrocks, 1995; Riechmann et al., 1996).

Alguns genes *MADS-box* fazem parte do modelo ABCE que foi proposto para explicar a formação dos órgãos florais. De acordo com esse modelo, a interação de quatro classes de genes homeóticos é necessária para a determinação floral. Para a formação das sépalas, é necessário a expressão de genes pertencentes à função A e E. Os genes das funções A, B e E são requeridos para a formação das pétalas. A combinação dos genes B, C e E para a formação dos estames e a junção dos genes C e E para a especificação do carpelo.

Os genes que fazem parte da função E são conhecidos como genes *SEPALLATA*, e são necessários para a formação de todos os verticilos. Em *Arabidopsis thaliana* são caracterizados quatro desses genes: *SEP1*, *SEP2*, *SEP3* e *SEP4*, que possuem função redundante na determinação floral. Em estudos com mutantes simples, as plantas obtidas possuem fenótipos sutis comparados ao selvagem, entretanto quando são formados mutantes triplos ocorre a mudança fenotípica, desenvolvendo flores com todos os verticilos transformados em sépalas. Além disso, quando são formados mutantes quádruplos, todos os verticilos são transformados em folhas (Pelaz et al., 2000; Honma & Goto, 2001;

Ditta et al., 2004). Mas como é estabelecido pelo modelo ABCE, eles não são requeridos sozinhos, cada proteína *SEP* forma um complexo multimérico com outros fatores de transcrição *MADS-box* para formar os verticilos florais e contribuir no desenvolvimento dos frutos.

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.

Devido à importância dos genes da classe E na formação dos verticilos florais e no desenvolvimento dos frutos, o objetivo desse estudo foi encontrar prováveis proteínas *SEP* de *Coffea arabica* no banco de dados *CAFEST*, identificando *in silico* os genes, seus domínios conservados, as bibliotecas em que se expressa e a relação deles com *SEP* de outras espécies.

2 MATERIAL E MÉTODOS

2.1 Identificação de sequências *SEPALLATA*

As sequências *SEP* (*SEPALLATA*) de café foram identificadas a partir do banco de dados de *ESTs* (*Expressed Sequence Tags*) gerado pelo projeto Genoma Brasileiro Café - *CAFEST* (Vieira et al., 2006). por meio do qual foi possível a procura por *reads*, montagem de *clusters* e análise dos prováveis genes.

Primeiramente foi realizada a procura no banco de dados *CAFEST* por palavras chaves, tendo-se em vista que todos os *reads* haviam sido anotados automaticamente por comparação com o banco de dados do *NCBI* (*National Center for Biotechnology Information*). Posteriormente, foi realizada a procura utilizando a ferramenta *BLAST* (*Basic Local Alignment Search*) com sequências obtidas pelo *NCBI* de genes *SEP* de outras espécies já identificados e publicados. 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 ($e\text{-value} < 10^{-5}$).

Todos os *reads* encontrados no banco de dados *CAFEST*, foram agrupados, formando os chamados *EST-contigs*, e submetidos à anotação manual. Todos esses *EST-contigs* foram utilizados para uma terceira busca no banco *CAFEST* utilizando-se a ferramenta *BLASTN*, visando-se encontrar novas sequências. Esse processo, também chamado de saturação, foi repetido até que não se encontrasse mais nenhum novo *read* significativo. Após todas as montagens, os possíveis *contigs SEP* foram submetidos à confirmação utilizando a ferramenta *BLAST* do *NCBI*, selecionando para as análises subsequentes, somente aqueles que possuíam similaridade com *SEP* de outras espécies e domínios completos.

2.2 Análise filogenética

O alinhamento dos *EST-contigs* foi realizado pelo programa *ClustalW* (Thompson et al., 1994)) com parâmetros padrões (default), utilizando-se as sequências de nucleotídeos traduzidas em aminoácidos. A árvore filogenética 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* pode ser medida pelo teste probabilístico de *bootstraps* (Sitnikova et al., 1995). Foram construídas duas árvores filogenéticas, uma comparando as sequências SEP identificadas de café e de outras espécies com diversas proteínas *MADS-box*. A outra somente com a comparação das prováveis proteínas SEP de café com as de outras espécies.

2.3 Identificação de *motifs* comuns entre as sequências

Para descobrir motivos de agrupamento entre as sequências SEP selecionadas no *CAFEST*, utilizou-se o programa *MEME* (*Multiple Expectation Minimization for Motif Elicitation*) 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* 6 e amplitude ótima entre 6 e 200 (Parenicova et al., 2003).

2.4 Northern eletrônico

Para a análise *in silico* da expressão, foram calculadas as frequências de *reads* formadores de cada *EST-contig* nas bibliotecas em que apareciam expressos. As bibliotecas analisadas foram FB1/FB2/FB3/FB4, botões florais em subsequentes estágios do florescimento e FR2, frutos. As bibliotecas

utilizadas foram descritas por Vieira et al. (2006). Os dados de expressão dos *reads* nas bibliotecas foram previamente normalizados para dar uma idéia exata do grau de expressão dos prováveis genes em cada tratamento e local da planta, já que, as bibliotecas formadas não possuíam o mesmo tamanho. 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. Com os dados foi feita uma matriz, relacionando os *reads* formadores de cada *contig* com as bibliotecas, posteriormente lançada nos programas Cluster e TreeView (Eisen et al., 1998). Os *EST-contigs* e bibliotecas foram agrupados por *hierachical clustering* e os resultados de expressão foram apresentados em uma escala de cinza, onde expressão zero ou negativa foi representada por coloração mais clara sendo aumentada gradativamente até atingir um cinza escuro, que representava o grau máximo de expressão positiva.

3 RESULTADOS E DISCUSSÃO

As sequências *SEP* de café foram identificadas através do banco de dados *CAFEST*, e encontrados aproximadamente 790 *reads*, realizando a busca por palavra chave e por sequências *SEP* de outras espécies depositadas no *NCBI*. A primeira clusterização (agrupamento) gerou 20 *EST-contigs*. Após o processo de saturação, com o objetivo de realizar novas buscas para corrigir *clusters* incompletos, foi gerado uma quantidade maior de *reads*, montados em 41 *EST-contigs*. Para cada possível gene *SEP* de café formado, foi analisada sua sequência protéica, a parte funcional de cada uma delas e seus domínios conservados. Entretanto, as sequências foram identificadas como possíveis *EST-contigs* verdadeiros, através da realização de um alinhamento e comparação com as sequências disponíveis no *NCBI*.

Após essas análises, alguns prováveis genes *SEP* não apresentaram similaridade com genes *SEP* de outras espécies, e outros não possuíam nem o domínio *MADS*. Desta forma, essas sequências foram consideradas incompletas e não foram utilizadas em análises posteriores. Dos 41 *EST-contigs*, apenas três foram mantidos na análise por apresentarem domínios completos e alta similaridade com proteínas *SEP* de outras espécies. Esse fato pode ser explicado por erros na metodologia aplicada aos processos de sequenciamento do banco de dados, isolamento dos genes expressos na construção das bibliotecas, e ineficiência dos métodos de seleção de *reads*, o que gerou *contigs* incompletos que continham apenas pequenas sequências conservadas dos genes da família *MADS-box*.

Os três *contigs* mantidos na análise, foram avaliados quanto à homologia, domínios, motivos e expressão em tecidos de café.

Em relação à homologia destes *contigs* com as proteínas SEP já descritas, pode ser observado na Tabela 1 um início de identificação desse *contigs* com os diferentes tipos de proteínas SEP, utilizando o *tblastn* do NCBI.

TABELA 1 Resultado do *tblastn* do NCBI, com os *contigs* de café identificados.

	Sequência	Tamanho (aa)	Identidade	e-value
Contig1	<i>Citrus</i> (CiSEP1)	433	90%	6e-41
108 aa	<i>Vitis vinifera</i>	431	87%	1e-40
	<i>Arabidopsis thaliana</i> (SEP1)	420	88%	2e-39
Contig14	<i>Populus tremuloides</i> (SEP3)	550	80%	9e-108
241 aa	<i>Syringa vulgaris</i> (SEP3)	689	84%	4e-112
Contig8	<i>Solanum lycopersicum</i> (rin)	733	78%	2e-75
279 aa	<i>Arabidopsis thaliana</i> (SEP4)	643	84%	1e-64

Utilizando a ferramenta *Blastp* e *Conserved domains* do NCBI, foram encontrados os domínios *MADS* e *K-box*, representantes das regiões mais conservadas do domínio *MIKC* da família *MADS-box* (Figura 1).

A identificação dos motivos de agrupamento das prováveis proteínas SEP de café, foi realizada por meio do alinhamento com proteínas SEP de outras espécies. Como pode ser observado na figura 2, seis motivos foram identificados em quase todas as sequências, representando os domínios conservados da família *MADS-box*. Segundo análise no programa *Conserved domains* do NCBI, o motivo 1 representa o domínio *MADS* e o motivo 2 o domínio *K-box*, que são os domínios mais conservados entre as espécies. No motivo 1, pode ser encontrado um potencial sítio de fosforilação, que possui a seguinte sequência de aminoácidos: *RQVTF*, encontrado no domínio *MADS* de todos os genes

MADS-box (Zahn et al., 2005; Kanno et al., 2006; Huang et al., 2009) (Figura 3). A partir de resultados similares obtidos por Malcomber & Kellogg (2005) e relacionando as figuras 2 e 3, o motivo 3 foi identificado como pertencente ao domínio I, e os motivos 4, 5 e 6 ao domínio C, que são domínios menos conservados. Como esses motivos representam os domínios característicos da família *MADS-box*, chamados de domínios *MIKC*, essa pode ser uma evidência que as sequências encontradas de café são pertencentes à família.

Os prováveis *contigs* representantes do gene *SEP* em café apresentaram todos os motivos e domínios característicos, entretanto o *contig1* mostrou-se incompleto, diferenciando pela falta do domínio conservado *K-box* e pela região C-terminal. Trabalhos posteriores devem ser realizados, a fim de tentar remontar este *contig*.

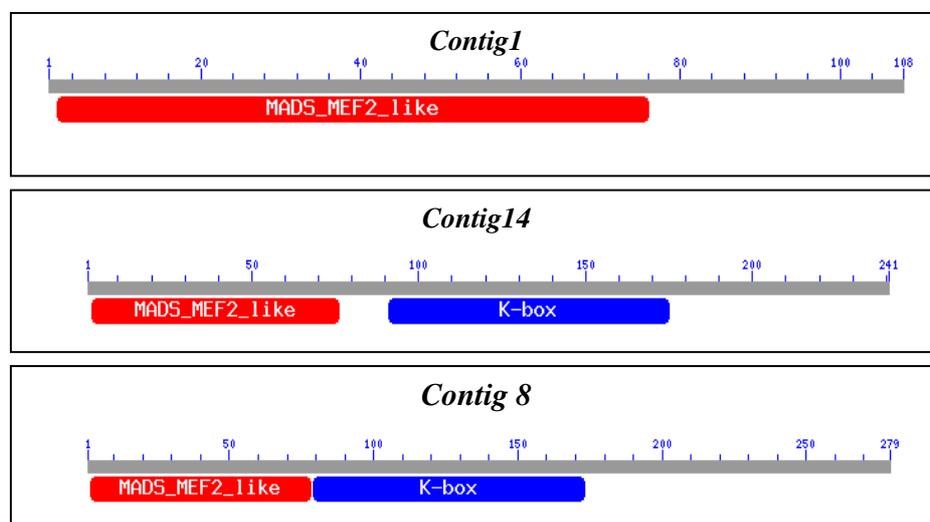


FIGURA 1 Representação esquemática dos domínios das supostas sequências *SEP* de café, utilizando a ferramenta *Blastp* do *NCBI*.

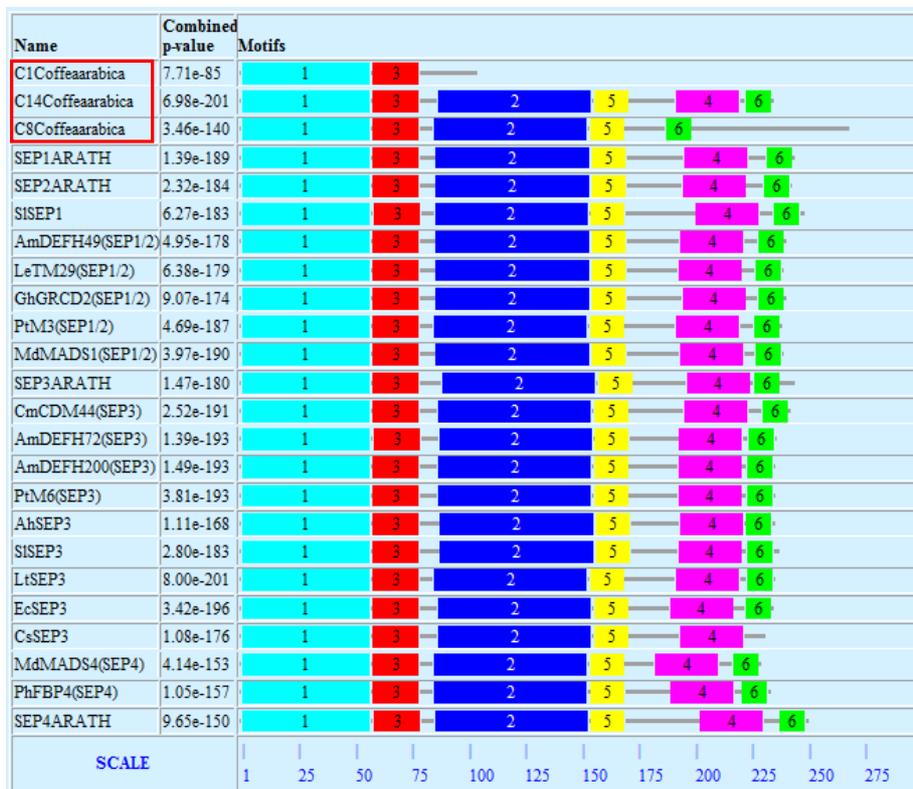


FIGURA 2 Análise dos MOTIF de agrupamentos presentes entre os *contigs* de café avaliados e outras proteínas SEP de espécies já depositadas no NCBI, utilizando o programa MEME. C1, C14 e C8 de *Coffea arabica*; AtSEP1, AtSEP2, AtSEP3, e AtSEP4 de *A. thaliana* (Acesso GenBank P29382, P29384, O22456 e P29383); AmDEFH49, AmDEFH72 e AmDEFH200 de *Antirrhinum majus* (S78015, S71756 e S71757); PhFBP4 de *Petunia hybrida* (AK21247); LeTM29 de *Lycopersicon esculentum* (CAC83066); CmCDM44 de *Chrysanthemum morifolium* (AAO22982); GhGRCD2 de *Gerbera hybrida* (AH04878); PtM3 e PtM6 de *Populus tremuloides* (AAL08423 e AAO49811); MdMADS1 e MdMADS4 de *Malus domestica* (AAC25922, AAD51423); AhSEP3 de *Alpinia hainanensis* (ACO72983); SISEP3 de *Silene latifolia* (BAD10945); LtSEP3 de *Liriodendron tulipifera* (AAX15920); EcSEP3 *Eschscholzia californica* (AAX15918); CsSEP3 *Crocus sativus* (ACB69512); SISEP1 *Silene latifolia* (BAD10944).

Entre as proteínas SEP a constituição de aminoácidos é bem parecida nos domínios *MIK*. Zahn et al. (2005) demonstraram a partir do alinhamento de sequências SEP de várias espécies, que os domínios *MIK* são altamente conservados, e em muitas posições são quase invariáveis. Existem apenas algumas diferenças na porção C-terminal, principalmente na região do domínio C, entretanto, são encontrados nessa região pequenos motivos altamente conservados (Kanno et al., 2006), que são representados como motivo 4 e 6 (Figura 3). Resultados similares foram encontrados em outros trabalhos, onde foram identificados os mesmos motivos comuns no domínio C (Malcomber & Kellogg, 2005; Kanno et al., 2006; Zhao et al., 2006; Xu et al., 2008).

Como esses motivos são conservados entre as diferentes espécies, eles podem desempenhar um papel importante na especificidade da formação do complexo de ligação (Kanno et al., 2006), que é fundamental para a função das proteínas SEP. O motivo 4 é mais conservado que o motivo 6, sugerindo que esse pode ter funções adicionais e aquele ser envolvido nas funções originais dos genes da subfamília (Vandenbussche et al., 2003). Para tentar caracterizar a função desses motivos, Melzer et al. (2009), demonstraram que a deleção do domínio C-terminal não prejudica a ligação do SEP3 ao DNA alvo ou a interação entre as proteínas. Eles também dividiram o domínio K em três partes distintas, onde foi observado que quando modificavam o subdomínio K3, conduziam a uma clara redução da ligação, determinando a sua importância. Isso indica que os domínios M, I, K são suficientes para determinar essa função. Os motivos conservados do domínio C, provavelmente desempenham um papel central na subfamília SEP (Lamb & Irish, 2003; Vandenbussche et al., 2004) porém, sua função específica ainda não é conhecida (Geuten et al., 2006). Zahn et al. (2005) determinaram que os motivos 4 e 6 possuem estruturalmente aminoácidos hidrofóbicos e polares e não assemelham a nenhum motivo com

função conhecida, entretanto eles podem ser funcionalmente importantes porque são localizados em porções similares a motivos do gene *AG* e *AP3*.

Ao analisar as diferenças presentes dentro desses motivos encontrados na região C-terminal (Figura 3) pode ser observada uma pequena variação característica dentro do motivo 4 (Figura 3) entre as sequências SEP1/2, SEP3 e SEP4. A maior parte das proteínas denominadas SEP1/2 possuem variações da sequência NPTLQ; na mesma posição, as proteínas SEP3 são variáveis na sequência EPTL; e as menos conservadas entre si, as SEP4 possuem variações na sequência NN-L, porém não muito constante. Essas variações nas proteínas SEP4 podem estar relacionadas a modificações na função. Variações similares foram encontradas por Malcomber & Kellogg (2005). Entretanto, isso não é observado entre as sequências SEP1 e SEP2 das várias espécies analisadas, pois possuem a constituição de aminoácidos muito parecida na região C-terminal, desta forma são denominadas SEP1/2, sendo diferencialmente identificadas somente em *Arabidopsis thaliana* (Pelaz et al., 2000).

Pela análise filogenética pode ser observada a homologia das sequências estudadas com proteínas SEP identificadas em outras espécies. A figura 4 demonstra o agrupamento esperado entre as proteínas SEP e outras proteínas *MADS-box*, formando dois grandes grupos distintos. Essa separação indica a diferenciação entre as proteínas da sub-família SEP das demais proteínas *MADS-box*, com altos valores de *bootstrap* (100), indicando que os *contigs* encontrados no genoma café são provavelmente homólogos a proteínas SEP.

A árvore filogenética formada entre proteínas SEP de outras espécies e das prováveis sequências de café, também indica essa homologia, determinando uma nítida separação entre as diferentes proteínas SEP (Figura 5). Essa separação pode ser determinada pelas variações constantes observadas na região C-terminal das sequências SEP, podendo indicar uma variação da função entre as espécies. Malcomber & Kellogg (2005), além de demonstrar que os motivos

encontrados na região C-terminal possui uma constante diferenciação entre as proteínas SEP, também propuseram essa região como o controle central na ativação da transcrição, ligação específica ao DNA, e localização sub-celular, porém a função ainda não está totalmente elucidada como discutido anteriormente. Entretanto, a falta dessa sequência não prejudicou no agrupamento do *contig1* com as proteínas SEP1/2 de outras espécies, assim como nas outras análises, indicando que os domínios M e I, podem ser suficientes para determinar qual classe de genes *SEP* elas pertencem.

De acordo com a figura 5, podemos agrupar o *contig1* com proteínas descritas na literatura como SEP1 e SEP2, o *contig14* com proteínas SEP3 e o *contig8* com proteínas caracterizadas como SEP4. Como a SEP1 e SEP2 mostram-se muito parecidas, Ermolaeva et al. (2003) indicou que esses genes recentemente sofreram uma duplicação na evolução do florescimento de plantas, mas a relação para o SEP3 e SEP4 não está clara. No agrupamento proposto não foi possível determinar diferenças entre esses SEP1 e SEP2, mas eles formaram um agrupamento mais próximo ao do SEP3 corroborando aos dados encontrados por outros autores (Purugganan, 1998; Lawton-Rauh et al., 2000; Sung et al., 2000; Becker & Theissen, 2003; Parenicova et al., 2003; Vandenbussche et al., 2003; Kanno et al., 2006). Porém, alguns estudos sugerem que o SEP1 e SEP2 são mais ligados ao SEP4 do que ao SEP3 (Yu & Goh, 2000; Lemmetyinen et al., 2004; Zahn et al., 2005), entretanto, a relação filogenética entre esses quatro genes ainda não está totalmente estabelecida (Kanno et al., 2006).

Esse perfil de agrupamento pode servir para prever os tipos de *SEP* que podem ser encontrados em café, porém os prováveis locais de expressão desses genes nos verticilos florais e possíveis funções não podem ser determinados apenas com essas análises, pois como descrito por vários autores, mesmo atuando em algumas espécies de forma redundante em outras, tanto o local de expressão, como as funções são variáveis. Por exemplo, estudos de

expressão do gene *SEP3* em petúnia e tomate, demonstram um rompimento das estruturas florais, entretanto em gérbera foi observado apenas a esterilidade dos estames (Kotilainen et al., 2000; Uimari et al., 2004). Além disso, em algumas plantas estes estudos de expressão modificam o fenótipo, em outras, não é observada nenhuma diferença em comparação à planta selvagem (Cseke et al., 2005; Xu et al., 2008). Desta forma são necessários estudos de expressão homeótica com as sequências identificadas para determinar a possível função de cada uma.

As funções das proteínas SEP também podem ser determinadas pelo tipo de ligação que elas estabelecem com outras proteínas *MADS-box*, e sabe-se que a maior parte desses complexos protéicos contém a proteína SEP3, que trabalha como importante mediador da formação dessas ligações (Immink et al., 2009). Dessa forma, o *contig14* de café que apresentou homologia com as proteínas SEP3 de outras espécies, pode ter inicialmente uma função predita como sendo uma proteína chave nesses complexos.

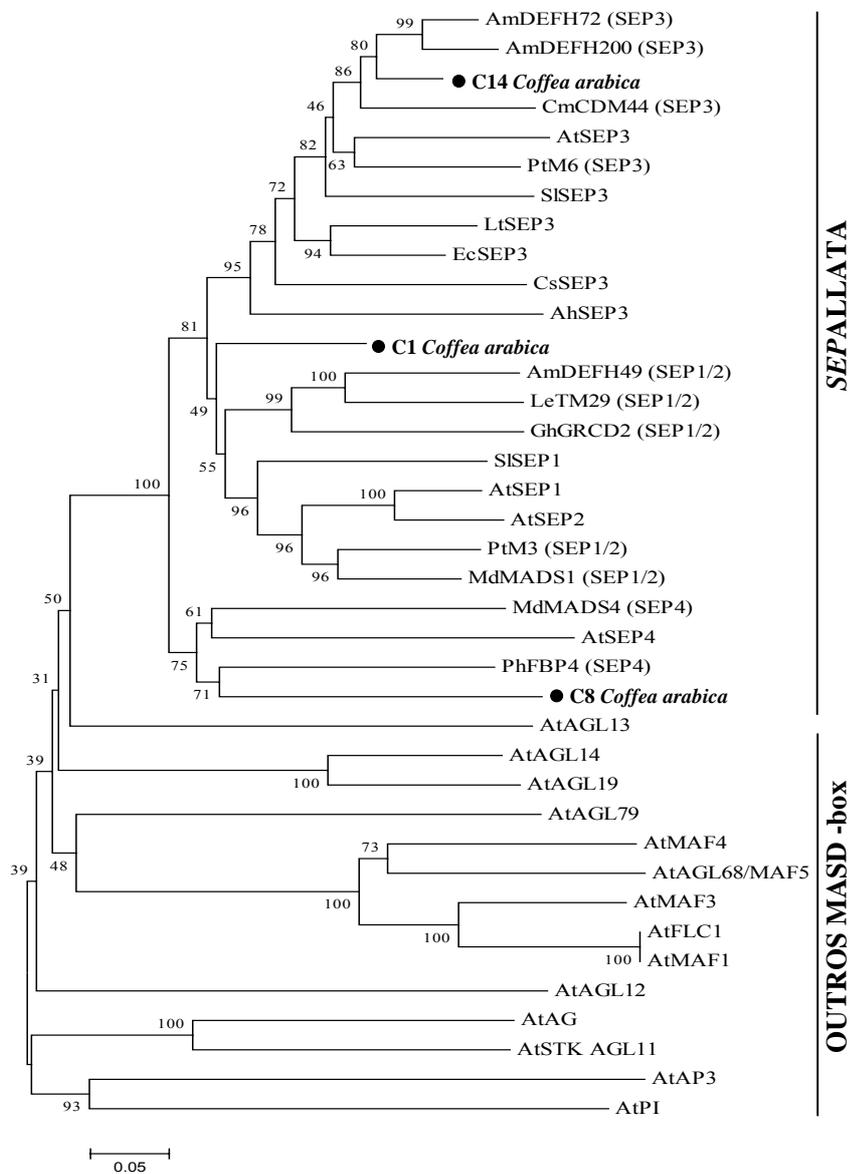


FIGURA 4 Relação filogenética entre as seqüências de aminoácidos das proteínas SEPALLATA com outras proteínas *MADS-box* do tipo II, publicadas e depositadas do *NCBI*. C1, C8 e C14 de *Coffea arabica*; *AtSEP1*, *AtSEP2*, *AtSEP3*, e *AtSEP4* de *A. thaliana* (Acesso GenBank P29382, P29384, (... Continuação...))

Figura 4, Cont.

O22456 e P29383); AmDEFH49, AmDEFH72 e AmDEFH200 de *Antirrhinum majus* (S78015, S71756 e S71757); PhFBP4 de *Petunia hybrida* (AK21247); LeTM29 de *Lycopersicon esculentum* (CAC83066); CmCDM44 de *Chrysanthemum morifolium* (AAO22982); GhGRCD2 de *Gerbera hybrida* (AH04878); PtM3 e PtM6 de *Populus tremuloides* (AAL08423 e AAO49811); MdMADS1 e MdMADS4 de *Malus domestica* (AAC25922, AAD51423); AhSEP3 de *Alpinia hainanensis* (ACO72983); SlSEP3 de *Silene latifolia* (BAD10945); LtSEP3 de *Liriodendron tulipifera* (AAX15920); EcSEP3 *Eschscholzia californica* (AAX15918); CsSEP3 *Crocus sativus* (ACB69512); SlSEP1 *Silene latifolia* (BAD10944). E como grupo externo outras proteínas *MADS-box* de *Arabidopsis thaliana*: AtAG (NP567569); AtFLC1 (AAK54440); AtMAF4 (NP_001119499); AtMAF3 (NP001032144); AtAGL68/MAF5 (NP001078799); AtMAF1 (NP177833); AtAGL14 (NP192925); AtAGL19 (NP194026); AtAGL13 (NP191671); AtAGL79 (NP189645); AtAGL12 (AAP37691); AtSTK (NP001078364); AtAP3 (NP191002); AtPI (NP197524)

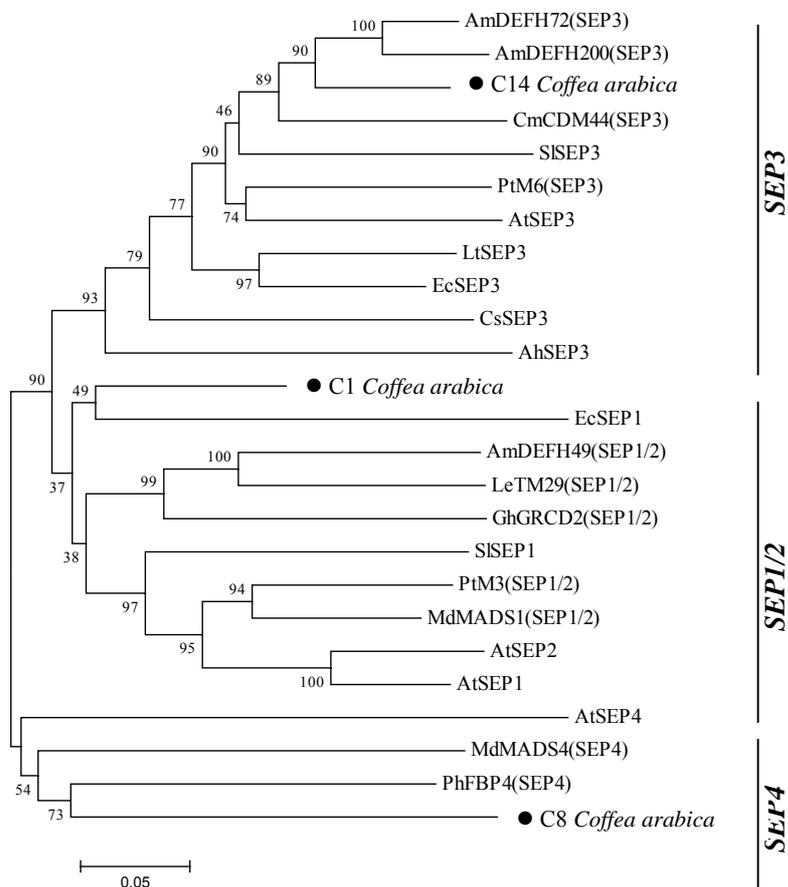


FIGURA 5 Árvore filogenética das prováveis proteínas SEPALLATA encontradas em café com as de outras espécies, publicadas e depositadas no *NCBI*. C1, C8 e C14 de *Coffea arabica*; AtSEP1, AtSEP2, AtSEP3, e AtSEP4 de *A. thaliana* (Acesso GenBank P29382, P29384, O22456 e P29383); AmDEFH49, AmDEFH72 e AmDEFH200 de *Antirrhinum majus* (S78015, S71756 e S71757); PhFBP4 de *Petunia hybrida* (AK21247); LeTM29 de *Lycopersicon esculentum* (CAC83066); CmCDM44 de *Chrysanthemum morifolium* (AAO22982); GhGRCD2 de *Gerbera hybrida* (AH04878); PtM3 e PtM6 de *Populus tremuloides* (AAL08423 e AAO49811); MdMADS1 e MdMADS4 de *Malus domestica* (AAC25922, AAD51423); AhSEP3 de *Alpinia hainanensis* (ACO72983); SISEP3 de *Silene latifolia* (BAD10945); LtSEP3 de *Liriodendron tulipifera* (AAX15920); EcSEP3 *Eschscholzia californica* (AAX15918); CsSEP3 *Crocus sativus* (ACB69512); SISEP1 *Silene latifolia* (BAD10944)

Pela análise do *Northern* eletrônico, foi observada a expressão dos *contigs* nas bibliotecas de botões florais em diferentes estádios de florescimento (FB1, FB2, FB3 e FB4) e frutos em estágio inicial de desenvolvimento (FR2), corroborando assim, com as funções determinadas para genes dessa subfamília, que são responsáveis pela identidade do meristema floral e desenvolvimento dos frutos, sendo dessa forma expressos na flor e fruto. Trabalhos realizados com diferentes espécies apresentaram resultados similares (Kanno et al., 2006; Matsubara et al., 2008), porém não foi possível observar uma expressão padrão entre os diferentes *contigs* analisados.

De acordo com a Figura 6, pode-se observar que o *contig1* apresentou um nível de expressão relativamente baixo nas fases iniciais de desenvolvimento da flor e um aumento na expressão quando é considerada a última fase. Isso pode estar relacionado ao acúmulo desse transcrito nessa fase e é similar aos resultados encontrados por Matsubara et al. (2008) que observou a expressão do gene *SEP* em botões florais e flores em vários estádios de desenvolvimento de petúnia (FBP2), indicando uma maior expressão no último estágio. Porém, não se pode relacionar os estádios de desenvolvimento considerados na construção das bibliotecas no *CAFEST* com os analisados por Matsubara et al. (2008).

Vários trabalhos observaram variações de expressão entre os diferentes verticilos (Malcomber & Kellogg, 2005; Zahn et al., 2005; Geuten et al., 2006). Entretanto essa diferença de expressão entre os quatro verticilos florais não pode ser analisada nesse trabalho, pois na construção das bibliotecas não foram utilizados tecidos individuais e sim, toda a estrutura floral, indicando apenas a expressão total na flor e no fruto

Essas evidências de expressão são mais um indício de que os *contigs* estudados são provavelmente da subfamília *SEP*, mas são necessários estudos de confirmação da expressão tecido específica *in vivo* para comparação com os resultados obtidos em todo o trabalho *in silico*.

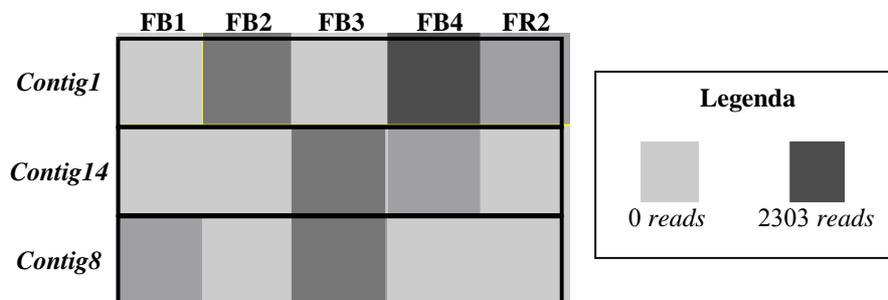


FIGURA 6 *Northern* eletrônico representando os níveis de expressão dos *EST-contigs* nas diferentes bibliotecas por meio de uma escala de vermelho. Bibliotecas: FB1/FB2/FB3/FB4 (botões florais em diferentes estádios do florescimento) e FR2 (frutos em estágio inicial). As bibliotecas utilizadas foram descritas por Vieira et al. (2006).

Por meio das análises *in silico*, os contigs apresentados do genoma café, puderam ser identificados quanto a sua similaridade com as proteínas SEP. O *contig1* foi identificado como SEP1/2 (C1SEP1/2), o *contig14* como SEP3 (C14SEP3), e o *contig8* como SEP4 (C8SEP4).

4 CONCLUSÃO

A partir de uma busca direcional no banco de dados *CAFEST* foi possível identificar três prováveis sequências *SEP*, que apresentaram com as análises de motivos porções determinantes na identidade das proteínas *SEP*.

Com a construção da árvore filogenética, foi possível demonstrar o correto agrupamento dessas sequências com proteínas *SEP* de outras espécies, indicando a homologia esperada, com valores altos de *bootstrap*.

Na análise de expressão pelo *Northern* eletrônico, foi possível observar uma expressão característica desses genes em tecidos de flor e fruto.

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**, Oxford, v.25, n.17, p.3389-3402, Sept. 1997.

ALVAREZ-BUYLLA, E.R.; LILJEGREN, S.J.; PELAZ, S.; GOLD, S.E.; BURGEFF, C.; DITTA, G.S.; VERGARA-SILVA, F.; YANOFSKY, M.F. MADS-box gene evolution beyond flowers: expression in pollen, endosperm, guard cells, roots and trichomes. **Plant Journal**, Oxford, v.24, n.4, p.457-466, Nov. 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. **Proceedings...** Menlo Park: AAAI, 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 Phylogenetics and Evolution**, Orlando, v.29, n.3, p.464-489, Dec. 2003.

CSEKE, L.J.; CSEKE, S.B.; RAVINDER, N.; TAYLOR, L.C.; SHANKAR, A.; SEN, B.; THAKUR, R.; KARNOSKY, D.F.; PODILA, G.K. SEP-class genes in *Populus tremuloides* and their likely role in reproductive survival of poplar trees. **Gene**, Amsterdam, v.358, n.26, p.1-16, Sept. 2005.

DITTA, G.; PINYOPICH, A.; ROBLES, P.; PELAZ, S.; YANOFSKY, M.F. The SEP4 gene of *Arabidopsis thaliana* functions in floral organ and meristem identity. **Current Biology**, London, v.14, n.21, p.1935-1940, Nov. 2004.

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 of the United States of America**, Washington, v.95, n.25, p.14863-14868, Nov. 1998.

ERMOLAEVA, M.D.; WU, M.; EISEN, J.A.; SALZBERG, S.L. The age of the *Arabidopsis thaliana* genome duplication. **Plant Molecular Biology**, Dordrecht, v.51, n.6, p.859-866, Apr. 2003.

GEUTEN, K.; BECKER, A.; KAUFMANN, K.; CARIS, P.; JANSSENS, S.; VIAENE, T.; THEISSEN, G.; SMETS, E. Petaloidy and petal identity MADS-box genes in the balsaminoid genera *Impatiens* and *Marcgravia*. **Plant Journal**, Oxford, v.47, n.4, p.501-518, Aug. 2006.

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

HIGGINS, D.G.; SHARP, P.M. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. **Gene**, Amsterdam, v.73, n.1, p.237-244, Dec. 1988.

HONMA, T.; GOTO, K. Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. **Nature**, London, v.409, n.6819, p.525-529, Jan. 2001.

HUANG, F.; CHI, Y.J.; GAI, J.Y.; YU, D.Y. Identification of transcription factors predominantly expressed in soybean flowers and characterization of GmSEP1 encoding a SEPALLATA1-like protein. **Gene**, Amsterdam, v.438, n.1/2, p.40-48, June 2009.

IMMINK, R.G.H.; TONACO, I.A.N.; FOLTER, S.; SHCHENNIKOVA, A.; DIJK, A.D.J. van; BUSSCHER-LANGE, J.; BORST, J.W.; ANGENENT, G.C. SEPALLATA3: the 'glue' for MADS box transcription factor complex formation. **Genome Biology**, London, v.10, n.2, p.24.1-24.16, Feb. 2009.

KANNO, A.; HIENUKI, H.; ITO, T.; NAKAMURA, T.; FUKUDA, T.; YUN, P.Y.; SONG, I.J.; KAMIMURA, T.; OCHIAI, T.; YOKOYAMA, J.; MAKI, M.; KAMEYA, T. The structure and expression of SEPALLATA-like genes in *Asparagus* species (Asparagaceae). **Sexual Plant Reproduction**, New York, v.19, n.3, p.133-144, Sept. 2006.

KOTILAINEN, M.; ELOMAA, P.; UIMARI, A.; ALBERT, V.A.; YU, D.; TEERI, T.H. *GRCD1*, an *AGL2*-like MADS box gene, participates in the C function during stamen development in *Gerbera hybrida*. **Plant Cell**, Oxford, v.12, n.10, p.1893-1902, Oct. 2000.

LAMB, R.S.; IRISH, V.F. Functional divergence within the APETALA3-PISTILLATA floral homeotic gene lineages. **Proceedings National Academy of Science of the United States of America**, Washington, v.100, n.11, p.6558-6563, May 2003.

LAWTON-RAUH, A.L.; ALVAREZ-BUYLLA, E.R.; PURUGGANAN, M.D. Molecular evolution of flower development. **Trends in Ecology & Evolution**, Amsterdam, v.15, n.4, p.144-149, Apr. 2000.

LEMMETYINEN, J.; HASSINEN, M.; ELO, A.; PORALI, I.; KEINONEN, K.; MAKELA, H.; SOPANEN, T. Functional characterization of SEPALLATA3 and AGAMOUS orthologues in silver birch. **Physiologia Plantarum**, Copenhagen, v.121, n.1, p.149-162, May 2004.

MALCOMBER, S.T.; KELLOGG, E.A. SEPALLATA gene diversification: brave new whorls. **Plant Science**, Shannon, v.10, n.9, p.427-435, Sept. 2005.

MATSUBARA, K.; SHIMAMURA, K.; KODAMA, H.; KOKUBUN, H.; WATANABE, H.; BASUALDO, I.L.; ANDO, T. Green corolla segments in a wild *Petunia* species caused by a mutation in FBP2, a SEPALLATA-like MADS box gene. **Planta**, Berlin, v.228, n.3, p.401-409, Aug. 2008.

MELZER, R.; VERELST, W.; THEIBEN, G. The class E floral homeotic protein SEPALLATA3 is sufficient to loop DNA in 'floral quartet'-like complexes *in vitro*. **Nucleic Acids**, Oxford, v.37, n.1, p.144-157, Jan. 2009.

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**, Rockville, v.15, n.7, p.1538-1551, July 2003.

PELAZ, S.; DITTA, G.S.; BAUMANN, E.; WISMAN, E.; YANOFSKY, M.F. B and C floral organ identity functions require SEPALLATA MADS-box genes. **Nature**, London, v.405, n.6783, p.200-203, May 2000.

PURUGGANAN, M.D. The molecular evolution of development. **BioEssays**, Chicago, v.20, n.9, p.700-711, Sept. 1998.

PURUGGANAN, M.D.; ROUNSLEY, S.D.; SCHMIDT, R.J.; YANOFSKY, M.F. Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. **Genetics**, Austin, v.140, n.1, p.345-356, May 1995.

RIECHMANN, J.; KRIZEK, B.; MEYEROWITZ, E. Dimerization specificity of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v.93, n.10, p.4793-4798, May 1996.

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

SHORE, P.; SHARROCKS, A.D. The MADS-box family on transcriptional factors. **European Journal Biochemistry**, Berlin, 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**, Lawrence, v.12, n.2, p.319-333, May 1995.

SUNG, S.K.; MOON, Y.H.; CHUNG, J.E.; LEE, S.Y.; PARK, H.G.; AN, G. Characterization of MADS box genes from hot pepper. **Molecular Cells**, Cambridge, v.11, n.3, p.352-359, June 2001.

TAMURA, K.; DUDLEY, J.; NEI, M.; KUMAR, S.H. MEGA4: molecular evolutionary genetics analysis: software: version 4.0. **Molecular Biology and Evolution**, Oxford, v.24, n.8, p.1596-1599, Aug. 2007.

THOMPSON, J.D.; 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**, Oxford, v.22, n.22, p.4673-4680, Nov. 1994.

UIMARI, A.; KOTILAINEN, M.; ELOMAA, P.; YU, D.Y.; ALBERT, V.A.; TEERI, T.H. Integration of reproductive meristem fates by a SEPALLATA-like MADS-box gene. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v.101, n.44, p.15817-15822, Nov. 2004.

VANDENBUSSCHE, M.; ZETHOF, J.; ROYAERT, S.; WETERINGS, K.; GERATS, T. The duplicated B-class heterodimer model: whorl-specific effects and complex genetic interactions in *Petunia hybrida* flower development. **The Plant Cell**, Rockville, v.16, n.3, p.741-754, Feb. 2004.

VANDENBUSSCHE, M.; ZETHOF, J.; SOUER, E.; KOES, R.; TORNIELLI, G.B.; PEZZOTTI, M.; FERRARIO, S.; ANGENENT, G.C.; GERATS, T. Toward the analysis of the petunia MADS box gene family by reverse and forward transposon insertion mutagenesis approaches: B, C, and D floral organ identity functions require SEPALLATA-like MADS box genes in petunia. **The Plant Cell**, Rockville, v.15, n.11, p.2680-2693, Oct. 2003.

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.

XU, Y.; ZHANG, L.; XIE, H.; ZHANG, Y.Q.; OLIVEIRA, M.M.; MA, R.C. Expression analysis and genetic mapping of three SEPALLATA-like genes from peach (*Prunus persica* (L.) Batsch). **Tree Genetics & Genomes**, Heidelberg, v.4, n.4, p.693-703, Oct. 2008.

YU, H.; GOH, C.J. Identification and characterization of three orchid MADS-box genes of the AP1/AGL9 subfamily during floral transition. **Plant Physiology**, Washington, v.123, n.4, p.1325-1336, Aug. 2000.

ZAHN, L.M.; KING, H.Z.; LEEBENS-MACK, J.H.; KIM, S.; SOLTIS, P.S.; LANDHERR, L.L.; SOLTIS, D.E.; DEPAMPHILIS, C.W.; MA, H. The evolution of the SEPALLATA subfamily of MADS-Box genes: a preangiosperm origin with multiple duplications throughout angiosperm history. **Genetics**, Austin, v.169, n.4, p.2209-2223, Apr. 2005.

ZHAO, X.Y.; CHENG, Z.J.; ZHANG, X.S. Overexpression of TaMADS1, a SEPALLATA-like gene in wheat, causes early flowering and the abnormal development of floral organs in *Arabidopsis*. **Planta**, Berlin, v.223, n.4, p.698-707, Apr. 2006.

CAPÍTULO 3

EFICIÊNCIA DE PROTOCOLOS DE EXTRAÇÃO DE RNA EM DIFERENTES TECIDOS DO CAFEIEIRO

RESUMO

Para a utilização de técnicas sensíveis de biologia molecular, como o estudo de genes diferencialmente expressos, é necessário a obtenção de um RNA de boa qualidade e em quantidades adequadas. Devido à presença de grandes variedades, e frequentemente grande quantidade de compostos secundários em plantas, não existe um método padrão para o isolamento de ácidos nucleicos que possa ser utilizado para todas as espécies. Os polifenóis e os polissacarídeos são os compostos de maior interferência no processo de extração, e quando presentes geram um RNA de baixa qualidade. Nesse trabalho foram testados quatro métodos de extração de RNA (Método CTAB, Borato quente, CONCERT e Tri Reagente), em quatro diferentes tecidos de café (raiz, folha, flor e fruto), com o objetivo de determinar qual método é mais eficiente na obtenção de um RNA de melhor qualidade. Foi observado que os métodos, CTAB e Borato quente, que possuíam a adição PVP e/ou β -mercaptoetanol, e precipitação com LiCl, foram os que apresentaram RNAs mais puros e sem degradação em todos os tecidos, e puderam ser utilizados para a análise de expressão gênica. Com a extração utilizando o TriReagente, que tem como base o fenol, não foi obtido RNA de boa qualidade em todos os tecidos e conseqüentemente não foi possível a análise de expressão. O método de macroextração CTAB, foi o que apresentou amostras com RNA de melhor qualidade e em grandes quantidades em todos os tecidos, assim, ao ser submetido as etapas de tratamento com DNase e construção cDNA, obteve-se os melhores resultados. Trabalhos posteriores precisam ser realizados com o objetivo de padronizar esse método para microextração. Porém para a extração de RNA de folhas de café, todos os métodos foram eficientes e podem ser usados para a obtenção de amostras com qualidade considerável.

ABSTRACT

In order to use sensitive techniques of molecular biology, such as the study of genes differentially expressed, it is necessary to obtain a RNA of great quality and in suitable quantities. Due to the presence of several varieties, and often expressive quantities of secondary compounds in plants, there is no standard method for the isolation of nucleic acids that can be used for all species. Polyphenols and polysaccharides are compounds that mostly interfere in the extraction process, and when they are present, a low quality RNA is produced. In this work it was tested four RNA extraction methods (CTAB method, Hot Borate, CONCERT and Tri Reagent), in four different coffee tissues (root, leaf, flower and fruit), aiming at determining which method is more efficient to obtain a high quality RNA. It was observed that the CTAB and Hot Borate methods, in which PVP was added and/or β -mercaptoethanol, and precipitation with LiCl was performed, presented more pure RNA and no degradation was observed in any tissue, and then, they could be used in further gene expression analysis. The extraction with TriReagent, which includes the use of phenol, a high quality RNA wasn't obtained in any tissue, and thus, the expression analysis was disturbed. CTAB macroextraction method was the one that presented samples with higher RNA quality, and with larger quantities in all tissues and therefore, when RNA samples were submitted to the DNase treatment and cDNA construction, the best results were possible. Future works need to be carried out aiming at the standardization of this macroextraction method. Nevertheless, in RNA extraction of coffee leaves, all methods were efficient and can be used to obtain samples with considerable quality.

1 INTRODUÇÃO

O isolamento de RNA de qualidade é um pré-requisito para estudos de expressão gênica incluindo transcriptase reversa (RT), PCR quantitativo em tempo real (qRT-PCR), construção de bibliotecas de cDNA ou análises de *microarray* (Cardillo et al., 2006).

O principal problema encontrado é que tecidos vegetais são caracterizados por uma alta variação na sua composição, e alguns tecidos são especialmente difíceis de isolar o RNA (Geuna et al., 1998). Para trabalhos que utilizam diferentes tecidos é de extrema importância um método de extração eficiente para todos, independente da composição do material, a fim de minimizar as variações que possam existir quando utilizados diferentes protocolos.

Devido à presença de metabólitos secundários, polissacarídeos e polifenóis, a extração de RNA de qualidade em tecidos do cafeeiro torna-se uma etapa de difícil padronização entre os diferentes tecidos e a obtenção de alta quantidade de RNA pode ser complicada pela suscetibilidade do RNA à degradação por RNAses. Além disso, uma das etapas essenciais é a eliminação total do DNA presente após a extração, pois, teoricamente uma cópia do DNA é suficiente para originar um falso positivo em estudos baseados em expressão gênica (Birtic & Kranner, 2006).

Estudos descrevem técnicas de extração para tecidos vegetais ricos em polifenóis ou polissacarídeos que apresentam problemas para isolamento de ampla quantidade de RNA de alta qualidade (Salzman et al., 1999; Azevedo et al., 2003). Contudo, esses métodos têm sido desenvolvidos para tecidos vegetais específicos, pois cada método depende da composição do tecido e alguns têm sido inadequados na redução de contaminação por polissacarídeos.

Entre as técnicas de extração mais relatadas para espécies com alto conteúdo de metabólitos secundários, estão os métodos utilizando polivinilpirrolidona solúvel (PVP) e precipitação com etanol (Salzman et al., 1999), borato quente (Wan & Wilkins, 1994), extração com fenol (Komjanc et al., 1999), precipitação com cálcio (Dal Cin et al., 2005), 2-butoxi-etanol (Manning, 1991; Malnoy et al., 2001) ou brometo de cetil trimetil amônio (CTAB) no tampão de extração (Meisel et al., 2005).

Nessas técnicas, são acrescentadas algumas substâncias para eliminar os contaminantes, como por exemplo, a adição de agentes antioxidantes, como PVP, polivinilpolipirrolidona (PVPP) e β -mercaptoetanol; o uso de solventes orgânicos como o clorofórmio e o fenol, que separam o RNA das proteínas; a adição de guanidina tiocianeto amplamente usada para inibir a atividade de RNAses (Chomczynski & Sacchi, 1987; Valenzuela-Avendaño et al., 2005) dentre outras.

Visando a obtenção de um protocolo ideal para extração de RNA de qualidade para vários tecidos de café, esse trabalho teve como objetivo identificar a metodologia mais adequada para isolar amostras de RNA em folha, flor, fruto e raiz de café, que possam ser utilizadas em estudos moleculares sensíveis.

2 MATERIAL E MÉTODOS

2.1 Material vegetal

O material vegetal utilizado foi coletado no campo experimental da Universidade Federal de Lavras (UFLA) em Lavras- MG. Para a extração de RNA de café foram utilizados quatro tecidos vegetais de *Coffea arabica*, cultivar Rubi, coletados no período de setembro de 2008: folhas (jovens), flores (estádios G4-G5 descritos por Moraes et al. (2008), frutos (em estádios diferentes de desenvolvimento) e raízes (secundárias). Foram coletados em nitrogênio líquido e posteriormente armazenados a -80 °C até o momento da extração. Os tecidos foram submetidos a quatro diferentes métodos de extração de RNA: Concert™ (Invitrogen), TRI Reagente® (Sigma), Borato quente (Wan & Wilkins, 1994; Birtic & Kranner, 2006), e CTAB (Chang et al., 1993).

Todos os materiais utilizados para extração foram tratados com solução diethylpirocarbonato (DEPC) a 0,5% para inativar RNAses. As soluções utilizadas foram preparadas com água destilada autoclavada RNase-free.

2.2 Isolamento do RNA

2.2.1 Reagente Concert™

O protocolo de extração de RNA com Concert foi realizado de acordo com manual: Concert™ Plant RNA Reagent (Invitrogen). A cada microtubo contendo aproximadamente 100 mg de tecido macerado foram adicionados 500 µL do reagente Concert gelado (4 °C) e homogeneizados em vórtex. Em seguida, os tubos foram incubados por 5 min a temperatura ambiente, deixados horizontalmente para maximizar a extração do RNA. Após esse período, o material foi submetido à centrifugação por 2 min à temperatura ambiente com velocidade 12.000 g e o sobrenadante foi transferido para um novo tubo. Foram adicionados 100 µL de NaCl 5 M e homogeneizados em vórtex. Posteriormente,

foram adicionados 300 μ L de clorofórmio e os tubos foram homogeneizados por inversão. Para separação das fases as amostras foram submetidas à centrifugação durante 10 min a 4 °C (12.000 g) e a fase aquosa superior foi transferida para um novo tubo. Em seguida, foi adicionado um volume equivalente à fase aquosa de isopropanol gelado (aproximadamente 400 μ L) e agitados em vórtex durante 5 seg. As amostras foram mantidas à temperatura ambiente por 10 min, e posteriormente submetidas à centrifugação também por 10 min a 4 °C (12.000 g). O sobrenadante foi descartado, o *pellet* foi lavado com 1 mL de etanol 75% gelado e os tubos foram submetidos à centrifugação por 1 min à temperatura ambiente (12.000 g). O líquido residual foi removido do tubo com a pipeta e o RNA foi ressuspensionado em 20 μ L de água Milli-Q autoclavada. As amostras foram armazenadas a -20 °C.

2.2.2 TRI Reagente®

O RNA total foi obtido utilizando TRI Reagente® (Sigma), conforme recomendações do fabricante. A cada microtubo contendo aproximadamente 100 mg de tecido macerado foram adicionados 1 mL do Tri Reagente, e homogeneizados em vórtex. Em seguida os tubos foram incubados por 5 min à temperatura ambiente. Após esse período, o material foi submetido à centrifugação por 10 min a 4 °C com velocidade 12.000 g. O sobrenadante foi transferido para um novo tubo onde foram adicionados 200 μ L de clorofórmio e homogeneizados por 15 seg. Para separação das fases as amostras foram submetidas à centrifugação durante 20 min a 4 °C (12.000 g) e a fase aquosa superior (aproximadamente 750 μ L) foi transferida para um novo tubo. Posteriormente foram adicionados 500 μ L de isopropanol, e os tubos foram homogeneizados por inversões e incubados por 60 min a -20 °C. As amostras foram centrifugadas durante 10 min a 4 °C (12.000 g), e o sobrenadante descartado. O *pellet* foi lavado com 1 mL de etanol 75% gelado e os tubos foram

submetidos à centrifugação por 5 min a 4 °C (12.000 g). O líquido residual foi removido do tubo com a pipeta e o RNA foi ressuspensionado em 20 µL de água Milli-Q autoclavada. As amostras foram armazenadas a -20 °C.

2.2.3 Método Borato quente

O método do Borato Quente foi adaptado do protocolo descrito por Birtic & Kranner (2006), onde o tampão de extração é formado de Sodium tetraborate decahydrate 0,2 M; EGTA - ethylene glycol bis (2-aminoethyl ether)-N,N,N',N'-tetraacetic acid 30 mM; SDS - sodium dodecyl sulfate 1%; Sodium deoxycholate 1%, ajustado para pH 9.0 com hidróxido de sódio. Para cada 1 mL de tampão, foi adicionado 0,06 g de polyvinylpyrrolidone (PVP) e 2% de dithiothreitol (DTT), e aquecidos a 80 °C para dissolução completa dos reagentes. Os tecidos foram macerados em nitrogênio líquido, aproximadamente 100 mg de tecido foram acrescentados 1 mL de tampão de extração. Imediatamente, foram adicionados 1 mg de Proteinase K, agitados e colocados em banho maria a 45 °C por 1 h 30 min. Posteriormente, foi adicionado 80 µL de Cloreto de Potássio 2 M e as amostras foram incubadas no gelo por 45 min. Em seguida, foram centrifugadas por 20 min a 4 °C (12.000 g) e recolhido o sobrenadante. Foi acrescentado 1/3 do volume correspondente ao sobrenadante, de Cloreto de Lítio 8 M (gelado) e as amostras foram incubadas a -20 °C por no mínimo 1 h para precipitação do RNA. Após esse período, foram centrifugadas por 20 min a 4 °C (12.000 g), e o sobrenadante descartado. O *pellet* formado foi lavado com 750 µL de Cloreto de Lítio 2 M pelo menos duas vezes. Novamente as amostras foram centrifugadas por 10 min, o líquido residual foi removido do tubo com a pipeta e o RNA foi ressuspensionado em 20 µL de água Milli-Q autoclavada. As amostras foram armazenadas a -20 °C.

2.2.4 Método CTAB

É um método de macroextração de RNA que foi descrito por Chang et al. (1993). Em cada tubo, contendo 2 g do material macerado em nitrogênio líquido, foram adicionados 25 mL de tampão de extração (2% (p/v) CTAB (cethyltrimethylammonium bromide), 2% (p/v) PVP, 100 mM de Tris-HCL, 25 mM de EDTA, 20 mM de NaCl e 2% de β -mercaptoetanol).

Após as amostras serem incubadas a 65 °C por 20 min, foram adicionados 20 mL de clorofórmio e, em seguida centrifugada por 5 min a 7.500 g, à temperatura ambiente. Aproximadamente 20 mL de sobrenadante foi transferido para um novo tubo, adicionado o mesmo volume de clorofórmio, seguido de homogeneização e novamente centrifugação por 5 min, à temperatura ambiente (7.500 g). A fase aquosa, aproximadamente 15 mL, foi transferida para um novo tubo, e adicionado Cloreto de Lítio a 12 M para uma concentração final de 2,5 M. As amostras foram deixadas *overnight* a 4 °C para a precipitação. Posteriormente, o material foi centrifugado durante 30 min a 4 °C (7.500 g) e o sobrenadante descartado. O precipitado formado foi lavado com Cloreto de Lítio a 2,5 M, seguido de nova centrifugação por 15 min a 4 °C (10.000 g).

Em seguida, o sobrenadante foi descartado, o *pellet* lavado com 1 mL de etanol 70% (v/v), transferidos para microtubos e centrifugados por 10 min a 4 °C (11.000 g). Finalmente, o sobrenadante foi descartado e o líquido residual removido do tubo com a pipeta. O RNA foi ressuscitado em 50 μ L de água Milli-Q autoclavada. As amostras foram armazenadas a -20 °C.

2.3 Tratamento com DNase

Todas as amostras foram tratadas com DNase DNA free (Ambion), em uma reação contendo 10 μ g de RNA total, 1 U da enzima rDNaseI, 0,1 volume do tampão DNaseI 10 X e água livre de DNase/RNase para completar o volume de 50 μ L. Essa reação foi incubada a 37 °C por 30 min aproximadamente e

adicionada 5 μL da solução para inativar DNase (DNase inactivation). As amostras foram incubadas por 2 min, centrifugadas a 10.000 g por 1,5 min e transferido o sobrenadante, aproximadamente 50 μL , para novos tubos.

Para a avaliação da integridade das amostras extraídas, o RNA foi submetido a eletroforese em gel de agarose 1,2% (m/v) corado com 2 μL de Brometo de Etídio (0,5 $\mu\text{g}/\text{mL}$), posteriormente visualizado sob luz ultravioleta e a imagem captada pelo fotodocumentador EDAS 290 (Kodak[®]). As amostras foram quantificadas em espectrofotômetro (Nanodrop[®] Espectrophotometer ND-1000) a A_{260} nm com o objetivo de determinar a quantidade e qualidade.

2.4 Síntese cDNA e qRT-PCR

Apenas com o objetivo de determinar se os RNAs extraídos possuem qualidade para ser submetidos a análises sensíveis, foi realizado a síntese de cDNA e a análise por qRT PCR.

Para síntese do cDNA foi utilizado o *Kit High-Capacity cDNA Reverse Transcription* (Applied Biosystems). Primeiramente, o RNA foi preparado a uma concentração de 1 μg em um volume final de 10 μL . Após essa etapa, foi preparado um mix contendo 2 μL do tampão 10X da enzima, 2 μL do primer RT Random Primers 10X, 0,8 μL do mix dNTP (100 mM), 1 μL *MultiScribeTM* Reverse Transcriptase, e água para um volume final de 10 μL /amostra. Para cada solução preparada de 10 μL de RNA a 1 μg , foram acrescentados 10 μL desse mix. Os tubos foram submetidos ao termociclador *Eppendorf Mastercycler gradient*, programado com três etapas: 10 min a 25 °C para o anelamento dos *primers*; 2 h a 37 °C para ação da enzima e 5 min a 85 °C para inativá-la. As amostras foram conservadas no freezer a -20 °C.

O cDNA foi utilizado como *template* para análise da expressão gênica quantitativa através do ABI PRISM 7500 Real-Time PCR (Applied Biosystems), pelo sistema de detecção SYBR Green.

Primeiramente foi realizado um ensaio de quantificação absoluta, onde foi determinada a curva padrão, a eficiência dos *primers* e a melhor diluição apresentada para as amostras. Os cDNA foram diluídos 1:10, 1:25 e 1:50.

Após determinar a diluição 1:10 e a eficiência dos *primers* entre 94% a 97%, foi realizado o ensaio da expressão relativa, através do método do C_T comparativo. As amostras foram processadas em triplicatas com três repetições biológicas. Foi analisada a expressão dos genes GAPDH e 14.3.3, que são considerados controles endógenos eficientes para café, segundo Barsalobres-Cavallari et al. (2009). A expressão do gene GAPDH foi normalizada com o controle 14.3.3, e a expressão do gene 14.3.3 foi normalizada com o gene controle GAPDH.

As condições térmicas da reação foram: 2 min a 50 °C, 10 min a 95 °C, seguidos por 40 ciclos de 15 seg a 95 °C e 1 min a 60 °C, e finalizando com 1 seg a 95 °C. Os dados foram analisados no programa *7500 Fast Software* (Versão 2.1). Para cada reação, foram utilizados 1 µL cDNA, 1,5 µM de cada *primer* e 5,0 µL de *Master Mix SYBR Green* (Applied) para um volume final 10,0 µL/amostra.

Para o cálculo da expressão, primeiramente cada amostra foi normalizada com o controle endógeno, utilizando a equação $\Delta C_T = C_T$ (gene alvo) - C_T (controle endógeno) e a quantificação relativa foi obtida pela fórmula $2^{-\Delta\Delta C_T}$. A linha de corte (*Threshold*) foi definida manualmente.

3 RESULTADOS E DISCUSSÃO

Para os procedimentos de extração, o mesmo material coletado foi utilizado e a mesma quantidade de tecido foi padronizada nas microextrações. Todos os métodos apresentaram uma quantidade de RNA suficiente para posteriores trabalhos, apenas no método CTAB, por ser uma macroextração, foi utilizado uma quantidade maior de tecido vegetal, conseqüentemente obtido uma quantidade de RNA elevada. Porém, na comparação entre os métodos, será considerada apenas a qualidade das amostras.

Para o sucesso do isolamento do RNA, é necessário impedir que polissacarídeos e polifenóis, encontrados em diferentes quantidades, nos diferentes tecidos e em espécies distintas, liguem-se aos ácidos nucleicos, contribuindo na determinação da quantidade, qualidade e integridade do RNA (Suzuki et al., 2003; Zamboni et al., 2008). Desta forma, esses parâmetros foram avaliados para determinar o método de extração de RNA mais eficiente para café.

As relações de pureza e as concentrações dos RNAs provenientes de todos os métodos de extração avaliados estão dispostas nas Tabelas 1. Para a avaliação da pureza, são observados os valores de absorbância a 280 nm, 260 nm e 230 nm. A relação entre a absorbância A260/A230 indica a contaminação por polissacarídeos ou polifenóis, e o valor da relação A260/A280, indica a contaminação por proteínas (Manning, 1991; Logemann et al., 1997). Segundo Asif et al. (2006) quando esses valores estiverem entre 1,8-2,1, respectivamente, indicam a descontaminação das amostras.

Ao comparar os métodos descritos no trabalho, pode-se observar que o método do Borato quente e o CTAB, em geral, apresentaram uma melhor qualidade das amostras. Em destaque, ao método do CTAB, que, como pode ser observado na Tabela 1, apresentou para todos os tecidos os melhores resultados,

com as relações de absorvância ideais, determinando amostras com pouca ou nenhuma contaminação com proteínas e polissacarídeos. Isso pode estar relacionado ao uso de substâncias como o PVP e β -mercaptoetanol, que possuem a função de reduzir as substâncias oxidantes e os compostos fenólicos (Jaakola, 2001), e pela precipitação com cloreto de lítio, que possui a função de reter substâncias protéicas, auxiliando na descontaminação das amostras. A utilização do cloreto de lítio foi descrita por Rubio-Piña & Vázquez-Flota (2008), como sendo a melhor substância utilizada nos protocolos para a precipitação do RNA. O uso do detergente CTAB no tampão de extração do método do CTAB é a base para a separação dos polissacarídeos dos ácidos nucleicos (Chang et al., 1993; Rogers & Benedich, 1994), podendo ser também um diferencial para a obtenção de um RNA com alta qualidade.

O método do Concert e principalmente o Tri Reagente, apresentaram amostras contaminadas com proteínas e polifenóis, segundo a relação dos valores de absorvância, que representaram valores menores aos recomendados (Tabela 1). Na extração realizada utilizando o Tri Reagente com o protocolo sugerido pelo fabricante, é observado um baixo rendimento entre os tecidos e conseqüentemente uma alta contaminação, podendo ser resultante da presença de proteínas e compostos polifenólicos. Resultado semelhante foi observado em frutos de *Vaccinium myrtillus* por Jaakola et al. (2001). Esse reagente é composto da substância química fenol, que pode também causar danos a cauda poli-A do mRNA extraído (Azevedo et al., 2003), dificultando a síntese do cDNA e análises posteriores.

Entretanto, ao analisar os tecidos separadamente (Tabela 1), pode-se observar que o tecido que possui melhor qualidade em quase todos os métodos foi o de folha, caracterizado pela presença de poucos polifenóis.

Os frutos por possuírem altos níveis de compostos fenólicos, polissacarídeos e elevados níveis de RNAses (Jaakola et al., 2001), dificultam o

processo de extração. Esses polissacarídeos formam uma estrutura gelatinosa que precipita durante a extração afetando tanto na quantidade quanto na qualidade do RNA (Sharma et al., 2003). Porém, apenas o método do Borato quente e CTAB foram eficientes na retirada desses constituintes, apresentando valores da relação A260/A230 maiores que 1 (Tabela 1), devido à ação das substâncias acrescentadas a esses métodos e que já foram discutidas anteriormente.

Em relação à integridade das amostras, pode ser observado na Figura 1, uma relação entre a integridade e a qualidade, determinando que nos métodos que apresentaram uma qualidade satisfatória, mostraram também uma integridade das bandas 28S e 18S característica do rRNA.

Foi possível observar um perfil expressão nas amostras extraídas com os métodos Concert, Borato quente e CTAB, determinando a qualidade dessas amostras (Figura 2 e 3). As amostras extraídas com Tri Reagente, depois do tratamento com DNase e síntese do cDNA, apresentaram quantidade e qualidade baixas mas suficientes para a realização do qRT-PCR, entretanto, para algumas, foram observados padrões de expressão característicos de amostras degradadas. Isso pode ser devido ao Tri Reagente não possuir a capacidade de inativar todos os contaminantes das amostras de café, permanecendo no material e posteriormente degradando a amostra. Ou pelo protocolo usado não ser suficiente para lavar as substâncias interferentes presentes no momento da reação. Como a técnica de PCR em tempo real possui uma sensibilidade de detecção alta, esses contaminantes podem interferir na observação da expressão, ou se as amostras estão degradadas não é realizada a amplificação do fragmento desejado.

De acordo com as Figuras 2 e 3, pode ser observado que os métodos de extração Borato quente, Concert e CTAB foram eficientes para a eliminação de todo e qualquer contaminante que possa impedir a reação, pois em todos os

tecidos foi observada a expressão dos genes GAPDH e 14.3.3, como era de se esperar já que são genes considerados constitutivos de café (Barsalobres-Cavallari et al., 2009).

Embora a extração de alguns tecidos de café utilizando o método do Tri Reagente não apresente RNAs com qualidade suficiente para ser submetido a análise de PCR em tempo real, pode ser observado que as amostras de RNA de folha, mesmo consideradas de baixa qualidade, foram suficientes para essa análise (dados não mostrados). Isso pode ser explicado pelo fato do tecido foliar de café não apresentar tantos compostos metabólicos e contaminantes, como nos outros tecidos utilizados no trabalho.

Analisando a Tabela 1 e o Figura 1, podemos inferir que o tecido foliar de café gerou amostras de RNA de qualidade quando utilizados todos os quatro protocolos. Entretanto, quando se compara os métodos para extração dos tecidos de folha, flor, raiz e fruto de café, os melhores resultados foram observados no método do CTAB, porém por ser um método de macroextração, é necessário a disponibilidade de uma quantidade maior de material vegetal. Os métodos Concert e Borato quente também foram eficientes para extração de RNAs de qualidade para análise de PCR em tempo real, porém, o método do Borato quente, além de ser um método mais barato, apresentou amostras com qualidade superior ao método de extração utilizando o reagente Concert.

TABELA 1 Comparação da qualidade dos RNAs isolados de diferentes tecidos de café através de quatro metodologias. Os dados representam o SE de três replicatas biológicas

Tecidos	<i>Borato Quente</i>			<i>CTAB</i>			<i>Concert™</i>			<i>Tri Reagente®</i>		
	Quantida de RNA (ng/μL)	Pureza A ₂₆₀ /A ₂₈₀	Pureza A ₂₆₀ /A ₂₃₀	Quantida de RNA (ng/μL)	Pureza A ₂₆₀ /A ₂₈₀	Pureza A ₂₆₀ /A ₂₃₀	Quantida de RNA (ng/μL)	Pureza A ₂₆₀ /A ₂₈₀	Pureza A ₂₆₀ /A ₂₃₀	Quantida de RNA (ng/μL)	Pureza A ₂₆₀ /A ₂₈₀	Pureza A ₂₆₀ /A ₂₃₀
Raiz	301 ± 52	1,5 ± 0,12	1,2 ± 0,01	176 ± 37	2,1 ± 0,02	2,1 ± 0,18	776 ± 106	1,6 ± 0,04	0,6 ± 0,08	327 ± 53	1,3 ± 0,03	0,7 ± 0,13
Folha	635 ± 60	2,1 ± 0,02	2,0 ± 0,03	2796 ± 761	2,0 ± 0,06	2,1 ± 0,08	2045 ± 176	1,9 ± 0,06	1,1 ± 0,16	647 ± 80	1,6 ± 0,06	0,5 ± 0,09
Flor	782 ± 179	2,0 ± 0,04	1,1 ± 0,07	4176 ± 52	1,4 ± 0,14	1,5 ± 0,16	1141 ± 20	1,9 ± 0,10	0,7 ± 0,02	1197 ± 149	1,3 ± 0,07	0,5 ± 0,03
Fruto	225 ± 46	1,9 ± 0,08	1,5 ± 0,19	3109 ± 74	2,1 ± 0,03	1,7 ± 0,08	430 ± 33	1,4 ± 0,06	0,5 ± 0,11	481 ± 175	1,0 ± 0,03	0,7 ± 0,24

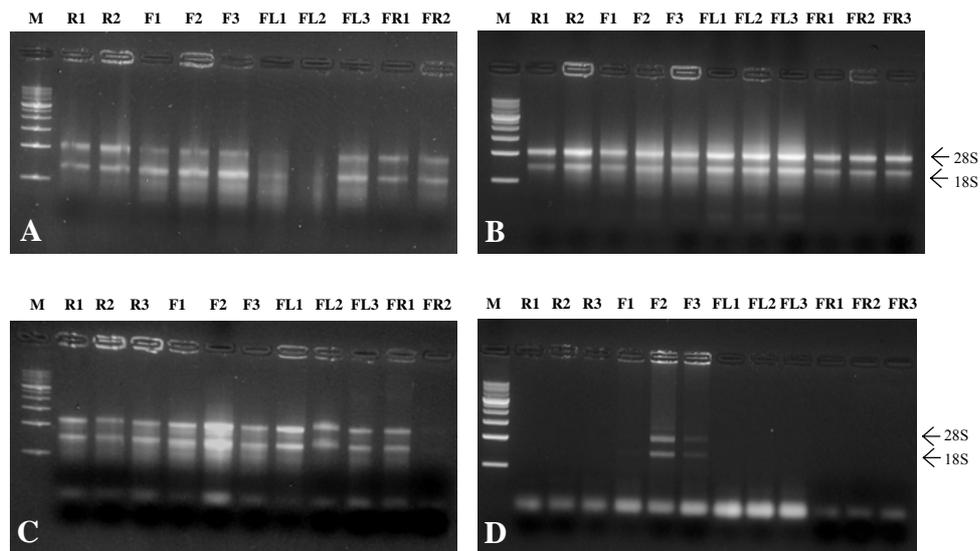


FIGURA 1 Gel de agarose 1,2% corado com brometo de etídio. As amostras utilizadas foram tratadas com DNase e aplicado 2 μ L no gel. A – extração pelo método Borato quente; B – CTAB; C – Concert; D- TriReagente. M – marcador molecular 1 Kb; Tecidos de *Coffea arabica*: 1-3 (Repetições biológicas). R (1-3) raiz; F (1-3) folha; FL (1-3) flor; FR (1-3) fruto.

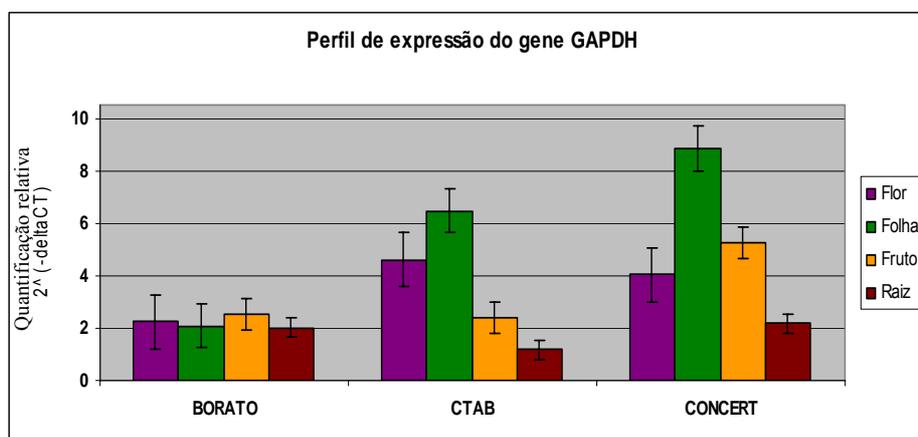


FIGURA 2 Exemplo do padrão de expressão do gene endógeno GAPDH, normalizado com 14.3.3, em tecidos vegetais de café: flor (roxo); folha (verde), fruto (laranja), raiz (vermelho); e utilizando diferentes métodos de extração, Borato quente, CTAB e Concert. Os valores representados são as médias dos valores de expressão das réplicas biológicas.

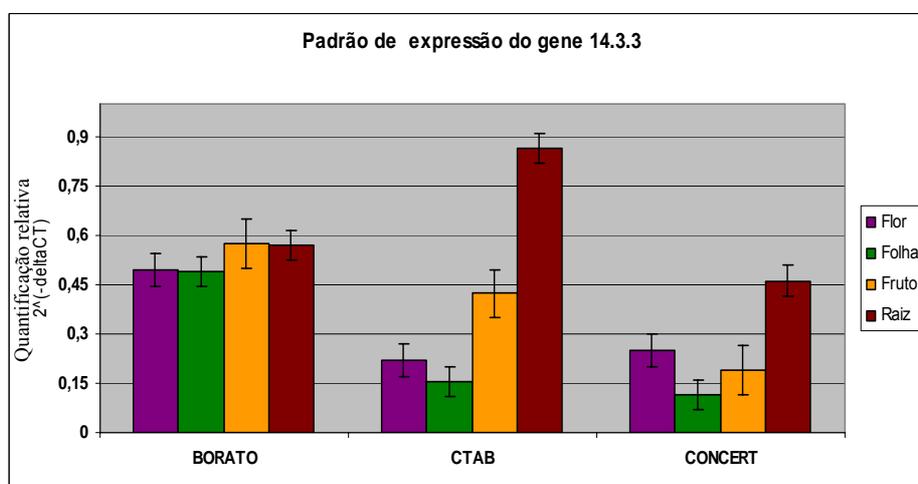


FIGURA 3 Exemplo do padrão de expressão do gene endógeno 14.3.3 normalizado com GAPDH, em tecidos vegetais de café: flor (roxo); folha (verde), fruto (laranja), raiz (vermelho); e utilizando diferentes métodos de extração, Borato quente, CTAB e Concert. Os valores representados são as médias dos valores de expressão das réplicas biológicas.

4 CONCLUSÕES

A utilização de determinados componentes na extração pelo método do Borato quente, CTAB e Concert, apresentam RNAs de raiz, folha, flor e fruto com poucos contaminantes, não interferindo na análise quantitativa pela qRT-PCR.

O método CTAB, foi considerado o que gerou RNAs de melhor qualidade em todos os tecidos e em todo o processo de manipulação.

A extração realizada com o Tri Reagente apresentou a maioria das amostras muito contaminadas, sendo considerado um método ineficiente para obter RNA de boa qualidade de café.

A extração de RNA de folha de café, tanto no método do Tri Reagente, como em todos os outros métodos é eficiente e pode ser utilizado para obtenção de um RNA de boa qualidade.

5 REFERÊNCIAS BIBLIOGRÁFICAS

ASIF, M.; TRIVEDI, P.; SOLOMOS, T.; TUCKER, M. Isolation of high-quality RNA from apple (*Malus domestica*) fruit. **Journal of Agricultural and Food Chemistry**, Washington, v.54, n.15, p.5227-5229, June 2006.

AZEVEDO, H.; LINO-NETO, T.; TAVARES, R.M. An improved method for high quality RNA isolation from needles of adult maritime pine trees. **Plant Molecular Biology Reporter**, Athens, v.21, n.4, p.333-338, Dec. 2003.

BARSALOBRES-CAVALLARI, C.F.; SEVERINO, F.E.; MALUF, M.P.; MAIA, I.G. Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions. **BMC Molecular Biology**, Bethesda, v.10, n.1, p.1-11, Jan. 2009.

BIRTIC, S.; KRANNER, I. Isolation of high-quality from polyphenol-, polysaccharide and lipid-rich seeds. **Phytochemical Analysis**, Kew, v.17, n.3, p.144-148, Mar. 2006.

CARDILLO, A.B.; GIULIETTI, A.M.; MARCONI, P.L. Analysis and sequencing of h6hmRNA, last enzyme in the tropane alkaloids pathway from anthers and hairy root cultures of *Brugmansia candida* (*Solanaceae*). **Electronic Journal of Biotechnology**, Cambridge, v.9, n.3, p.196-198, June 2006.

CHANG, S.; PURYEAR, J.; CAIRNEY, J. A simple and efficient method for isolating RNA from pine trees. **Plant Molecular Biology Reporter**, Athens, v.11, n.2, p.113-116, June 1993.

CHOMCZYNSKI, P.; SACCHI, N. Single-step method of RNA isolation by acid guanidium thiocyanate-phenolchloroform extraction. **Analytical Biochemistry**, New York, v.162, n.1, p.156-159, Apr. 1987.

DAL CIN, V.; DANESIN, M.; RIZZINI, F.M.; RAMINA, A. RNA extraction from plant tissues: the use of calcium to precipitate contaminating pectic sugars. **Molecular Biotechnology**, Totowa, v.31, n.2, p.113-119, July 2005.

GEUNA, F.; HARTINGS, H.; SCIENZA, A. A new method for rapid extraction of high quality RNA from recalcitrant tissues of grapevine. **Plant Molecular Biology Reporter**, Athens, v.16, n.1, p.61-67, Mar. 1998.

JAAKOLA, L.; PIRTTILA, A.M.; HALONEN, M.; HOHTOLA, A. Isolation of high quality RNA from bilberry (*Vaccinium myrtillus* L.) fruit. **Molecular Biotechnology**, Washington, v.19, n.2, p.201-203, Oct. 2001.

KOMJANC, M.; FESTI, S.; RIZZOTTI, L.; CATTIVELLI, L.; CERVONE, F.; LORENZO, G. de. A leucine-rich repeat receptor-like protein kinase (LRPKm1) gene is induced in *Malus x domestica* by *Venturia inaequalis* infection and salicylic acid treatment. **Plant Molecular Biology**, Athens, v.40, n.6, p.945-957, Aug. 1999.

LOGEMANN, J.; SCHELL, J.; WILLMITZER, L. Improved method for isolation of RNA from plant tissue. **Analytical Biochemistry**, New York, v.163, n.1, p.16-20, May 1987.

MALNOY, M.; REYNOIRD, J.P.; MOURGUES, F.; CHEVREAU, E.; SIMONEAU, P. A method for isolating total RNA from pear leaves. **Plant Molecular Biology Reporter**, Athens, v.19, n.1, p.69a-69f, Mar. 2001.

MANNING, K. Isolation of nucleic acid from plants by differential solvent precipitation. **Analytical Biochemistry**, New York, v.195, n.1, p.45-50, May 1991.

MEISEL, L.; FONSECA, B.; GONZÁLEZ, S.; BAEZAYATES, R.; CAMBIAZO, V.; CAMPOS, R.; GONZALEZ, M.; ORELLANA, A.; RETAMALES, J.; SILVA, H. A rapid and efficient method for purifying high quality total RNA from peaches (*Prunus persica*) for functional genomics analyses. **Biological Research**, Santiago, v.38, n.1, p.83-88, Jan. 2005.

MORAIS, H.; CARAMORI, P.H.; KOGUISHI, M.S.; RIBEIRO, A.M.A. Escala fenológica detalhada da fase reprodutiva de coffeea arabica. **Bragantia**, Campinas, v.67, n.1, p.257-260, jan. 2008.

ROGERS, S.O.; BENEDICH, A.J. **Plant molecular biology manual**. Netherlands: Kluwer Academic, 1994. 8p.

RUBIO-PIÑA, C.; VAZQUEZ-FLOTA, D. Isolation of functional total RNA from *Argemone mexicana* tissues. **Plant Biotechnology**, Santiago, v.11, n.4, p.1-5, Oct. 2008.

SALZMAN, R.A.; FUJITA, T.; ZHU-SALZMAN, K.; HASEGAWA, P.M.; BRESSAN, R.A. An improved RNA isolation method for plant tissues containing high levels of phenolic compounds or carbohydrates. **Plant Molecular Biology Reporter**, Athens, v.17, n.1, p.11-17, Mar. 1999.

SHARMA, A.D.; GILL, P.K.; SINGH, P. RNA isolation from plant tissues rich in polysaccharides. **Analytical Biochemistry**, New York, v.314, n.2, p.319-321, Mar. 2003.

SUZUKI, M.; KETTERLING, M.G.; MCCARTY, D.R. Quantitative statistical analysis of *cis*-regulatory sequences in ABA/VP1- and CBF/DREB1-regulated genes of arabidopsis. **Plant Physiology Review**, Washington, v.139, n.1, p.437-447, Aug. 2005.

VALENZUELA-AVENDAÑO, J.P.; ESTRADA-MOTA, I.A.; LIZAMA-U.C.G.; SOUZA-PERERA, R.; VALENZUELA-SOTO, E.M.; ZÚÑIGA-AGUILAR, J.J. Use of a simple method to isolate intact RNA from partially hydrated *Selaginella lepidophylla* plants. **Plant Molecular Biology Reporter**, Athens, v.23, n.2, p.199-201, June 2005.

WAN, C.Y.; WILKINS, T.A. A modified hot borate method significantly enhances the yield of high-quality RNA from cotton (*Gossypium hirsutum* L.). **Analytical Biochemistry**, New York, v.233, n.1, p.7-12, Jan. 1994.

ZAMBONI, A.; PIERANTONI, L.; FRANCESCHI, P. de. Total RNA extraction from strawberry tree (*Arbutus unedo*) and several other woody-plants. **Journal of Biogeosciences and Forestry**, Bologna, v.1, n.1, p.122-125, Aug. 2008.

CAPÍTULO 4

ISOLAMENTO E CARACTERIZAÇÃO PARCIAL DOS GENES *SEPALLATA* EM CAFÉ

RESUMO

O florescimento é determinado pela expressão regulada de uma rede complexa de genes, que estão interligados a fatores ambientais. Essa rede é composta pelo modelo molecular do florescimento ABCE, formado por vários genes da família *MADS-box*, dentro dos quais os genes *SEPALLATA* (*SEP*) apresentam um papel essencial, determinando a identidade dos órgãos florais em cada um dos quatro verticilos, e no desenvolvimento dos frutos. Quatro tipos foram totalmente definidos em *Arabidopsis thaliana* (*SEP1*, *SEP2*, *SEP3* e *SEP4*), todos atuando de maneira redundante formando complexos proteicos com outros genes. O objetivo deste trabalho foi isolar, clonar, sequenciar e comprovar a expressão tecido-específica, dos prováveis genes *SEP* de *C. arabica*, identificados pela análise *in silico*. Foram desenhados *primers* específicos para cada uma das três sequências identificadas, e a clonagem desses fragmentos foi realizada no vetor *TOPO TA*. O resultado obtido a partir do sequenciamento mostrou que as sequências de café foram similares às sequências *in silico*, caracterizando parcialmente duas sequências como possíveis genes *SEP1/2* e *SEP3*, demonstrando que a análise preliminar, utilizando bioinformática, é uma ferramenta eficiente em estudos moleculares. Além disso, as análises de expressão, realizadas pela qRT-PCR em diferentes tecidos, detectaram a expressão apenas nos tecidos de flores e frutos, corroborando com os resultados obtidos pela análise do *Northern* eletrônico e também com vários trabalhos de expressão dos genes *SEP* em outras espécies. Deste modo, essas análises permitiram a identificação e caracterização de dois genes *SEP* em café, entretanto trabalhos subsequentes são necessários para confirmar e melhor caracterizar estes genes.

ABSTRACT

The flowering process is determined by a balanced expression of a complex gene network, interconnected to environmental factors. This network is formed by the flowering molecular model, so called ABCE, composed by several genes of the *MADS-box* gene family, among of which the *SEPALLATA* (*SEP*) genes play an essential role, determining the identity of floral organs in each of the whorls, and fruit development. It has been identified four *SEPALLATA* genes in *Arabidopsis* (*SEP1*, *SEP2*, *SEP3*, *SEP4*), all acting in a redundant manner, forming protein complexes with other genes. This work aimed at isolating, cloning, sequencing and proving the tissue-specific expression of putative *C. arabica* *SEP* genes, previously identified by the *in silico* analysis. Specific primers were designed for each of the three sequences identified, and the cloning process was performed using the *TOPO TA* vector. The obtained results from the sequencing process showed that the genes sequences were similar to the *in silico* sequences, characterizing two sequences, *SEP1/2* and *SEP3* gene, demonstrating that previous analyses using bioinformatics are efficient tools for molecular studies. Moreover, the expression analyses, performed by the qRT-PCR, detected the expression only in flower and fruit tissues, corroborating with the results obtained by the electronic *Northern* and also with expression analyses of *SEP* genes in other species. Thus, these analyses allowed the identification and characterization of two coffee *SEP* genes, however further experiments are necessary to confirm and better characterization of these genes.

1 INTRODUÇÃO

Desde o isolamento do primeiro fator de transcrição *MADS-box* em plantas, várias funções biológicas para esses genes têm sido descritas em várias espécies (Ng & Yanofsky, 2001; Portereiko et al., 2006). Os mecanismos moleculares responsáveis pela especificação dos órgãos florais são realizados por genes da família *MADS-box*, pelo modelo ABCE, que possuem cinco classes de genes homeóticos para a formação desses órgãos.

Os genes *SEPALLATA* (*SEP*) são exemplos desses genes homeóticos e pertencem à classe E, sendo necessários para a formação de todos os verticilos. Em *Arabidopsis thaliana* são caracterizados quatro genes dessa classe: *SEP1*, *SEP2*, *SEP3* e *SEP4*, que possuem função redundante na determinação floral. Mas como é estabelecido pelo modelo ABCE, eles não são requeridos sozinhos, cada proteína *SEP* forma um complexo multimérico com outros fatores de transcrição *MADS-box* para formar os verticilos florais e contribuir no desenvolvimento dos frutos.

Desta forma, a expressão desses genes geralmente é encontrada em flores e frutos. Tanto a expressão quanto a função são dependentes do tipo de *SEP* analisado e principalmente da espécie estudada, porque mesmo atuando de forma redundante possuem expressões diferentes nos diversos tecidos em espécies distintas. Porém, em algumas espécies foram encontradas expressões em tecidos vegetativos (Malcomber & Kellogg, 2005), que mesmo sendo considerada baixa, não deixa de ser indício de uma possível função desses genes em tecidos vegetativos, além dos órgãos reprodutivos.

Para esses estudos de expressão gênica, várias técnicas são utilizadas, e com o desenvolvimento da técnica de PCR quantitativo em tempo real (qRT-PCR), o campo da expressão de genes foi transformado, principalmente pela elevada sensibilidade e especificidade que a técnica apresenta.

Devido a importância dos genes *SEP* na formação da flor e suas variações de função e expressão nas diferentes espécies, é de fundamental importância estudos iniciais de caracterização desses genes, uma vez que o florescimento em café é um processo bastante complexo. Desta forma, o objetivo desse trabalho foi isolar, clonar, sequenciar e comprovar a expressão tecida específica dos genes *SEP* identificados pela análise *in silico*, realizando uma caracterização parcial desses genes. Além de definir os íntrons e éxons através da comparação das sequências de DNA e cDNA.

2 MATERIAL E MÉTODOS

2.1 Material vegetal

O material vegetal utilizado foi coletado no campo experimental da Universidade Federal de Lavras (UFLA) em Lavras- MG. Foram utilizados quatro tecidos com três repetições biológicas, sendo cada repetição representada por uma planta, de *Coffea arabica*, cultivar Rubi. Os tecidos foram coletados no período de setembro de 2008: folhas (jovens), flor (estádios G4-G5 descritos por Moraes et al. (2008), fruto (em estádios diferentes de desenvolvimento) e raiz (secundárias). Armazenados em freezer -80° C.

Todos os procedimentos de extração, clonagem, sequenciamento e qRT-PCR foram realizados no Laboratório Central de Biologia Molecular (LCBM) da UFLA.

2.2 Extração de DNA

Para extração do DNA foi utilizado o protocolo CTAB descrito por Doyle & Doyle (1991) com modificações, e extraído apenas do tecido foliar de café. As folhas foram maceradas em nitrogênio líquido, transferido 1 g de tecido para microtubos e acrescentado 800 µL do tampão de extração (2% CTAB, 1,4 M NaCl, 0,2% 2-mercaptoetanol, 20 mM EDTA, 100 mM Tris-HCl- pH 8,0). As amostras foram agitadas no vórtex e permaneceram no banho-maria a 65° C por 1 h, homogeneizadas ocasionalmente. Após 5 min à temperatura ambiente, foram adicionados às amostras 600 µL de clorofórmio/álcool isoamílico (24:1) (v:v) para desproteíntização e agitados levemente por 10 min. Após centrifugado a 13.000 g por 5 min à temperatura ambiente, o sobrenadante foi transferido para novos tubos e adicionados novamente clorofórmio/álcool isoamílico na mesma quantidade anterior. Agitados e centrifugados por 5 min, o sobrenadante obtido foi transferido novamente para novos tubos e acrescentado 900 µL de

isopropanol gelado para precipitação do DNA, no freezer a 20° C por no mínimo 30 min. Após esse período foram centrifugadas por 5 min, descartado o sobrenadante e adicionado 300 µL de etanol 70% gelado para lavagem do DNA. Centrifugado por 5 min e descartado o sobrenadante, o *pellet* formado foi seco à temperatura ambiente e ressuspendido em 50 µL água Mili-Q autoclavada. Para obtenção de amostras puras de DNA foram acrescentados 2,5 µL de RNase (100 µg/mL) em banho maria a 37° C por 30 min. As amostras foram precipitadas novamente para eliminação da enzima a partir da adição do isopropanol, e estocadas no freezer a -20° C.

As amostras de DNA foram aplicadas em gel de agarose 0,8% corado com brometo de etídio para analisar a integridade do material e realizadas as leituras no espectrofotômetro NanoDrop® Espectrophotometer ND-1000 para relação de absorvância A_{260}/A_{280} nm com o objetivo de analisar a pureza das amostras.

2.3 Extração de RNA

Para isolamento do RNA foram utilizados como tecido vegetal, folha, flor, raiz e fruto de café. O método de extração, Borato Quente, foi adaptado do protocolo descrito por Birtic & Kranner (2006), onde o tampão de extração é formado de Sodium tetraborate decahydrate 0,2 M; EGTA - ethylene glycol bis (2-aminoethyl ether) -N,N,N',N'-tetraacetic acid 30 mM; SDS - sodium dodecyl sulfato 1%; Sodium deoxycholate 1%, ajustado para pH 9.0 com hidróxido de sódio. Para cada 1 mL de tampão, foi adicionado 0,06 g de polyvinylpyrrolidone (PVP) e 2% de dithiothreitol (DTT), e aquecidos a 80° C para dissolução completa dos reagentes. Os tecidos foram macerados em nitrogênio líquido, e a cada 100 mg de tecido foram acrescentados 1 mL de tampão de extração. Imediatamente, foram adicionados 1 mg de Proteinase K, agitados e colocados em banho maria a 45° C por 1 h 30 min. Posteriormente, foi adicionado 80 µL

de Cloreto de Potássio 2 M e as amostras foram incubadas no gelo por 45 min. Em seguida, foram centrifugadas por 20 min a 4 °C (12.000 g) e recolhido o sobrenadante. Foi acrescentado 1/3 do volume, correspondente ao sobrenadante, de Cloreto de Lítio 8 M (gelado) e as amostras foram incubadas no freezer a -20 °C por, no mínimo, 1 h para precipitação do RNA. Após esse período, foram centrifugadas por 20 min a 4 °C (12.000 g), e o sobrenadante descartado. O *pellet* formado foi lavado com 750 µL de Cloreto de Lítio 2 M pelo menos duas vezes. Novamente as amostras foram centrifugadas por 10 min, o líquido residual foi removido do tubo com a pipeta e o RNA foi ressuspenso em 20 µL de água Milli-Q autoclavada. As amostras foram armazenadas a -20°C.

Em seguida as amostras foram tratadas com DNase DNA free da Ambion. Em uma reação contendo 10 µg de RNA total, 1 U da enzima rDNaseI, 0,1 volume do tampão DNaseI 10X e água livre de DNase/RNase para completar o volume de 50 µL. Essa reação foi incubada a 37 °C por 30 min aproximadamente e adicionada 5 µL da solução para inativar DNase (DNase inactivation). As amostras foram incubadas por 2 min, centrifugadas a 10.000 g por 1,5 min e transferido o sobrenadante para novos tubos.

2.4 Síntese cDNA

A síntese de cDNA foi realizada a partir do RNA total dos tecidos de café, utilizando o Kit Super Script III First – Strand Synthesis Super Mix da Invitrogen. Padronizando 1 µg de RNA, foram feitas as sínteses de cDNA de todos os tecidos utilizando *primer Oligo(dT)*. Para um volume final de 8 µL, foram acrescentados a 1 µg de RNA, 1 µL de Primer Oligo(dT), 1 µL do tampão e água livre de DNase/RNase. Essas amostras foram incubadas no termociclador Eppendorf Mastercycler gradient a 65 °C por 5 min e imediatamente colocadas no gelo por 1 min. Após um *spin*, foram adicionados às amostras, 10 µL do tampão 2X First-Strand Reaction Mix e 2 µL da enzima SuperScript III. A

reação de transcriptase reversa, para um volume final de 20 µL, foi realizada no termociclador, em duas etapas: a 50 °C por 50 min para a ação da enzima, e 85 °C por 5 min para inativá-la. As amostras foram conservadas no freezer a -20° C até o uso.

2.5 Desenho dos primers

A partir das prováveis sequências *SEPALLATA* de café obtidas do estudo *in silico*, foram desenhados os *primers*, para clonagem e para análise de qRT-PCR (Tabelas 1 e 2), utilizando o programa “OligoPerfect™ Designer” (Invitrogen). Como gene controle para análise qRT-PCR, foi escolhido o GAPDH, segundo Barsalobres-Cavallari et al. (2009), (Tabela 2).

TABELA 1 *Primers* utilizados na clonagem dos genes *SEPALLATA* (C1, C8 e C14).

GENE	SEQUÊNCIA DO PRIMER	<i>Amplicon (pb)</i>
C1 F	5' CACCATGGGTAGGGGGAG 3'	327
C1 R	5'CTAAATTATCGTCAGAATCACTGTACA 3'	327
C14SEP3 1F	5' CACCATGGGAAGAGGTAGG 3'	726
C14SEP3 1R	5' TCATGGGAACCAACCTGC 3'	726
C8SEP4 1F	5' CACCATGGGAAGGGGTAAA 3'	857
C8SEP4 1R	5' TTATGCCTGTTGGCCCAA 3'	857

TABELA 2 *Primers* utilizados na análise de qRT-PCR.

GENE	SEQUENCIA DO PRIMER	<i>Amplicon</i> (<i>pb</i>)
qC1 F	5' GCAGTTATGGAGCAGTGGAA 3'	101
qC1 R	5' CGTCAGAATCACTGTACAAGCA 3'	101
qC14SEP3 F	5' CCAGATTGGATCCCAGAATG 3'	104
qC14SEP3 R	5' CCTGCCATGGAGTTATTCACA 3'	104
qC8SEP4 F	5' TTGGAAGGCCAACAAATACG 3'	102
qC8SEP4 R	5' CCTGTTGGCCCAAGAACATA 3'	102
qGapdh F	5' TTGAAGGGCGGTGCAA A 3'	59
qGapdh R	5' AACATGGGTGCATCCTTGCT 3'	59

2.6 Amplificação dos fragmentos e eluição das bandas do gel

As amostras de DNA e cDNA foram amplificadas utilizando uma enzima de alta fidelidade que possui atividade de exonuclease 3' - 5', a Tli DNA polymerase (Promega), e *primer* específicos (“forward” e “reverse”) para os três genes *SEP* identificados em café (Tabela 1).

Para a amplificação dos fragmentos, foi realizada para cada gene, uma reação de volume final de 50 μ L, onde foram adicionados 5 μ L tampão 10 X *buffer* (sem MgCl₂), 25 mM MgCl₂ (2 mM concentração final), 10 mM dNTP mix (0.2 mM final), *primer* específicos para os genes *SEPALLATA* (“forward” e “reverse”) (1 μ M final), e 0.2 μ L Tli DNA polymerase (5.0 u).

A reação foi submetida ao termociclador Eppendorf Mastercycler gradient, programado com temperatura de desnaturação inicial de 94 °C, por 2 min, seguidos de 38 ciclos de amplificação compostos de 3 etapas: 94 °C, por 20 seg; 60 °C, por 50 seg; 72 °C, por 1 min. Após os 34 ciclos, as amostras foram submetidas a uma etapa final de 10 min a 72 °C.

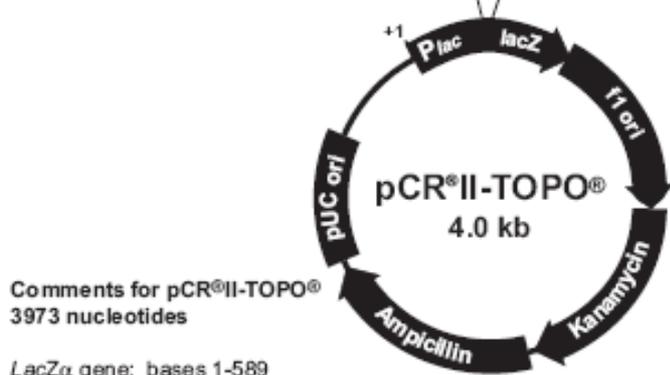
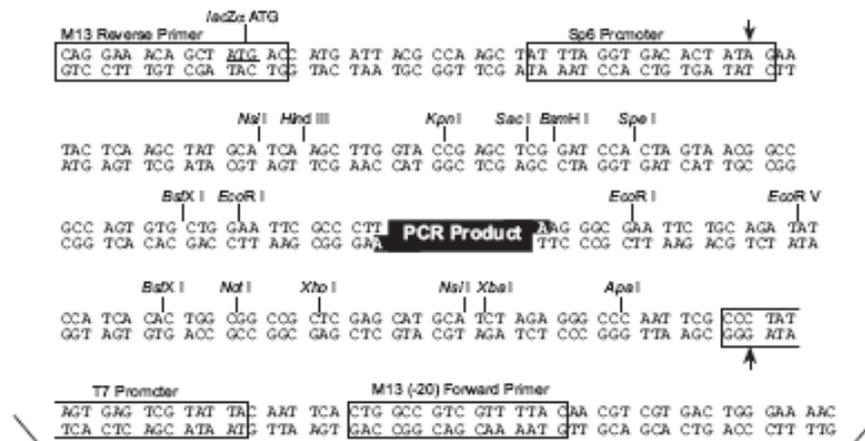
Os produtos das ampliações foram submetidos à eletroforese em gel de agarose 1,2% (m/v) corado com 2 μ L de Brometo de Etídio (0,5 μ g/mL), sob

corrente elétrica de 110 V em Tampão TAE (0,001 M EDTA pH 8,0; 0,04 M TRIS pH 8,0; 0,02 M ácido acético), por um período de aproximadamente 45 min, sendo visualizado sob luz ultravioleta e a imagem captada pelo fotodocumentador EDAS 290 (Kodak®).

Os fragmentos obtidos foram eluídos do gel de agarose por meio do Kit MinElute (QUIAGEN®), segundo as recomendações 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 min ou até a diluição completa do gel, agitando-se no vórtex a cada 3 min, durante a incubação. Foi adicionado 1 volume de isopropanol à amostra, agitando cuidadosamente várias vezes o microtubo. A amostra foi aplicada em uma coluna MinElute inserida dentro de um microtubo de 2 mL e centrifugada, a 16.000 g por 1 min, sendo o filtrado posteriormente descartado. Foram adicionados 500 µL do tampão QG e em seguida centrifugado por 1 min. Para lavar os fragmentos, foram adicionados 750 µL do tampão PE à coluna que foi centrifugada por 1 min. O filtrado foi descartado e a coluna foi centrifugada novamente a 16.000 g por 1 min. A coluna foi colocada em um tubo limpo de 1,5 mL e, para eluir o cDNA, foram adicionados 10 µL do tampão EB (10 mM Tris-HCl, pH 8,5) no centro da membrana. A coluna foi deixada em repouso por 2 min e centrifugada por 1 min. O produto da eluição foi armazenado em freezer a -20 °C.

2.7 Clonagem e transformação bacteriana

Para a clonagem foi utilizado o vetor *pCR[®]II TOPO TA Cloning* da Invitrogen (Figura 1). Os produtos da eluição foram adicionados ao vetor, na proporção de 1:1 e acrescentados 1 µL solução de sal e água para um volume final de 6 µL. A reação foi incubada por 20 min à temperatura ambiente e colocadas no gelo para a utilização na transformação bacteriana.



Comments for pCR^{II}-TOPO[®]
 3973 nucleotides

LacZα gene: bases 1-589
 M13 Reverse priming site: bases 205-221
 Sp6 promoter: bases 239-256
 Multiple Cloning Site: bases 269-383
 T7 promoter: bases 406-425
 M13 (-20) Forward priming site: bases 433-448
 f1 origin: bases 590-1027
 Kanamycin resistance ORF: bases 1361-2155
 Ampicillin resistance ORF: bases 2173-3033
 pUC origin: bases 3178-3851

FIGURA 1 Mapa do vetor pCR^{II} TOPO TA Cloning (Invitrogen)

Para a transformação foram utilizadas as células da bactéria *Escherichia coli* DH5αTM-T1^R quimicamente competente (Invitrogen). A essas células foram adicionados 2 μL da reação de clonagem com o vetor *TOPO TA*, e incubados no

gelo por 30 min. Para a introdução do vetor na bactéria foi utilizado o processo de choque térmico que consistiu em colocar a 42 °C por 30 s e imediatamente transferir os tubos para o gelo. Foram acrescentados 250 µL do meio S.O.C e deixados por 1 h a 37 °C em constante agitação a 200 g. Logo após a incubação, o plaqueamento foi realizado utilizando 40 e 60 µL da solução de transformação em placa de Petri, contendo 20 mL de meio de cultura Luria-Bertani (LB) + ágar, 50 µg/mL ampicilina e 40 mg/mL de X-Gal (5-bromo-4-cloro-3-indolil-D-galactosídeo).

As placas foram mantidas a 37 °C por aproximadamente 24 h, para permitir o crescimento das colônias bacterianas. Após esse período, 10 colônias brancas possivelmente transformadas foram selecionadas aleatoriamente, transferidas a outra placa para subcultivo com o mesmo meio, e submetidas à PCR de colônia, para comprovar a presença do fragmento de interesse. Para isso, com auxílio de ponteiras esterilizadas, 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 para cada colônia, foi realizada em um volume final de 20 µL contendo tampão da enzima 10 X, 25 mM MgCl₂ (2 mM concentração final), 10 mM dNTP mix (0.2 mM final), os *primers SEP* correspondentes a cada gene (1µM final de cada), *Paq DNA polymerase* (5.0 U) e água destilada. A reação foi submetida ao termociclador Eppendorf Mastercycler gradient, programado com temperatura de desnaturação inicial de 94 °C, por 1 min, seguidos de 34 ciclos de amplificação compostos de 3 etapas: 94 °C, por 1 min; 60 °C, por 1 min; 72 °C, por 1 min. Após os 34 ciclos, as amostras foram submetidas a uma etapa final de 5 min 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.

As placas do subcultivo foram colocadas a 37 °C por aproximadamente 24 h para crescimento das colônias brancas isoladas. Após esse crescimento e a partir da confirmação da inserção do fragmento pela PCR, as colônias foram transferidas para meio líquido, possibilitando o crescimento das células, que serão utilizadas para a extração de DNA plasmidial e posterior sequenciamento. Com o auxílio de palitos esterilizados, cada colônia foi transferida para 1 mL de meio LB (Lauria Bertani) suplementados com 50 µg/mL ampicilina. Esse procedimento foi realizado em placa com 96 poços, que foi mantida a 37 °C, sob agitação, durante 22 h, em agitador orbital ajustado para 300 g.

2.8 Miniprep

O miniprep foi realizado utilizando as 10 colônias selecionadas de cada placa, e para facilitar, o procedimento foi realizado em placas Deepwell de 96 poços. Após o crescimento das colônias, a placa foi centrifugada por 6 min a 3700 g, para sedimentação das células. O sobrenadante foi removido e a placa foi seca por 5 min. A cada poço foi adicionado 240 µL de tampão GET (20% de glicose, 0,5 M de EDTA e 1 M de Tris-HCl) e a placa foi agitada em vórtex por 2 min para a ressuspensão das células. Após essa etapa a placa foi centrifugada por 10 min a 3700 g, descartado o sobrenadante e novamente seca por 5 min. Foi adicionado novamente 80 µL do tampão GET, agitada em vórtex. Em uma microplaca de 250 µL foi adicionado 2,5 µL de RNase em cada poço, e todo o material ressuspendido anteriormente foi acrescentado nessa placa. Posteriormente, em cada poço foi adicionado 80 µL de NaOH 0,2 N, homogeneizado por inversões suaves, e incubado por 10 min à temperatura ambiente. Após a adição de 80 µL de KOAc 3 M, a placa foi novamente homogeneizado por inversões, incubada à temperatura ambiente por 10 min e aquecida em estufa a 90 °C por 30 min. A placa foi submetida a resfriamento no gelo por 10 min e centrifugada a 20 °C , por 10 min a 3700 g. O conteúdo foi

transferido para uma placa Milipore com filtro, centrifugando uma contra a outra com os poços correspondentes, a 20 °C por 6 min e a 3700 g, até que todo o conteúdo tenha passado para a outra placa. Foi adicionado 100 µL de isopropanol ao filtrado e homogeneizado por inversões suaves. A placa foi centrifugada a 20 °C por 45 min a 3700 g e descartado o sobrenadante. Posteriormente foi adicionado 200 µL de etanol 70%, gelado, centrifugado por 5 min, a 3700 g, a 20 °C, novamente retirado o sobrenadante, e o *pellet* foi seco a temperatura ambiente por 30 min. Finalmente as amostras foram ressuspensas em 60 µL de água mili-Q autoclavada para posterior análise de sequenciamento.

2.9 Sequenciamento

Os plasmídeos contendo os insertos de DNA e cDNA, dos fragmentos correspondentes aos *contigs* C1, C8 e C14, assim como os mesmos fragmentos eluídos diretamente do gel de agarose sem a inserção em vetores (*amplicons*), foram utilizados para o sequenciamento. Todos esses materiais foram sequenciados no MegaBace 1000 (Amersham Biosciences), utilizando o método enzimático baseado na síntese de DNA *in vitro*, na presença de nucleosídeos trifosfatados terminadores de cadeia (Sanger et al., 1977).

Para a amplificação dos fragmentos inseridos no vetor, foi utilizado um *primer* universal (M13) que anela à sequência do *TOPO TA* utilizado para a clonagem, e para os *amplicons* foram usados os *primers* correspondentes a cada *contig* (Tabela1). O sequenciamento foi realizado com o kit “Big Dye Terminator” (Applied Biosystems) segundo as especificações do fabricante. Para a reação de sequenciamento, foram adicionados a uma microplaca de 96 cavidades, 1 µL de cada plasmídeo (na concentração máxima de 350 ng/µL), 4 µL 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), e 5 picomoles de *primer forward* e *reverse* M13. Para a reação com

os *amplicons*, foram diferencialmente utilizados 3 μL do fragmento (na concentração aproximada de 20 $\text{ng}/\mu\text{L}$) e os *primers* específicos para cada sequência. A reação foi realizada em termociclador Eppendorf Mastercycler gradient, utilizando-se o seguinte programa: 30 ciclos de 95 °C, por 20 seg; 50°C, por 15 seg e 60 °C, durante 1 min. Os produtos amplificados na reação de sequenciamento foram precipitados adicionando 1 μL de Acetato de Amônio (7,5 mol.L-1) em cada amostra. Posteriormente, foram adicionados aproximadamente 30 μL de etanol 95%, que corresponde a 1/3 do volume. A placa foi vedada com adesivo, agitada manualmente por inversão, permanecendo à temperatura ambiente por 15 min. Após a centrifugação, por 45 min, a 4.000 g, 20 °C, os sobrenadantes foram descartados e os sedimentos foram lavados com 100 μL de etanol 70%. A placa foi centrifugada, por 15 min, a 4.000 g, 20 °C. Os sobrenadantes foram descartados e a placa foi novamente centrifugada de forma invertida sobre papel absorvente (1.000 g, 1 min) para remover o excesso de etanol, sendo, em seguida, mantidas a 37 °C, por 10 min, para secagem.

Posteriormente, foram adicionados 10 μL de *loading solution* (kit “Big Dye Terminator” - Applied Biosystems) 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 seg; voltagem de corrida: 6 Kv e tempo de corrida: 240 min. As sequências foram determinadas utilizando-se o programa *Sequence Analyzer*.

2.10 Análises das sequências

Para pesquisa de similaridade, as sequências de DNA e cDNA obtidas dos fragmentos isolados de café foram comparadas e alinhadas com as previamente obtidas *in silico*, utilizando o *ClustalW*, 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).

2.11 Análise quantitativa da expressão dos genes

A análise da expressão gênica quantitativa foi realizada através do ABI PRISM 7500 Real-Time PCR (Applied Biosystems), pelo sistema de detecção SYBR Green. Os cDNAs sintetizados após a extração do RNA dos tecidos de café, pelo método descrito no item 2.3, foram utilizados como *template* para a análise. Os *primers* foram desenhados a partir das sequências obtidas pela análise *in silico* (Tabela 2), utilizando a parte terminal como molde, para a amplificação de pequenos fragmentos, geralmente entre 60 a 120 pb, maximizando a eficiência da amplificação.

Primeiramente foi realizado um ensaio de quantificação absoluta, onde foi determinada a curva padrão, a eficiência dos *primers* e a melhor diluição para as amostras testadas. Os cDNA foram diluídos 1:10, 1:25 e 1:50.

Após determinar a diluição das amostras de 1:10 e a eficiência dos *primers* de 94% a 97%, foi realizado o ensaio de quantificação relativa. O gene utilizado como controle endógeno foi baseado em Barsalobres-Cavallari et al. (2009), que sugeriu o gene GAPDH como um gene referência em café por apresentar um nível constante de expressão nas condições propostas no trabalho.

As condições térmicas da reação foram 2 min a 50 °C, 10 min a 95 °C, seguidos por 40 ciclos de 15 seg a 95 °C e 1 min a 60 °C, e finalizando com 15 seg a 95 °C. Os dados foram coletados e armazenados no programa 7500 Fast Software (Versão 2.1). Para cada reação, foram utilizados 1 µL cDNA, 1,5 µM de cada *primer* e 5,0 µL de Master Mix SYBR Green (Applied) para um volume final 10,0 µL/amostra.

As amostras foram processadas em triplicatas e realizadas três repetições biológicas. 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*, que foi determinada manualmente. O método usado foi o C_T comparativo.

Para o cálculo da expressão, primeiramente cada amostra foi normalizada com o controle endógeno, utilizando a equação $\Delta C_T = C_T$ (gene alvo) - C_T (controle endógeno) e a quantificação relativa foi obtida pela fórmula $2^{-\Delta\Delta C_T}$.

3 RESULTADOS E DISCUSSÃO

Para uma possível caracterização dos genes *SEP* em café, foram desenhados *primers* a partir das sequências identificadas pela análise *in silico*. Esses *contigs* possuíam os seguintes tamanhos: *contig1* identificado como possível *SEP1/2* com 327 pb (C1SEP1/2), *contig14* (*SEP3*) com 726 pb (C14SEP3), e o *contig8* (*SEP4*) com 857 pb (C8SEP4). Primeiramente, por meio da análise da PCR utilizando como molde o cDNA de flores de café, foram obtidas bandas correspondentes aos tamanhos especificados (Figura 2). Entretanto, quando são analisadas as bandas correspondentes ao DNA de café, podemos observar que os tamanhos dos fragmentos geralmente são maiores que os fragmentos correspondentes de cDNA. Isso pode ser explicado pelo fato do DNA possuir em sua extensão, sequências que são removidas durante o processamento do mRNA, denominadas íntrons. Desta forma, como o cDNA é sintetizado a partir do mRNA, ele também não possui essas sequências, podendo apresentar o tamanho um pouco menor comparado ao DNA. Esse fato pode ser melhor exemplificado, ao observar a Figura 2B, que representa os *amplicons* derivados do *primer* C14SEP3, onde o fragmento derivado de um DNA possui aproximadamente 1000 pb, e o derivado de um cDNA possui o tamanho esperado da análise *in silico*, de aproximadamente 700 pb.

Após a eluição das prováveis sequências *SEP* de café do gel de agarose, a clonagem e a transformação bacteriana, foi realizado o ‘miniprep’ e posterior sequenciamento desses fragmentos. Para o sequenciamento, além dos fragmentos ligados ao vetor, foram utilizados também os *amplicons* derivados diretamente do gel de agarose, sem a inserção em vetores. Porém, foi obtido somente o sequenciamento completo para alguns desses *amplicons*. Para os fragmentos que foram clonados no vetor, o processo de sequenciamento não

obteve sucesso. Desta forma, foram sequenciados os genes correspondentes ao *contig1* e *contig14*, derivados de cDNA de flores de café.

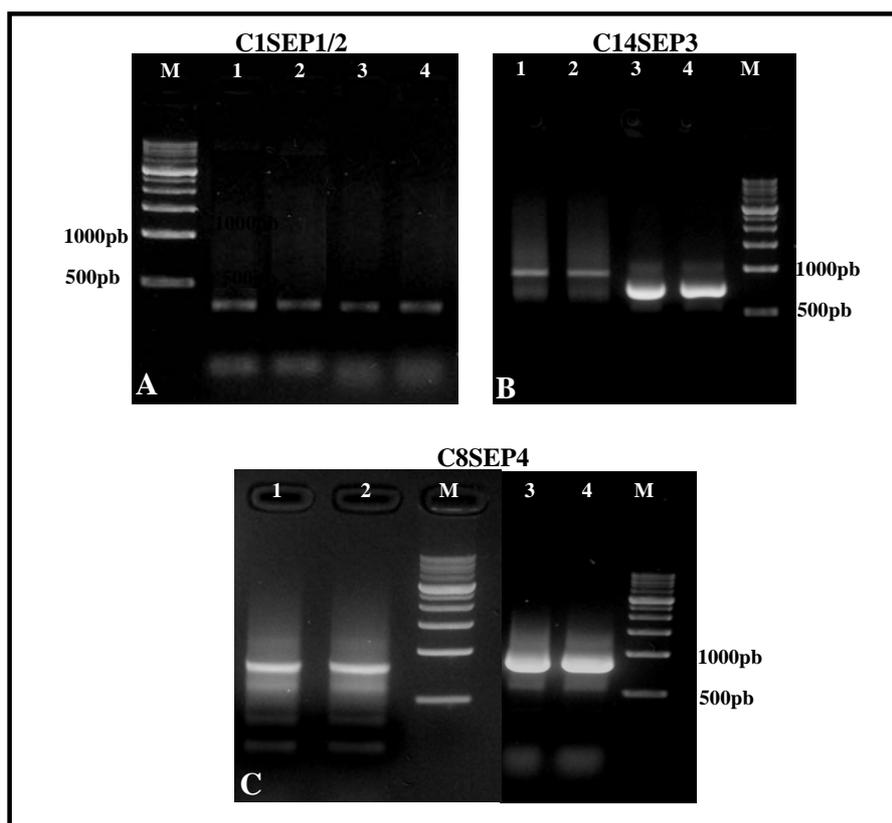


FIGURA 2 Gel de agarose 1,2% com produtos de amplificação de amostras de DNA (1-2) e cDNA (3-4) utilizando os *primers SEP*. Marcador Labor 1Kb (M). A – amplificação de fragmentos com *primer C1SEP1/2*; B – amplificação dos fragmentos com *primer C14SEP3*; C - amplificação dos fragmentos com *primer C8SEP4*.

Essas sequências foram alinhadas com as sequências obtidas na análise *in silico*. Como pode ser observado na Figura 3, o alinhamento obtido da sequência *in silico* do *contig14* com o C14SEP3 derivado do sequenciamento

representou um alinhamento perfeito, porém como o sequenciamento foi obtido a partir de *amplicons* derivados diretamente do gel, onde foram usados os *primers* específicos do gene, o início da sequência não foi detectado. Resultado semelhante pode ser observado para a sequência do C1SEP1/2 (Figura 4). Desta forma é de extrema importância repetir o sequenciamento com os fragmentos clonados em vetores, para a obtenção dos fragmentos completos, e para o uso desses fragmentos em trabalhos posteriores de transformação de plantas a fim de determinar a real função desses genes em café.

Na comparação com sequências depositadas no banco de dados do *NCBI*, foi encontrado um valor de identidade da sequência do *contig14*, de 85% com *SEP3* de *Lycopersicon esculentum*, e 80% com a de *Antirrhinum majus*. Para a comparação da sequência representante do *contig1*, foi estabelecida uma identidade de 77% entre *SEP1* de *Cucumis sativus*, *Citrus* e de *Antirrhinum majus*. Estes resultados corroboram com os encontrados nas árvores filogenéticas e nas comparações realizadas com as sequências *in silico* apresentadas no Capítulo 2, onde é determinado homologia também entre essas sequências.

Para os *contig1* e *contig14* foi possível obter uma caracterização parcial de suas sequências. Ao observar os resultados obtidos no Capítulo 2 e a confirmação dessas sequências nesse capítulo, pode-se afirmar que provavelmente a sequência apresentada como *contig1* é um representante do *SEP1* e *SEP2*, e o *contig14* representante do *SEP3* em café arábica. Entretanto, esses resultados têm que ser confirmados com estudos de mutantes para esses genes. Como não foi obtido o sequenciamento completo do *contig8*, não foi possível a confirmação dessa sequência.

Nesse trabalho foi possível analisar uma grande similaridade com os *contigs* obtidos através da análise *in silico*, com as sequências dos fragmentos obtidos pelo sequenciamento, demonstrando a eficiência da análise genômica

feita previamente usando bioinformática, através de bancos de dados, que se constitui numa ferramenta muito poderosa para estudos moleculares.

```

CLUSTAL W (1.81) multiple sequence alignment

C14SEP3_cDNA      -----GGGGAGGATGAGACAG-----
contig14          ACGGAAAGAGAGGTAGGAGAGAGGATGGGAAGAGGTAGGGTGGAGCTAAAGAGGATAGAG
                    **, .***;***. **

C14SEP3_cDNA      -----ATCATAGGCAAG-TACTTTTGCTAAGCGAAGAAATGGGCTATTGAAGAAAGCT
contig14          AACAGATCAATAGGCAAGTAACTTTTGCTAAGCGAAGAAATGGGCTATTGAAGAAAGCT
                    : .***** :*****

C14SEP3_cDNA      TATGAGCTCTCAGTCTAATGTGATGCTGAGGTTGCTTTGATCATCTTCTCCAATAGAGGA
contig14          TATGAGCTCTCAGTCTAATGTGATGCTGAGGTTGCTTTGATCATCTTCTCCAATAGAGGA
                    *****

C14SEP3_cDNA      AAGCTCTATGAGTTTTGCAGTAGCTCTAGCATGGTCAAGACGCTAGAGAGGTATCAGAAA
contig14          AAGCTCTATGAGTTTTGCAGTAGCTCTAGCATGGTCAAGACGCTAGAGAGGTATCAGAAA
                    *****

C14SEP3_cDNA      TGCAACTATGGAGCACCTGAGCCCAACATATCCACCAGGGAAGCACTGGAGCTGAGTAGT
contig14          TGCAACTATGGAGCACCTGAGCCCAACATATCCACCAGGGAAGCACTGGAGCTGAGTAGT
                    *****

C14SEP3_cDNA      CAGCAGGAGTATTTGAAGCTTAAAGCACGTTACGAAGCTCTACAAAGATCACAAGGAAT
contig14          CAGCAGGAGTATTTGAAGCTTAAAGCACGTTACGAAGCTCTACAAAGATCACAAGGAAT
                    *****

C14SEP3_cDNA      CTTTGGGCGAGGACCTTGGTCCTTTAAACAGCAAGGAACTTGAATCATTGGAAAGGCAG
contig14          CTTTGGGCGAGGACCTTGGTCCTTTAAACAGCAAGGAACTTGAATCATTGGAAAGGCAG
                    *****

C14SEP3_cDNA      CTTGATATGTCTCTGAAGCAAATCAGATCAACACGGACTCAAGTAATGCTGGATCAGCTC
contig14          CTTGATATGTCTCTGAAGCAAATCAGATCAACACGGACTCAAGTAATGCTGGATCAGCTC
                    *****

C14SEP3_cDNA      ACTGATCTCCAGAGAAAGGAACATGCTTTGAACGAGGCAAACTCTAAAACAAAGG
contig14          ACTGATCTCCAGAGAAAGGAACATGCTTTGAACGAGGCAAACTCTAAAACAAAGG
                    *****

C14SEP3_cDNA      CTGATGGAAGGAAACCAAGTAAATCTCCAGTGAATCCGAATGCACAGGATGTTGGGTAC
contig14          CTGATGGAAGGAAACCAAGTAAATCTCCAGTGAATCCGAATGCACAGGATGTTGGGTAC
                    *****

C14SEP3_cDNA      GGGCGACAACCAGCTCATGCTCAGGGGTGATGGTTTTCTTCATCCATTGGATTGTGAACC
contig14          GGCCGACAACCAGCTCATGCTCAGGG-TGATGGTTTTTTCATCCATTGGATTGTGAACC
                    ** *****

C14SEP3_cDNA      GACGATTCACAGATTGGATACCAGAATGATNCAATAACAGTGGCAGCGGCAGGCCGAGT
contig14          GCC--ATTCCAGATTGGATCCAGAATGATCCAATACCAGTGGCAGCGGCAGGCCGAGT
                    *. * .*****.***** *****

C14SEP3_cDNA      GTGAATAACTACAIGGCAGGTGG--TCCCAA--
contig14          GTGAATAACTCCAIGGCAGGTGGTCCCATGA
                    *****.***** * *****;

```

FIGURA 3 Alinhamento pelo programa ClustalW (Thompson et al., 1994), do produto do sequenciamento (C14SEP3 cDNA) e a sequência identificada *in silico* contig14.

```

CLUSTAL W (1.81) multiple sequence alignment

C1SEP1/2_cDNA      -----AGTAGCTGAAGAAATAAGCACAGATAATNGNGCTGGTTACG
contig1            ATGGGTAGGGGAGAGTGGAGCTGAAGAGAATAGAGAACAAGATAAATAGGCAGTTACG
                    ** :*:*:*:*:* *.....*****:..**:*:*****

C1SEP1/2_cDNA      TATTGCTAAGAGGAGGAATGGACTGCTCAINGATAGCTATATGATGCTCTCAGATACTGT
contig1            -TTTGCTAAGAGGAGGAATGGACTGCTCAAGAAAG---CTTATGAGCTCTCAG-ITCTGT
                    ;*****:..*:. .*: :***** *:*:****

C1SEP1/2_cDNA      GTGACNCTGAGAGCTAGGCCCTCATCCTCTTCTCTAACCAGGGGCAAGCTTTATGAGTTCT
contig1            GTGATGCTGAGS--TGGCCCTCATCATCTTCTTAACCAGGGGCAAGCTTTATGAGTTCT
                    **** .*****. :*****.*****.*****.*****.*****

C1SEP1/2_cDNA      GTAGCAGCTCCAACATGCTCAGGACTCTTGAAAGGTACCAAAGATGCAGTTATGGAGCAG
contig1            GTAGCAGCTCCAACATGCTCAGGACTCTTGAAAGGTACCAAAGATGCAGTTATGGAGCAG
                    *****

C1SEP1/2_cDNA      TGGAAAGTCAGCCACTCAGCCAAAGAGATTGAGGTAACCTACCATAATTAAGCATCACTTGG
contig1            TGGAAAGTCAGCCACTCAGCCAAAGAGATTGAGGTAACCTACCATAATTAAGCATCACTTGG
                    *****

C1SEP1/2_cDNA      CCCATGCTTGTACAGTGATTCTGA-----
contig1            CCCATGCTTGTACAGTGATTCTGACGATAATTAG
                    *****

```

FIGURA 4 Alinhamento pelo programa ClustalW (Thompson et al., 1994) do produto do sequenciamento (C1SEP1/2 cDNA) e a sequência identificada *in silico* contig1.

Para a análise de qRT-PCR primeiramente foi definida pelo ensaio da curva padrão, a diluição de 1:10 como a diluição padrão para a análise de quantificação relativa com os *primers* analisados. A expressão relativa dos prováveis genes *SEP* em café, demonstrou uma alta expressão em flores e frutos, como pode ser observado na Figura 5, onde os valores de expressão foram normalizados pelo controle endógeno GAPDH. O padrão de expressão entre todos os genes foi relativamente o mesmo e para todos foi observado uma maior expressão no tecido floral, do que em fruto. Entretanto essa diferença não pode ser considerada significativa segundo os valores do desvio padrão.

Esses resultados podem estar associados às funções descritas para essa subfamília, pois como são responsáveis pela determinação do meristema floral e desenvolvimento dos frutos, possuem consequentemente uma expressão nesses

tecidos. Esses resultados corroboram com vários trabalhos de expressão desses genes em outras espécies (Kanno et al., 2006; Matsubara et al., 2008).

Em plantas de trigo, foi observado por análise de *Northern blot*, que o candidato ao gene *SEP* (*TaMADS1*) somente foi expresso em órgãos florais primordiais, nos quatro verticilos florais e no óvulo do trigo, e não foram encontrados nenhum sinal em raiz, caule, folhas e sementes (Zhao et al., 2006). Já Matsunaga et al. (2008), analisaram a expressão dos genes *SEP* por qRT-PCR em *Silene latifolia*, utilizando flores masculinas e femininas, e observaram que os dois genes identificados *SiSEP1* e *SiSEP3* foram expressos em ambos botões florais.

Xu et al. (2008) estudaram a expressão dos genes *SEP* em vários tecidos de pêssogo, e observaram a não expressão em folhas, mas como esperado, a expressão estava presentes nos órgãos reprodutivos, entretanto diferentes padrões de expressões foram observados em frutos e flores.

O que muitas vezes pode variar é o padrão de expressão dos diferentes genes *SEP*, nos diferentes verticilos florais. O *SEP3* em *Arabidopsis* é expresso em flores primordiais e nos três verticilos florais, mas o transcrito não pode ser detectado nos meristemas florais. Resultados similares são observados para os genes *SEP3* e outros *SEP* em distintas espécies de plantas, como o *FBP2* em petúnia, o *TaMADS1* em trigo, o *TM5* em tomate e o *OsMADS8* em arroz (Angenent et al., 1992; Pnueli et al., 1994; Grego et al., 1997; Mandel & Yanofsky, 1998; Pelucchi et al., 2002; Zhao et al., 2006). A partir dos resultados de todos esses trabalhos, pode-se afirmar que esses padrões de expressão são diversos nas diferentes espécies e com os vários *SEP*.

Uma expressão pouco característica pode ser observada referente ao gene *C8SEP4* (Figura 5), onde foi identificada uma pequena expressão no tecido foliar, podendo relacionar esse gene a uma possível função em folhas de café. Resultados semelhantes corroboram com essa hipótese, descrevendo a expressão

de um provável gene *SEP* de soja (*GmSEPI*) no caule, o que pode indicar uma possível participação desses genes em tecidos vegetativos, além de tecidos florais e fruto (Huang et al., 2009). Em gérbera o gene homólogo ao *SEP* (*GhGRCD2*) também foi expresso em tecidos vegetativos (Malcomber & Kellogg, 2005) e uma pequena expressão também foi encontrada em folhas de arroz e sorgo (Malcomber & Kellogg, 2004). Essa expressão é considerada baixa, mesmo assim, não deixa de ser indício de uma possível função desses genes em tecidos vegetativos.

Estudos subsequentes são necessários para realizar uma caracterização completa de todos esses possíveis genes *SEP* de café, de acordo com suas funções e localização celular, utilizando para isso, estudos com mutantes e localização *in situ*.

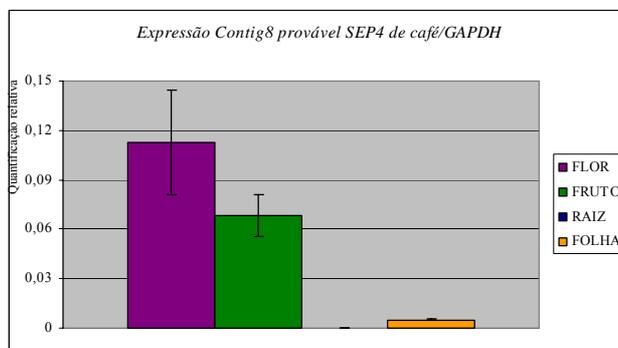
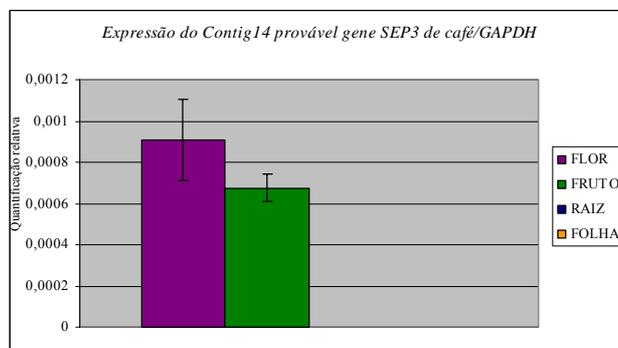
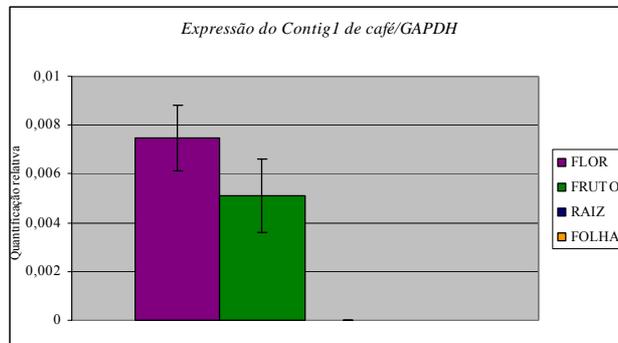


FIGURA 5 Perfil da expressão quantitativa relativa por qRT-PCR dos genes *SEP*. As colunas representam a expressão de transcritos gênicos em diferentes tecidos de plantas de café, com valores dados a partir da média dos valores de expressão das réplicas biológicas; flor (roxa), fruto (verde), raiz (azul), folha (laranja). A abundância de transcritos gênicos foi normalizada pela expressão dos gene *GAPDH*.

4 CONCLUSÕES

A partir das análises realizadas foi possível a caracterização parcial das sequências de dois genes *SEP* em café arábica, identificados anteriormente em análises *in silico*.

Não foi possível a identificação de íntrons e éxons através da comparação entre as sequências de DNA e cDNA.

Os genes *SEP* apresentam um padrão de expressão característico em flores e frutos de plantas de café arábica.

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**, Oxford, v.25, n.17, p.3389-3402, Sept. 1997.

ANGENENT, G.C.; BUSSCHER, M.; FRANKEN, J.; MOL, J.N.M.; TUNEN, A.J. van. Differential expression of two MADS box genes in wild-type and mutant petunia flowers. **Plant Cell**, Rockville, v.4, n.8, p.983-993, Aug. 1992.

BARSALOBRES-CAVALLARI, C.F.; SEVERINO, F.E.; MALUF, M.P.; MAIA, I.G. Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions. **BMC Molecular Biology**, Bethesda, v.10, n.1, p.1-11, Jan. 2009.

BIRTIC, S.; KRANNER, I. Isolation of high-quality RNA from polyphenol, polysaccharide-and lipid-rich seeds. **Phytochemical Analysis**, Chichester, v.17, n.3, p.144-148, Mar. 2006.

DOYLE, J.J.; DOYLE, J.L. Isolation of plant DNA from fresh tissue. **Focus**, Rockville, v.1, n.12, p.13-15, Feb. 1991.

GRECO, R.; STAGI, L.; COLOMBO, L.; ANGENENT, G.C.; SARI-GORLA, M.; PE, M.E. MADS box genes expressed in developing inflorescences of rice and sorghum. **Molecular Genetics Genetic**, Ribeirão Preto, v.253, n.5, p.615-623, Feb. 1997.

HUANG, F.; CHI, Y.J.; GAI, J.Y.; YU, D.Y. Identification of transcription factors predominantly expressed in soybean flowers and characterization of GmSEP1 encoding a SEPALLATA1-like protein. **Gene**, Amsterdam, v.438, n.1/2, p.40-48, June 2009.

KANNO, A.; HIENUKI, H.; ITO, T.; NAKAMURA, T.; FUKUDA, T.; YUN, P.Y.; SONG, I.J.; KAMIMURA, T.; OCHIAI, T.; YOKOYAMA, J.; MAKI, M.; KAMEYA, T. The structure and expression of SEPALLATA-like genes in *Asparagus* species (Asparagaceae). **Sexual Plant Reproduction**, New York, v.19, n.3, p.133-144, Sept. 2006.

MALCOMBER, S.T.; KELLOGG, E.A. Heterogeneous expression patterns and separate roles of the SEPALLATA gene LEAFY HULL STERILE1 in grasses. **Plant Cell**, Rockville, v.16, n.7, p.1692-1706, June 2004.

MALCOMBER, S.T.; KELLOGG, E.A. SEPALLATA gene diversification: brave new whorls. **Trends in Plant Science**, Oxford, v.10, n.9, p.427-435, Sept. 2005.

MANDEL, M.A.; YANOFSKY, M.F. The Arabidopsis AGL9 MADS box gene is expressed in young flower primordia. **Sexual Plant Reproduction**, New York, v.11, n.1, p.22-28, Feb. 1998.

MATSUBARA, K.; SHIMAMURA, K.; KODAMA, H.; KOKUBUN, H.; WATANABE, H.; BASUALDO, I.L.; ANDO, T. Green corolla segments in a wild Petunia species caused by a mutation in FBP2, a SEPALLATA-like MADS box gene. **Planta**, Berlin, v.228, n.3, p.401-409, Aug. 2008.

MORAIS, H.; CARAMORI, P.H.; KOGUISHI, M.S.; RIBEIRO, A.M.A. Escala fenológica detalhada da fase reprodutiva de coffea arabica. **Bragantia**, Campinas, v.67, n.1, p.257-260, jan. 2008.

NG, M.; YANOFSKY, M.F. Function and evolution of the plant MADS-box gene family. **Nature Reviews Genetics**, London, v.2, n.3, p.186-195, Mar. 2001.

PELUCCHI, N.; FORMARA, F.; FAVALLI, C.; MASIERO, S.; LAGO, C.; COLOMBO, L.; KATE, M.M. Comparative analysis of rice MADS-box genes expressed during flower development. **Sexual Plant Reproduction**, New York, v.15, n.3, p.113-122, Oct. 2002.

PNUELI, L.; HAREVEN, D.; BRODAY, L.; LIFSCHITZ, E. The *TM5* [*tdr5*] MADS-box gene mediates organ differentiation in the three inner whorls of tomato flowers. **The Plant Cell**, Rockville, v.6, n.2, p.175-186, Feb. 1994.

PORTEREIKO, M.F.; LLOYD, A.; STEFFEN, J.S.; PUNWANI, J.A.; OTSUGA, D.; DREWS, G.N. AGL80 is required for central cell and endosperm development in arabidopsis. **The Plant Cell**, Rockville, v.18, n.8, p.1862-1872, June 2006.

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

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**, Oxford, v.22, n.22, p.4673-4680, Nov. 1994.

XU, Y.; ZHANG, L.; XIE, H.; ZHANG, Y.Q.; OLIVEIRA, M.M.; MA, R.C. Expression analysis and genetic mapping of three SEPALLATA-like genes from peach (*Prunus persica* (L.) Batsch). **Tree Genetics & Genomes**, Heidelberg, v.4, n.4, p.693-703, Oct. 2008.

ZHAO, X.Y.; CHENG, Z.J.; ZHANG, X.S. Overexpression of TaMADS1, a SEPALLATA-like gene in wheat, causes early flowering and the abnormal development of floral organs in Arabidopsis. **Planta**, Berlin, v.223, n.4, p.698-707, Apr. 2006.