



CARLOS HENRIQUE CARDON

**IDENTIFICAÇÃO DE POLIMORFISMOS NO
FLOWERING LOCUS T (FT) E SUA EXPRESSÃO
GÊNICA EM DIFERENTES ESPÉCIES DO
GÊNERO *COFFEA*.**

**LAVRAS - MG
2016**

CARLOS HENRIQUE CARDON

**IDENTIFICAÇÃO DE POLIMORFISMOS NO *FLOWERING LOCUS T*
(*FT*) E SUA EXPRESSÃO GÊNICA EM DIFERENTES ESPÉCIES DO
GÊNERO *COFFEA*.**

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

Prof. Antonio Chalfun Júnior, PhD
Orientador
Prof. Dr. Horllys Gomes Barreto
Coorientador

**LAVRAS – MG
2016**

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Cardon, Carlos Henrique.

Identificação de polimorfismos no *FLOWERING LOCUS T*
(*FT*) e sua expressão gênica em diferentes espécies do gênero
Coffea. / Carlos Henrique Cardon. – Lavras : UFLA, 2016.
78 p. : il.

Dissertação (mestrado acadêmico)—Universidade Federal de
Lavras, 2016.

Orientador(a): Antonio Chalfun Júnior, PhD.

Bibliografia.

1. Cafeeiro. 2. Florescimento. 3. *FLOWERING LOCUS T* (*FT*).
4. Região promotora. 5. Polimorfismo. I. Universidade Federal de
Lavras. II. Título.

CARLOS HENRIQUE CARDON

**IDENTIFICAÇÃO DE POLIMORFISMOS NO *FLOWERING LOCUS T*
(*FT*) E SUA EXPRESSÃO GÊNICA EM DIFERENTES ESPÉCIES DO
GÊNERO *COFFEA*.**

**IDENTIFICATION OF POLYMORPHISMS IN *FLOWERING LOCUS T*
(*FT*) AND ITS GENE EXPRESSION IN DIFFERENT *COFFEA* GENUS
SPECIES.**

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

APROVADO EM 29 de Julho de 2016

Antonio Chalfun Júnior, PhD (UFPA/DBI)
Dr. Horllys Gomes Barreto (UFT)
Dr. Raphael Ricon de Oliveira (UCO, ESPANHA)

Prof. Antonio Chalfun Júnior, PhD
Orientador

Prof. Dr. Horllys Gomes Barreto
Coorientador

**LAVRAS – MG
2016**

Dedico a Deus por me conceder a vida e dar-me a graça de poder usufruir desse bem tão maravilhoso. A minha família, especialmente a minha mãe e ao meu irmão que sempre me deram força e acreditaram em mim, a minha namorada que também sempre esteve ao meu lado me apoiando. E todos aqueles que contribuíram de alguma forma para a conclusão deste trabalho.

DEDICO!

AGRADECIMENTOS

Agradeço a Deus por conceder o dom da vida e da sabedoria e me dar saúde e forças para vencer as batalhas da vida.

A minha família que sempre estiveram ao meu lado, acreditaram em mim e sempre me apoiaram especialmente a minha mãe Eneida e ao meu irmão Claudio que contribuíram grandemente para que eu pudesse chegar até aqui.

Agradeço a minha namorada Maysa, que sempre esteve ao meu lado me dando forças, me apoiando e se fazer presente não só nos momentos de felicidades, mas também nas horas que precisei de ajuda.

A Universidade Federal de Lavras especialmente ao Departamento de Biologia, ao Setor de Fisiologia Vegetal e ao Laboratório de Fisiologia Molecular de Plantas que me possibilitaram desenvolver esse trabalho e realizar meu mestrado.

Aos professores que no decorrer da minha vida me passaram seus conhecimentos me capacitando para que pudesse chegar até aqui..

Aos membros da banca de avaliação que dispuseram de seu tempo para poder acrescentar mais ao meu trabalho com seus conhecimentos.

Ao professor Antonio Chalfun Júnior por sempre se fazer presente e dispor a sanar dúvidas, ter acreditado em mim e além de orientador ser um amigo.

Ao professor Horllys Gomes Barreto pela coorientação e amizade, por ter ajudado durante esse tempo tirando dúvidas e ensinando.

Aos meus amigos de infância e da graduação que de certa forma contribuíram para o meu crescimento e fazem uma parte fundamental da minha vida.

Aos meus amigos fora do grupo de pesquisa que adquiri após a minha chegada em Lavras e que fizeram parte dessa conquista.

A todos os amigos membros e ex-membros do LFMP que de forma direta e indireta contribuíram para que eu pudesse concluir esse trabalho.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela bolsa de estudo, ao INTCcafé e a FAPEMIG pelos incentivos financeiros para a realização deste trabalho e apoio.

E a todos que de alguma forma contribuíram para a realização e conclusão deste trabalho.

MUITO OBRIGADO!

RESUMO

Devido à grande importância econômica e social do café, principalmente para o Brasil por ser o maior produtor e exportador do mundo, é objetivo de todos entenderem melhor e aprimorar as técnicas de manejo a fim de melhorar a qualidade do produto e reduzir custos na produção. Um dos maiores problemas encontrado pelos produtores na cafeicultura é a florada sequencial e desuniforme do cafeeiro, agregando custos elevados para a produção da bebida com parâmetros de boa qualidade. O aprimoramento dos conhecimentos em genes envolvidos na floração pode ser a chave para o entendimento dessa característica do florescimento do cafeeiro que apesar de vários estudos, até hoje não se tem definido a principal causa desse fenômeno. O gene *FLOWERING LOCUS T* (*FT*), é um dos fundamentais durante a indução do florescimento em várias espécies. Polimorfismo na região promotora do gene pode ser determinante para a alteração do comportamento da sua expressão, uma vez que na região promotora existem sequências motivos de ligação de fatores de transcrição fundamentais na regulação da expressão. Com o objetivo de entender melhor a ação do gene *FT* sobre o florescimento de cafeeiro, ao comparar as regiões promotoras de genótipos representantes das espécies *Coffea arabica*, *Coffea canephora*, *Coffea eugenioides* e *Coffea racemosa*, encontramos vários motivos dentro das sequências em um comprimento de aproximadamente 1kb. Alguns dos motivos encontrados são de ligação de elementos *cis*-reguladores ligados ao florescimento que de certa forma atuam sobre o controle da expressão do gene *FT* em cafeeiro. Apesar de não ter tido relação da expressão do gene *FT* com as características de florescimento dos genótipos estudados, a identificação de motivos de ligação de fatores de transcrição controlados por estímulos ligados a indução do florescimento em cafeeiro, mostra que o gene *FT* pode ter uma papel fundamental no entendimento da florada desuniforme do cafeeiro.

Palavras-Chave: Cafeeiro. Florescimento. Polimorfismo gênico.

ABSTRACT

Due to the great economic and social importance of coffee, mainly to Brazil, the largest producer and exporter in the world, a better understanding of management techniques is of great importance in order to improve product quality and reduce production costs. One of the biggest problems faced by coffee growers is the sequential and asynchronous flowering of coffee trees, which increases production costs for the production of a good quality product. A better knowledge of the genes involved in flowering can be the key to the understanding the feature of the sequential flowering found in coffee trees, which although several studies have been conducted to improve the knowhow on the phenomenon, little is known so far. about its main causes. The gene *FLOWERING LOCUS T (FT)* is one of the key genes during the induction of flowering in various species. Polymorphism in the promoter region of the genes may be decisive in changing the behavior of their expression, since on their promoter regions are located binding motifs sequences of transcription factors essential in the regulation of expression. In order to better understand the action of FT on coffee flowering, this study aimed to compare the putative polymorphisms of the FT promoter region, as well as its expression pattern, in four different coffee species, *Coffea arabica*, *Coffea canephora*, *Coffea eugenioides* and *Coffea racemosa*, with different flowering features. Some of the motifs found were from binding regions of cis-regulatory elements linked to flowering regulation, which may act on the control of *FT* expression in coffee trees. Although it was not found a relationship between the *FT* expression pattern and the flowering features of the genotypes analyzed in this study, the identification of transcription factor binding motifs controlled by stimuli related to coffee flowering induction, shows that coffee *FT* may play a fundamental role in understanding the asynchronous coffee flowering.

Key-Words: Coffee tree. Flowering. Gene polymorphism.

SUMÁRIO

	CAPÍTULO 1.....	10
1	INTRODUÇÃO GERAL.....	11
2	REFERÊNCIAL TEÓRICO.....	14
2.1	O Cafeeiro: Aspectos morfológicos e econômicos.....	14
2.2	Fenologia do cafeeiro.....	15
2.3	O florescimento e o gene <i>FT</i>.....	19
2.4	Região promotora de um gene.....	22
2.5	Fatores de transcrição.....	24
2.6	Polimorfismo gênico.....	26
2.7	Características das Espécies.....	28
2.7.1	<i>Coffea arabica</i>.....	28
2.7.2	<i>Coffea canephora</i>.....	29
2.7.3	<i>Coffea eugenioides</i>.....	29
2.7.4	<i>Coffea racemosa</i>.....	30
3	CONSIDERAÇÕES GERAIS.....	31
4	REFERÊNCIAS.....	32
	CAPITULO 2.....	45
5	INTRODUÇÃO.....	44
6	MATERIAL E MÉTODOS.....	46
6.1	Material vegetal.....	46
6.2	Extração de DNA e RNA.....	46
6.3	Tratamento com DNAase e Síntese de cDNA.....	48
6.4	Análises <i>in-silico</i>.....	49
6.5	Desenho de <i>primers</i>.....	51
6.6	Isolamento da região promotora e codante do gene <i>FT</i>.....	53
6.7	Sequenciamento.....	53
6.8	Análise de polimorfismo na região promotora do gene <i>FT</i>.....	54

6.9	Análise da expressão gênica por RT-qPCR.....	54
7	RESULTADOS E DISCUSSÃO.....	56
7.1	Análise <i>in-silico</i>	56
7.1.1	Identificação da sequência codante do gene <i>FT</i>	56
7.1.2	Busca da região promotora do gene <i>FT</i>	57
7.2	Isolamento das sequências codantes e promotoras do gene <i>FT</i> ...	58
7.3	Sequências <i>cis</i> -reguladoras.....	59
7.4	Polimorfismo da região promotora do gene <i>FT</i>	63
7.5	Sequências codantes.....	66
7.6	Análise da expressão gênica.....	69
8	CONCLUSÕES.....	72
9	CONCLUSÃO FINAIS.....	73
10	REFERÊNCIAS.....	74

CAPÍTULO 1

Aspectos socioeconômicos e características morfológicas do cafeeiro

1. INTRODUÇÃO GERAL

O café é uma das principais commodities do mundo, sendo a segunda mercadoria mais comercializada no mundo após o petróleo (CONAB, 2016). Dentre as mais de 100 espécies já identificadas, somente as espécies *Coffea arabica* e *Coffea canephora* são cultivadas com fins comerciais. Originado no continente africano (Etiópia), com o tempo o cafeeiro foi sendo distribuído para o resto do mundo até chegar ao Brasil, onde encontrou condições favoráveis para seu cultivo. Atualmente, o Brasil é o maior produtor e exportador de café do mundo (CONAB, 2016), se tornando uma das culturas mais importantes para o país.

O cafeeiro tem características peculiares de desenvolvimento, tais como, ciclo bianual e florescimento sequencial. Um dos maiores desafios na produção do café está na maturação desuniforme dos frutos, decorrente do seu florescimento sequencial durante seu estágio reprodutivo (ECCARDI & SANDALJ, 2003). A maturação desuniforme dos frutos acarreta na agregação de custos para se produzir uma bebida de boa qualidade, pois exigem uma etapa de separação de frutos no estágio apropriado, ou seja, de cereja como descrito por MORAIS et al., (2008). Estes processos podem levar a aumentos no custo do produto dificultando o seu consumo para grande parte da população (ALVES & LIVRAMENTO, 2003).

O florescimento do cafeeiro sofre influência de muitos fatores abióticos formando uma rede complexa de interação genética que dificultam uma compreensão mais precisa. O florescimento é induzido pelas primeiras chuvas, ou pelo retorno da irrigação, após um período de déficit hídrico, o qual parece ser necessário para aquisição de competência das gemas florais (ALVIM, 1960). Dessa forma, em cultivos irrigados, alguns estudos mostram que o momento e a extensão, ou uniformidade, do florescimento podem ser controlados de acordo com o a intensidade do déficit hídrico ao qual as plantas são submetidas e o

momento de retorno da irrigação (GUERRA et al., 2005; DA SILVA et al., 2009)

Até o momento, pouco se sabe sobre o controle molecular do florescimento do cafeeiro (BARRETO et al., 2012; DE OLIVEIRA et al., 2014). O ambiente influencia a expressão e atividade de certos genes que, por sua vez, regulam o florescimento na tentativa de garantir a melhor época para reprodução. Entre esses genes, *FLOWERING LOCUS T (FT)*, tem um papel muito importante para ativação do mecanismo molecular do florescimento, fundamental para regulação de outros genes responsáveis pela formação dos órgãos florais. *FT* é controlado principalmente pelo acúmulo da proteína CO (CONSTANS), por sua vez regulada pelo comprimento do dia. (SAMACH et al., 2000; ABE et al., 2005). As horas de luz regulam a expressão do gene CO na folha que após atingir um limite tal ativa a expressão de *FT*. O mRNA de *FT* é translocado para os meristemas onde regulam a expressão do gene *SOC1* e *API* iniciando o processo de formação dos órgãos florais (AMASINO, 2010). Com isso a planta é capaz de perceber as horas de do dia e ajustar à época mais adequada para florescimento.

Polimorfismos no DNA decorrente de mutações durante o processo evolutivo das espécies, ou por exposição a elementos mutagênicos, apresentam grande influência na regulação da expressão gênica nos organismos. O tipo de mutação mais comum são os SNPs (*Single Nucleotide Polymorphism*), os quais estão presentes em várias regiões do DNA, alterações em regiões maiores ocorrem com menos frequência, mas de grande importância também. Esses polimorfismos, quando localizados na região promotora dos genes, podem interferir de forma significativa na regulação da expressão dos mesmos (ADRIAN et al., 2010). Na região promotora, existem sequências motivos às quais fatores de transcrição se ligam e controlam de forma positiva ou negativa a expressão gênica. Alterações nessas sequências podem impedir a ligação correta

dos fatores de transcrição afetando diretamente, ou até mesmo impedido, a sua ação (NOTAGUCHI et al., 2008).

Levando em consideração o problema enfrentado pelos produtores de café quanto ao florescimento sequencial do cafeeiro, a necessidade de se entender melhor quais os fatores que levam a essa característica, a importância do gene *FT* na participação da rota de indução do florescimento e os motivos no qual pode afetar sua expressão, esse trabalho teve como objetivo comparar possíveis polimorfismos na região promotora do gene *FT*, assim como sua expressão, em quatro espécies de café, *Coffea arabica*, *Coffea canephora*, *Coffea eugenioides* e *Coffea racemosa*.

2. REFERENCIAL TEÓRICO

2.1 O Cafeeiro: Aspectos morfológicos e econômicos

Pertencente a família *Rubiaceae*, o gênero *Coffea* conta com mais de 100 espécies já registradas, sendo que duas delas, *Coffea arabica* (Arabica) e *Coffea canephora* (Robusta/Conilon), se destacam por apresentarem importância econômica. O cafeeiro possui porte arbustivo, sistema radicular pivotante, ciclo bianual com florada sequencial, e ramos plagiotrópicos de crescimento indeterminado (BRIDSON & VERDCOURT, 1988; ECCARDI & SANDALJ, 2003).

Apesar de não ser originário do Brasil, o cafeeiro se adaptou bem as condições ambientais que o país apresenta (ECCARDI & SANDALJ, 2003), e hoje o país ocupa o posto de maior produtor, exportador e segundo maior consumidor do mundo, atrás apenas dos Estados Unidos (CONAB, 2016). No Brasil a área total cultivada com café corresponde a 2.248.565,8 hectares, sendo que 79,2% são da espécie *C. arabica* e o restante da espécie *C. canephora*. A região Sudeste é a maior produtora de café do país, predominando a espécie *C. arabica*. O estado de Minas Gerais se destaca por ser o maior produtor de café do Brasil com 67,8% da área de *C. arabica* cultivada no país, seguido do Espírito Santo, São Paulo, Bahia, Paraná, Rondônia e Goiás. (CONAB, 2016).

A produção nacional de café para 2016 está estimada entre 49.126,1 e 51.943,9 mil sacas beneficiadas de café. A área produtiva para todo o país será de 1.977,5 mil hectares. Com base na bialidade do cafeeiro, esse ano para a maioria dos estados haverá aumento da produção principalmente para o estado de Minas Gerais nas regiões Sul, Sudeste e Cerrado. O único estado que apresentará bialidade negativa será o estado do Paraná resultado de uma geada

rigorosa no ano de 2015 que prejudicou o ciclo normal da planta (CONAB, 2016).

O cafeeiro é uma planta com altas exigências nutricionais e climáticas e necessita de um bom manejo para atingir uma boa produção, desenvolve bem em solos profundos e porosos com boa permeabilidade por possuir raízes profundas e espalhadas (ECCARDI & SANDALJ, 2003). A altitude é determinante para a qualidade e produção de algumas espécies como *C. arabica* que se destaca em regiões entre 800 a 1200 metros de altitude, podendo variar de cultivar para cultivar. Já a espécie *C. canephora* se desenvolve melhor em altitudes menores, abaixo de 800 metros (CARVALHO et al., 1991). O cafeeiro se desenvolve bem em temperaturas entre 10 °C a 27 °C, não suportando temperaturas superiores a 37 °C ou inferiores a 5 °C, o que ocasionam um desequilíbrio fisiológico da planta (CARVALHO et al., 1991; ALVES & LIVRAMENTO, 2003).

O ciclo de vida do cafeeiro é bastante complexo, despertando grande interesse e necessidade de um melhor entendimento morfológico, fisiológico e genético do seu desenvolvimento. Um dos principais desafios da cafeicultura nacional e mundial é a florada seqüencial do cafeeiro, que resulta na maturação desuniforme dos frutos e afeta de forma significativa a qualidade da bebida, tornando-a indesejável ao consumidor ou de baixa classificação no mercado, além de agregar custos na produção por se fazer necessário mais de um evento de colheita para colher todo o café (MELO & SOUSA, 2011).

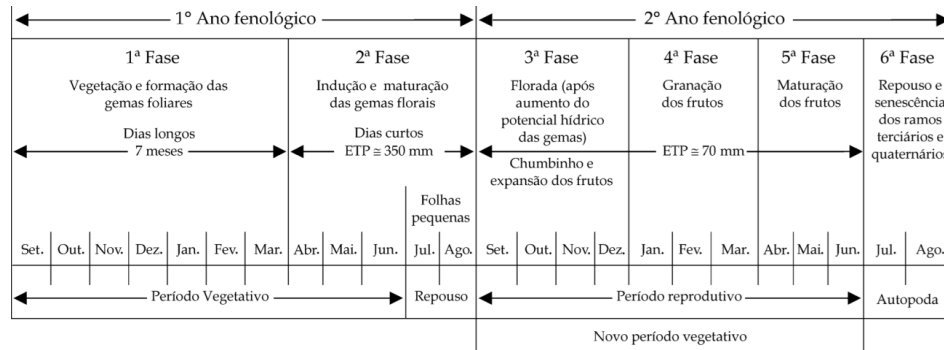
2.2 Fenologia do cafeeiro

O cafeeiro completa seu ciclo fenológico em dois anos, contados no início do estágio vegetativo ao término do estágio reprodutivo, determinando-o como bianual. Durante a bienalidade do cafeeiro, o momento que precede a

reprodução e os estádios durante esse período, deixa várias interrogações devido ao seu tipo de florescimento. (GOUVEIA, 1981). O ciclo fonológico do cafeeiro é dividido em seis fases, desde o desenvolvimento dos ramos até a maturação dos frutos. Um estudo sobre a fenologia do cafeeiro resultou numa tabela que divide cada um dos estádios durante o ciclo da planta (CAMARGO & CAMARGO, 2001).

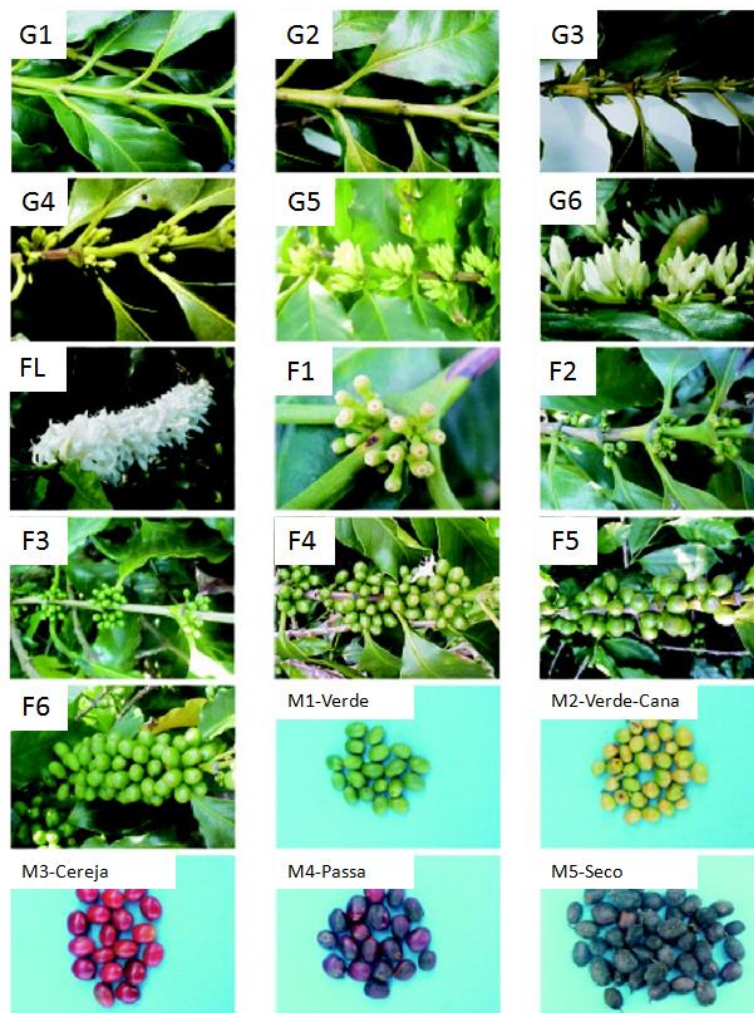
O primeiro ano é caracterizado pelo desenvolvimento dos ramos plagiotrópicos a partir dos ramos ortotrópicos. A primeira fase, que compreende os meses de setembro a março, tem início a partir da formação das gemas axilares nos nós dos ramos plagiotrópicos primários. Portanto, o tamanho e a qualidade da safra no ano seguinte é impactada diretamente pela quantidade de flores que a planta será capaz de produzir a partir dos ramos e gemas desenvolvidos no ano anterior (CAMARGO; CAMARGO, 2001; LIVRAMENTO, 2010; THOMAZIELLO et al., 2000;). Na segunda fase, abril a agosto, a ocorrência de dias curtos estimula a indução floral transformando gemas axilares (vegetativas), produzidas na fase anterior, em gemas florais. Nos meses finais desta fase (julho- agosto) as gemas iniciam um período de repouso (dormência), permanecendo nesse estado por semanas até o início da próxima fase no segundo ano fenológico (CAMARGO; CAMARGO, 2001; CAMARGO; FRANCO, 1985).

FIGURA 1 – Esquemática das seis fases fenológica de cafeeiro arabica, durante 24 meses, nas condições climáticas tropical do Brasil (CAMARGO & CAMARGO, 2001).



A fase reprodutiva é influenciada por diversos fatores endógenos e exógenos, sendo que alguns deles em particular, como temperatura e estresse hídrico, são determinantes para a formação dos frutos e determinação do tempo de abertura das flores (MELO & SOUSA, 2011). Assim como ocorre a divisão das fases do ciclo fenológico do cafeeiro, também há uma divisão das fases reprodutivas do cafeeiro, da formação das gemas florais até a maturação dos frutos, a fim de facilitar o entendimento e avaliações quando se estuda os estádios reprodutivos do cafeeiro. As fases são divididas pelo desenvolvimento das flores, dos frutos e maturação dos frutos (Fig. 2) (MORAIS et al., 2008).

FIGURA 2 - Escala das fases reprodutivas em *C. arabica* em que: 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); F1 - até 3 mm de comprimento; F2 – 3,1 a 4 mm; F3 – 4,1 a 5 mm; F4 – 5, 1 a 10 mm; F5 – 10 a 15 mm; F6 – maior que 15 mm (MORAIS et al., 2008).



2.3 O florescimento e o gene *FT*

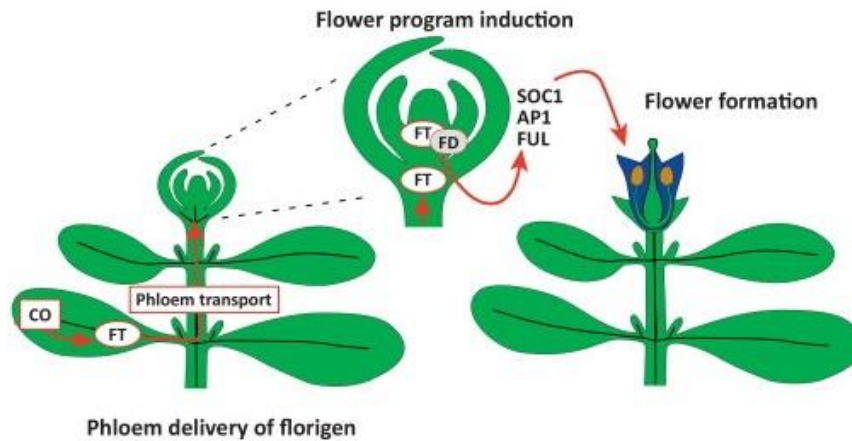
O florescimento é iniciado pela transição do meristema vegetativo para o reprodutivo e termina com a formação e abertura das flores (antese). Essa transição é controlada por múltiplas rotas genéticas influenciadas direta e indireta por fatores exógenos e endógenos e coordenam o desenvolvimento dos órgãos florais (LEVY & DEAN, 1998; ANDRÉS & COUPLAND, 2012). A ausência ou a não suficiência de elementos e condições como nutrientes, água, temperatura, luminosidade e outros exigidos pela planta para o desenvolvimento reprodutivo completo, afetam o ciclo fenológico da planta, inibindo ou alterando o processo de indução floral (ABU-ASAB et al., 2001).

Durante esse estágio de transição de fases, várias interações gênicas estão associadas à regulação da expressão dos genes responsáveis por esse mecanismo. Em plantas modelo como *Arabidopsis*, o florescimento é controlado através de seis rotas já conhecidas: idade, vernalização, giberelina (GA), temperatura, fotoperíodo-dependente e vias autônomas (FORNARA et al., 2010). Estudos mostraram a associação de genes de forma direta e indireta às fases do florescimento. Dentre esses genes o *FLOWERING LOCUS T (FT)* é um dos genes fundamentais para a indução indireta do florescimento (VARDY et al., 2005). A proteína do *FT* é um dos componentes do sistema de sinalização florigênico da planta que se transloca por longa distância através do floema até o meristema onde desencadeia sua função (CORBESIER et al., 2007; LIN et al., 2007).

A expressão do gene *FT* é estreitamente regulada e central no desenvolvimento fisiológico que define o início do florescimento. Essa é uma característica conservada em várias espécies de plantas (WIGGE et al., 2005; TURCK et al., 2008). Conhecido desde o início como um dos principais genes envolvido na rota do florescimento, hoje existem vários estudos, em que o

descreve como responsável por diferente mecanismo ou ação em algumas espécies como promotor do florescimento (WICKLAND & HANZAWA, 2015). Ele pode reprimir o florescimento também, como exemplo em beterraba que existem o *FT1* que reprime e o *FT2* que induz o florescimento (PIN et al., 2010). O entendimento do balanço da atividade dos elementos florígenos e ante-florígenos é muito importante para o crescimento geral, arquitetura e florescimento da planta. Isso foi elucidado pela primeira vez pelos genes *FT* e *TFL1-like* em tomates modelo (SHALIT et al., 2009; LIFSCHITZ et al., 2014; CAO et al., 2015).

FIGURA 3 – Esquema adaptado de (MCGARRY & KRAGLER, 2013) da ação do gene *FT* sobre a indução do florescimento controlado pela proteína CO em *Arabidopsis thaliana*. O *FT* é regulado pelo CO na folha e a proteína FT é translocada pelo floema até o meristema apical, onde interage com a proteína FD e ativa a expressão dos genes de formação dos órgãos florais.



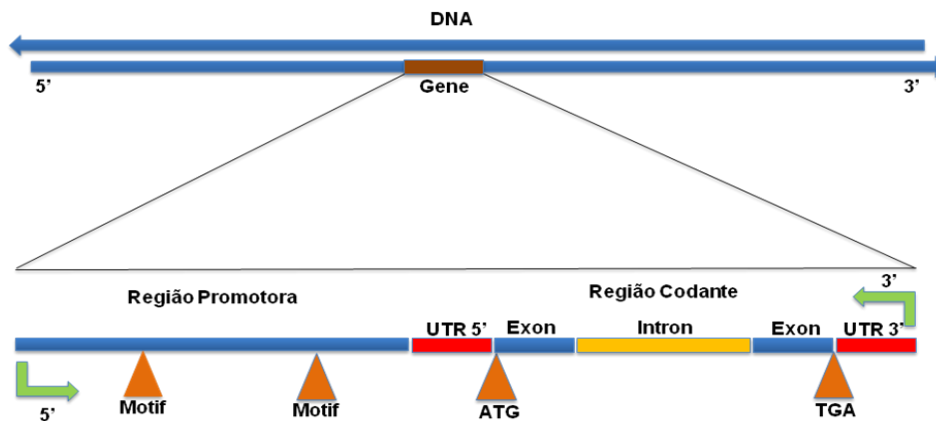
Tanto o gene *FT* quanto outros genes envolvidos na indução e formação dos órgãos florais, são controlados por fatores de transcrição fundamentais para a determinação da quantidade de transcritos e do momento da transcrição desses genes (ADRIAN et al., 2010). O gene *CONSTANS* (*CO*) é considerado um dos principais reguladores do gene *FT* (Fig.3), por sua vez regulado por outros fatores de transcrição. (ABE et al., 2005; WIGGE et al., 2005). Entretanto, existem diferenças nessa regulação entre espécies, por exemplo, a regulação do *FT* pela proteína CO não se aplica a todas as espécies, em *Arabidopsis* CO regula *FT* em dias longos enquanto em Arroz a homóloga CO (Hd3), em dias curtos (PUTTERILL et al., 2004). Além disso, genes *CO-like* em alguns casos não controlam a expressão do *FT* em plantas como

ervilhas e Medicago (PUTTERILL et al., 2013; WONG et al., 2014; WELLER & ORTEGA, 2015).

2.4 Região promotora de um gene.

Entendido como uma sequência mínima necessária do DNA para que haja a transcrição de um RNA, que pode ser ou não traduzido em uma cadeia polipeptídica e formar uma proteína para uma determinada função no organismo. Assim, o gene é a informação que codifica para formação das partes do organismo e dos elementos regulatórios do metabolismo envolvido no desenvolvimento e manutenção da vida (WATSON, 1970). A sequência do DNA que compreende um gene é dividida basicamente por quatro partes, seguindo o sentido 5' – 3': região promotora, região 5' UTR (Região não traduzida), região codante e a região 3' UTR. A figura abaixo mostra as divisões de um gene.

FIGURA 4 - Esquema demonstrativo da estrutura de um gene com suas respectivas regiões. Na parte superior da figura é esquematizado a fita dupla de DNA contendo um gene típico no sentido 5' e 3'; Na parte inferior mostra em maior escala as partes que um gene é composto, no sentido 5' para o 3' linha indicado pela seta estão: Região promotora; Região UTR 5', Região codante (com exon e intron) e região UTR 3'.



A região promotora é fundamental para a regulação da transcrição, é a parte do gene não transcrita, antecedente ao mRNA (RNA mensageiro). Nessa parte estão presentes sítios específicos de ligação de várias proteínas que tem papel fundamental na regulação da expressão do gene. Essas proteínas são conhecidas como fatores de transcrição que, através da ação direta ou intermediadas por outras proteínas, se ligam a sítios específicos da região promotora, conhecidas como motivos, e, assim, controlam a expressão gênica. Os motivos e fatores de transcrição são conhecidos como *cis*-elementos de regulação da transcrição (SEGAL et al., 2002).

A região promotora geralmente é a maior parte, que pode variar de um gene para outro, podendo ter 10 K pb (10 mil pares de bases) ou mais, e outros menos que 1 K pb dependendo do genoma da espécie ou da distância em que um

gene se encontra do outro no genoma. Apesar de alguns genes possuírem a região promotora longa, em alguns as regiões com predominância de motivos fundamentais para sua transcrição estão próximo da região TATA-BOX (Região *cis*-reguladora da RNA polimerase II), essa região geralmente está próximo de -100 pb (GLEZER et al., 2000). Como já observado em *Arabidopsis thaliana* ao testar diferentes tamanhos da região promotora na expressão do gene *FT* (LIU et al., 2014).

2.5 Fatores de transcrição

Em organismos eucariotos a transcrição da informação que dará origem a um mRNA para uma determinada função na célula ocorre através da RNA polimerase II regulada por fatores de transcrição (GLEZER et al., 2000). Fatores de transcrição são proteínas cuja atividade é influenciada por sinais do desenvolvimento e ambientais que se ligam em sequências específicas da região promotora do gene, chamadas de motivos, e controlam a transcrição de outros genes através da formação do complexo PIC e atração da PolII (Ref.) Assim, fatores de transcrição podem induzir ou inibir a transcrição do DNA em mRNA e formam uma rede complexa de informação para resposta celular coordenada (SEGAL et al., 2002). Em vários casos existem diversos fatores de transcrição inativos preexistentes que podem ser ativados quando sob o estímulo específico (BEG & BALDWIN, 1993).

Fatores de transcrição podem exercer sua função ao se ligarem diretamente ao DNA ou por interação com outras proteínas formando dímeros capazes de se ligar ao DNA (MANTOVANI, 1999). Como exemplo de fator de transcrição, a proteína CO codificada pelo gene *CONSTANS* reconhece proteínas nuclear que se ligam aos domínios B-box e CCT (*CONSTANS*, *CO-Like*, *TIME*

OF CAB1) e atuam na regulação da transcrição do gene *FT* em algumas espécies (ROBSON et al., 2001).

A interação dos fatores TGA4/OBF4 com o *CO* desempenha um papel fundamental na regulação de genes de florescimentos dependentes de fotoperíodo (MOURADOV et al., 2002; BÄURLE & DEAN, 2006) No gene *FT* eles se ligam na região promotora e são alvos primários de *CO* (SCHMID et al., 2003). Além de outros fatores, o domínio CCT de ligação do *CO* também está apto a interagir com fatores do complexo nuclear Y (NF-Y) (BEN-NAIM et al., 2006; WENKEL et al., 2006). Os elementos NF-Y se ligam ao DNA como um complexo heterotrimerico através de motivos do grupo CCAAT-BOX (MANTOVANI, 1999; MCNABB & PINTO, 2005).

Além do *CO* e dos fatores relacionados a ele, existem outros fatores de transcrição induzidos através de estímulos hormonais e de estresse que controlam a expressão de genes do florescimento como o *FT*. Alguns desses fatores são: o EIN3 e EIL (Ethylene-Insensitive 3 e Ethylene-Insensitive like), que estão relacionados a rota do etileno (CHANG et al., 2013); Os MADS-box que estão envolvidos no controle do florescimento, de respostas às rotas de vernalização, fotoperíodo e rotas autônomas de indução floral (BORNER et al., 2000; MATHELIER et al., 2013); Fatores de transcrição do grupo B3, que podem estar associados ao florescimento (FRANCO-ZORRILLA et al., 2014); Os elementos de transcrição VOZ, ativador da transcrição em rotas sinalizadas por fitocromos B, supressor do gene *FLOWERING LOCUS C (FLC)* e ativador da expressão do gene *FT*; E os SRS que estão relacionados a fatores que controlam o tempo de florescimento (CHANG et al., 2008).

Com o avanço da bioinformática e comparação de motivos conservados entre espécies, foi possível organizar bancos de dados para identificação e predição *in silico* de regiões *cis*-reguladoras. Esses bancos de dados possuem inúmeras sequências onde, através da comparação com a

sequência de interesse, é possível encontrar indicações de motivos presentes na sequência e a especificação de cada um deles. Alguns bancos de dados são mais completos que outros, mas todos apresentam no mínimo as sequências mais conhecidas. Como exemplos mais conhecidos temos o PlantCARE (Plant cis-acting regulatory element), PlantPAN (Plant promoter analysis navigator) e PlantTFDB (Plant transcription factor database).

2.6 Polimorfismo Gênico

Polimorfismo são alterações de bases na sequência do DNA decorrente de mutações induzidas ou não pelo ambiente. Essas mutações podem ser caracterizadas de várias formas, alteração de uma base ou de várias bases ao mesmo tempo. A alteração de uma única base é denominada de SNP (single nucleotide polymorphism), microssatélites são alterações de pequenas sequências do DNA, de quatro a oito nucleotídeos, já os minissatélites caracterizam-se por alterações no DNA de 10 a 100 nucleotídeos. Essas alterações podem ser tanto de deleção, adição de sequências ou troca de bases, e são principalmente causadas pela movimentação de elementos de transposição, conhecidos como transposons, que acabam por carregar ou deixar pedaços da sequência do DNA quando migram ou se copiam de uma parte para outra no genoma dentro ou entre cromossomos e acabam por alterar a constituição da sequência genômica (SIM et al., 2012; BROCCANELLO et al., 2015).

As mutações presentes no genoma das espécies podem não ser prejudiciais e são fundamentais para o processo de especiação dos organismos, pois através das modificações (polimorfismos) permite novas combinações, funcionalizações e através da seleção constitui um mecanismo fundamental para a evolução da vida (BROCCANELLO et al., 2015). Sendo marcas distinguíveis e herdáveis entre gerações, alguns tipos de polimorfismos são usados como

marcadores genético para a identificação e a separação de indivíduos. Isso passou a ser uma ferramenta fundamental e de grande importância para os estudos de melhoramento genético.

Polimorfismo presente entre indivíduos da mesma espécie ou de espécies diferentes podem gerar alterações na expressão gênica quando localizados em motivos que controlam a transcrição dos genes (LIAO & LEE, 2010). Alterações nas sequências motivos que controlam esses elementos *cis*-reguladores, podem impedir a ligação das proteínas e mudar os padrões de transcrição do gene (RAMALHO ET AL., 1990; MATIOLI & FERNANDES, 2012).

Polimorfismos no DNA podem ser observados de várias maneiras, através do uso de enzimas de restrição, amplificação de fragmentos através da reação em cadeia da polimerase (PCR) e análise dos fragmentos em eletroforese, e um dos mais usados por possibilitar visualizar qual a alteração na sequência, é através da fragmentação do DNA e sequenciamento da molécula (LIAO & LEE, 2010). Dentre os métodos de sequenciamento, o mais usado é o método de inibição de cadeia por terminação ou mais conhecido como método de Sanger devido ao baixo custo e boa eficiência.

O método de Sanger para sequenciamento do DNA é baseado no processo de replicação de um fragmento específico de DNA por PCR com *primers* que flanqueiam a sequência alvo e uso de desoxiribonucleotídeos (ddNTPs) marcados fluorescentemente, os quais se diferenciam dos ribonucleotídeos (dNTPs), pela falta do grupo OH no terceiro carbono. Sendo que, quando esses nucleotídeos são adicionados, ocorre a finalização da síntese de DNA. Cada ddNTP é marcado com um fluoróforo diferente e a sequência dos fragmentos de tamanhos diferentes é determinado pela separação deste fragmentos por eletroforese capilar em gel de alta resolução. À medida que os fragmentos são separados, eles passam por um laser e um detector. O laser

promove a ativação dos corantes fluorescentes, levando a uma emissão de fluorescência, a qual apresenta um comprimento de ondas característico para cada corante, possibilitando o sequenciamento do fragmento em questão (SANGER et al., 1977).

2.7 Características das espécies.

2.7.1 *Coffea arabica*

Distribuição: Nordeste da África Tropical (Sudoeste da Etiópia (Oeste do vale Great Rift) e Sudeste do Sudão (Boma Plateau)); Leste da África Tropical (Kenya (Mr. Marsibit), naturalizado na África Tropical e em outras regiões tropicais. Ecologia: Região úmida de florestas sempre verdes, altitude média de 950 a 1950 m (DAVIS et al., 2006). A espécie *Coffea arabica* é a única tetraploide ($2n = 4x = 44$) do gênero *Coffea*. Surgiu do cruzamento natural, mas de uma hibridação incomum de gametas entre as espécies diplóides *Coffea canephora* e *Coffea eugenioides*, e apresenta também a característica de autofecundação diferente das demais (LASHERMES et al., 1999).

Devido à grande importância econômica da espécie *C. arabica* hoje há diversos genótipos lançados no mercado, oriundo de vários cruzamentos. Um dos genótipos comerciais é o Acauã, oriundo do cruzamento entre a cultivar Mundo Novo IAC 388-17 e o Sarchimor IAC 1668, durante um processo de melhoramento visando a resistência a ferrugem. O resultado do cruzamento deu origem a esse genótipo que apresentou boa adaptação a condições adversas (MATTIELO, 2000). A cultivar Acauã produz frutos e sementes com formato alongado, sua bebida é classificada como de boa qualidade e possui maturação tardia (CARVALHO, 2008). Entre os genótipos da espécie *C. arabica*, o Semperflorens se destaca de todos os outros por não apresentar um período

exato de florescimento e não sofrer influência de fotoperíodo curtos e longos (WENT & ASHBY, 1959).

2.7.2 *Coffea canephora*

Distribuição: Oeste da África tropical (Gana, Guinéa, Guinéa Bissau, Costa do Marfim, Libéria e Nigéria), Centro-Oeste da África Tropical (Cabinda, Camarões, Congo, República da África Central, República Democrática do Congo, Gabão), Nordeste da África Tropical (Tanzânia e Uganda), Sul da África tropical (Angola). Ecologia: Umidade, florestas sempre verdes, às vezes em florestas sazonais úmidas secas, ocasionalmente em florestas de galeria (DAVIS et al., 2006).

A espécie *Coffea canephora* é diplóide ($2n = 22$) e apresenta autoincompatibilidade na mesma planta. Devido essa característica de incompatibilidade, existe uma grande variabilidade dentro de uma mesma lavoura, dificultando a existência de uma característica própria entre as plantas cultivadas, pois as sementes produzidas não reproduzem necessariamente as características da planta matriz (BRAGANÇA et al., 2001). Apesar da alta variabilidade encontrada nos genótipos de *C. canephora*, estudos com a cultivar Conilon mostraram que através da propagação assexuada as plantas apresentaram características boas para tamanho de fruto, uniformidade de maturação e alta capacidade produtiva (SINOCCK, 1965; FERRÃO et al., 2008).

2.7.3 *Coffea eugenioides*

Distribuição: Centro-Oeste da África Ocidental (Burundi, Ruanda, República Democrática do Congo), Nordeste da África Tropical (Sudão), Leste da África Tropical (Quênia, Tanzânia, Uganda). Ecologia: Local úmido, florestas verde, florestas de galeria, ou, florestas sazonalmente seca, às vezes em savanas e matas (DAVIS et al., 2006). Apesar dessa espécie ter contribuído para a

origem da espécie *Coffea arabia*, ela não possui boas características de produção, mas é muito usada em cruzamentos para obter aspectos de qualidade por possuir baixo teor de cafeína e frutos pequenos (MAZZAFERA & CARVALHO, 1991; YUYAMA et al., 2016).

2.7.4 *Coffea racemosa*

Distribuição: Sul da África Tropical (Moçambique, Zimbábue); África Austral (KwaZulu-Natal); Oceano Índico Ocidental (Canal de Moçambique). Ecologia: Sazonalmente seca floresta verde mista decídua, ou na floresta verde sazonalmente seca, principalmente em floresta de litoral incluindo florestas estabilizadas em dunas de areia, às vezes em matagal, muitas vezes encontrada em associação com *Microphylla brachystegia*; Altitude de 0-600m (DAVIS et al., 2006).

A espécie *C. racemosa* apresenta florescimento intenso e maturação muito precoce, leva apenas 90 a 100 dias da antese à maturação completa do fruto, enquanto, em *C. arabica*, esse período é de 220 dias, em média (MEDINA FILHO et al., 1977). Esta é uma espécie de resistência à seca e ao bicho mineiro. Estudos de cruzamentos com *C. arabica* mostraram que a precocidade de maturação e abundância de florescimento dessa espécie pode ser passada de geração em geração. (GUERREIRO FILHO et al., 1990).

3. CONSIDERAÇÕES GERAIS

Não há ainda um conhecimento específico sobre a ação do gene *FT* em cafeeiro. Considerando sua conservação de sequência e importância para a indução do florescimento em angiospermas, um entendimento mais profundo da expressão desse gene em espécies de cafeeiro com características reprodutivas distintas representa uma chave importante na elucidação do florescimento em plantas perenes. Além disso, devido a sua importância agrônômica, o entendimento do florescimento em *Coffea* sp. ajudaria na compreensão do principal problema da cultura que consiste no florescimento sequencial e desuniforme.

4. REFERÊNCIAS

ABE, M. et al. FD, a bZIP protein mediating signals from the floral pathway integrator *FT* at the shoot apex. **Science**, v. 309, n. 5737, p. 1052-1056, 2005. ISSN 0036-8075.

ABU-ASAB, M. et al. Earlier plant flowering in spring as a response to global warming in the Washington, DC, area. **Biodiversity & Conservation**, v. 10, n. 4, p. 597-612, 2001. ISSN 0960-3115.

ADRIAN, J. et al. cis-Regulatory Elements and Chromatin State Coordinately Control Temporal and Spatial Expression of FLOWERING LOCUS T in Arabidopsis. **Plant Cell**, v. 22, n. 5, p. 1425-1440, May 2010. ISSN 1040-4651.

ALVES, J. D.; LIVRAMENTO, D. Morfologia e fisiologia do cafeeiro. **Lavras: UFLA**, v. 1, 2003.

ALVIM, P. D. T. Moisture stress as a requirement for flowering of coffee. **Science**, v. 132, n. 3423, p. 354-354, 1960. ISSN 0036-8075.

AMASINO, R. Seasonal and developmental timing of flowering. **The Plant Journal**, v. 61, n. 6, p. 1001-1013, 2010. ISSN 1365-313X.

ANDRÉS, F.; COUPLAND, G. The genetic basis of flowering responses to seasonal cues. **Nature Reviews Genetics**, v. 13, n. 9, p. 627-639, 2012. ISSN 1471-0056.

BARRETO, H. G. et al. In silico and quantitative analyses of the putative FLC-like homologue in coffee (*Coffea arabica* L.). **Plant Molecular Biology Reporter**, v. 30, n. 1, p. 29-35, 2012. ISSN 0735-9640.

BÄURLE, I.; DEAN, C. The timing of developmental transitions in plants. **Cell**, v. 125, n. 4, p. 655-664, 2006. ISSN 0092-8674.

BEG, A. A.; BALDWIN, A. S. The I kappa B proteins: multifunctional regulators of Rel/NF-kappa B transcription factors. **Genes & development**, v. 7, n. 11, p. 2064-2070, 1993. ISSN 0890-9369.

BRAGANÇA, S. M. et al. Variedades clonais de café conilon para o Estado do Espírito Santo. **Pesquisa Agropecuária Brasileira**, v. 36, n. 5, p. 765-770, 2001.

BRIDSON, D.; VERDCOURT, B. Flora of tropical East Africa: Rubiaceae (part 2). **Rotterdam: AA Balkema**, 1988.

BROCCANELLO, C. et al. A new polymorphism on chromosome 6 associated with bolting tendency in sugar beet. **BMC genetics**, v. 16, n. 1, p. 1, 2015. ISSN 1471-2156.

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

CAO, K. et al. Four Tomato FLOWERING LOCUS T-Like Proteins Act Antagonistically to Regulate Floral Initiation. **Frontiers in plant science**, v. 6, 2015.

CARVALHO, A. et al. GENETIC-ASPECTS OF THE COFFEE TREE. **Revista Brasileira De Genetica**, v. 14, n. 1, p. 135-183, Mar 1991. ISSN 0100-8455.

CARVALHO, C. D. Cultivares de café: origem, características e recomendações. **Brasília: Embrapa Café**, 2008.

CHANG, K. N. et al. Temporal transcriptional response to ethylene gas drives growth hormone cross-regulation in Arabidopsis. **Elife**, v. 2, p. e00675, 2013. ISSN 2050-084X.

CHANG, W.-C. et al. PlantPAN: Plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups. **BMC genomics**, v. 9, n. 1, p. 1, 2008. ISSN 1471-2164.

CONAB. **Monitoramento agrícola – Café – Safra 2016**. CONAB. Brasília. v2: 1-68 p. 2016.

DA SILVA, F. M. et al. Uso de ethrel na colheita mecanizada e seletiva de café arábica (*Coffea arabica* L.). **Coffee Science**, v. 4, n. 2, p. 178-182, 2009. ISSN 1984-3909.

DAVIS, A. P. et al. An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). **Botanical Journal of the Linnean Society**, v. 152, n. 4, p. 465-512, 2006. ISSN 1095-8339.

DE OLIVEIRA, R. R. et al. Flower development in *Coffea arabica* L.: new insights into MADS-box genes. **Plant reproduction**, v. 27, n. 2, p. 79-94, 2014. ISSN 2194-7953.

ECCARDI, F.; SANDALJ, V. **O café: Ambientes e diversidade**. Casa da Palavra, 2003. ISBN 8587220713.

FERRÃO, R. G. et al. Parâmetros genéticos em café Conilon. **Pesquisa Agropecuária Brasileira**, v. 43, n. 1, p. 61-69, 2008.

FORNARA, F.; DE MONTAIGU, A.; COUPLAND, G. SnapShot: control of flowering in Arabidopsis. **Cell**, v. 141, n. 3, p. 550-550. e2, 2010. ISSN 0092-8674.

FRANCO-ZORRILLA, J. M. et al. DNA-binding specificities of plant transcription factors and their potential to define target genes. **Proceedings of the National Academy of Sciences**, v. 111, n. 6, p. 2367-2372, 2014. ISSN 0027-8424.

GLEZER, I. et al. O fator de transcrição NF-kB nos mecanismos moleculares de ação de psicofármacos. **Rev Bras Psiquiatr**, v. 22, n. 1, p. 26-30, 2000.

GOUVEIA, N. Florescimento e maturação em Coffea arabica cv. Catuai vermelho. 9. Congresso Brasileiro de Pesquisas Cafeeiras 27-30 Oct 1981 São Lourenço, MG (Brasil), 1981, Instituto Brasileiro do Café, Rio de Janeiro (Brasil).

GUERRA, A.; ROCHA, O.; RODRIGUES, G. Manejo do cafeeiro irrigado no Cerrado com estresse hídrico controlado. **Irrigação & Tecnologia Moderna**, v. 65, n. 66, p. 42-45, 2005.

GUERREIRO FILHO, O. et al. Melhoramento do cafeeiro: XLIII. Seleção do cafeeiros resistentes ao bichomineiro. **Bragantia**, v. 49, p. 291-304, 1990.

LASHERMES, P. et al. Molecular characterisation and origin of the Coffea arabica L. genome. **Molecular and General Genetics MGG**, v. 261, n. 2, p. 259-266, 1999. ISSN 0026-8925.

LEVY, Y. Y.; DEAN, C. The transition to flowering. **The Plant Cell**, v. 10, n. 12, p. 1973-1989, 1998. ISSN 1532-298X.

LIAO, P.-Y.; LEE, K. H. From SNPs to functional polymorphism: The insight into biotechnology applications. **Biochemical Engineering Journal**, v. 49, n. 2, p. 149-158, 2010. ISSN 1369-703X.

LIFSCHITZ, E.; AYRE, B. G.; ESHED, Y. Florigen and anti-florigen—a systemic mechanism for coordinating growth and termination in flowering plants. **Frontiers in plant science**, v. 5, p. 465, 2014. ISSN 1664-462X.

LIU, L. et al. Induced and natural variation of promoter length modulates the photoperiodic response of FLOWERING LOCUS T. **Nature communications**, v. 5, 2014.

LIVRAMENTO, D. E. Morfologia e Fisiologia do Cafeeiro. In: REIS, P. R.; CUNHA, R. L. (Ed). **Café Arábica do plantio á colheita**. Lavras: U.R. EPAMIG SM, 896 p., 2010.

MANTOVANI, R. The molecular biology of the CCAAT-binding factor NF-Y. **Gene**, v. 239, n. 1, p. 15-27, 1999. ISSN 0378-1119.

MATIOLI, S. R.; FERNANDES, F. M. D. C. Biologia molecular e evolução. In: (Ed.). **Biologia molecular e evolução**: Holos/Sociedade Brasileira de Genética, 2012.

MATTIELO, J. B. A. et al. **Acauã-variedade de café adaptada para regiões mais secas e quentes**. CONGRESSO BRASILEIRO DE PESQUISAS CAFEIRAS. Marília, SP: PROCAFÉ. 26: 1 p. 2000.

MCGARRY, R. C.; KRAGLER, F. Phloem-mobile signals affecting flowers: applications for crop breeding. **Trends in plant science**, v. 18, n. 4, p. 198-206, 2013. ISSN 1360-1385.

MCNABB, D. S.; PINTO, I. Assembly of the Hap2p/Hap3p/Hap4p/Hap5p-DNA complex in *Saccharomyces cerevisiae*. **Eukaryotic Cell**, v. 4, n. 11, p. 1829-1839, 2005. ISSN 1535-9778.

MEDINA FILHO, H.; CARVALHO, A.; MEDINA, D. Germoplasma de *Coffea racemosa* e seu potencial no melhoramento do cafeeiro. **Bragantia**, v. 36, p. 43-46, 1977.

MELO, B.; SOUSA, L. B. BIOLOGIA DA REPRODUÇÃO DE *Coffea arabica* L. e *Coffea canephora* Pierre. **Revista Verde de Agroecologia e Desenvolvimento Sustentável**, v. 6, n. 2, 2011. ISSN 1981-8203.

MORAIS, H. et al. Escala fenológica detalhada da fase reprodutiva de *Coffea arabica*. **Bragantia**, v. 67, n. 1, p. 257-260, 2008. ISSN 0006-8705.

MOURADOV, A.; CREMER, F.; COUPLAND, G. Control of flowering time interacting pathways as a basis for diversity. **The Plant Cell**, v. 14, n. suppl 1, p. S111-S130, 2002. ISSN 1532-298X.

NOTAGUCHI, M. et al. Long-distance, graft-transmissible action of Arabidopsis FLOWERING LOCUS T protein to promote flowering. **Plant and Cell Physiology**, v. 49, n. 11, p. 1645-1658, 2008. ISSN 0032-0781.

PIN, P. A. et al. An antagonistic pair of *FT* homologs mediates the control of flowering time in sugar beet. **Science**, v. 330, n. 6009, p. 1397-1400, 2010. ISSN 0036-8075.

PUTTERILL, J.; LAURIE, R.; MACKNIGHT, R. It's time to flower: the genetic control of flowering time. **Bioessays**, v. 26, n. 4, p. 363-373, 2004. ISSN 1521-1878.

PUTTERILL, J. et al. *FT* genes and regulation of flowering in the legume *Medicago truncatula*. **Functional Plant Biology**, v. 40, n. 12, p. 1199-1207, 2013. ISSN 1445-4416.

RAMALHO, M.; DOS SANTOS, J. B.; PINTO, C. B. **Genética na agropecuária**. FAEPE, 1990. ISBN 8525006777.

ROBSON, F. et al. Functional importance of conserved domains in the flowering-time gene *CONSTANS* demonstrated by analysis of mutant alleles and transgenic plants. **The Plant Journal**, v. 28, n. 6, p. 619-631, 2001. ISSN 1365-313X.

SAMACH, A. et al. Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. **Science**, v. 288, n. 5471, p. 1613-1616, 2000. ISSN 0036-8075.

SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences**, v. 74, n. 12, p. 5463-5467, 1977. ISSN 0027-8424.

SCHMID, M. et al. Dissection of floral induction pathways using global expression analysis. **Development**, v. 130, n. 24, p. 6001-6012, 2003. ISSN 0950-1991.

SEGAL, E. et al. From promoter sequence to expression: a probabilistic framework. Proceedings of the sixth annual international conference on Computational biology, 2002, acm. p.263-272.

SHALIT, A. et al. The flowering hormone florigen functions as a general systemic regulator of growth and termination. **Proceedings of the National Academy of Sciences**, v. 106, n. 20, p. 8392-8397, 2009. ISSN 0027-8424.

SINOCCK, J. La rénovation de la caféiculture malgache a partir de clones sélectionnés. **Café, Cacao, Thé (France)** v. 12 (3) p. 223-235, 1965.

THOMAZIELLO, R. A. et al. **Café Arábica: cultura e técnicas de produção**. Campinas: IAC (Boletim Técnico, 187), 82p., 2000.

TURCK, F.; FORNARA, F.; COUPLAND, G. Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. **Annu. Rev. Plant Biol.**, v. 59, p. 573-594, 2008. ISSN 1543-5008.

VARDY, M. et al. Conservation of Arabidopsis flowering genes in model legumes. **Plant Physiol**, v. 137, n. 14201434, p. 104.057018, 2005.

WATSON, J. D. Molecular biology of the gene. **Molecular biology of the gene.**, n. 2nd edn, 1970.

WELLER, J. L.; ORTEGA, R. Genetic control of flowering time in legumes. **Frontiers in plant science**, v. 6, p. 207, 2015. ISSN 1664-462X.

WENT, F.; ASHBY, W. **The Experimental Control of Plant Growth.** Ronald, 1959.

WICKLAND, D. P.; HANZAWA, Y. The FLOWERING LOCUS T/TERMINAL FLOWER 1 gene family: functional evolution and molecular mechanisms. **Molecular plant**, v. 8, n. 7, p. 983-997, 2015. ISSN 1674-2052.

WIGGE, P. A. et al. Integration of spatial and temporal information during floral induction in Arabidopsis. **Science**, v. 309, n. 5737, p. 1056-1059, 2005. ISSN 0036-8075.

WONG, A. et al. Isolation and functional analysis of CONSTANS-LIKE genes suggests that a central role for CONSTANS in flowering time control is not evolutionarily conserved in *Medicago truncatula*. **Frontiers in plant science**, v. 5, p. 486, 2014. ISSN 1664-462X.

CAPITULO 2

**Identificação de polimorfismo na região promotora do gene *FT* em
cafeeiros das espécies *Coffea arabica*, *Coffea canephora*, *Coffea eugenioides*
e *Coffea racemosa***

RESUMO

O Brasil é o maior produtor e exportador de café do mundo, o entendimento e aprimoramento das técnicas de manejo da cultura a fim de melhorar a qualidade do produto e reduzir custos na produção, são de grande relevância para os agricultores e para o país, devida sua grande importância social e econômica. Um dos maiores problemas encontrados pelos produtores na cafeicultura é a floração sequencial e desuniforme do cafeeiro, agregando custos elevados para a produção da bebida com parâmetros de boa qualidade. O conhecimento de genes envolvidos na floração pode ser a chave para o entendimento dessa característica de florescimento do cafeeiro. *FLOWERING LOCUS T (FT)* é um dos genes fundamentais durante a indução do florescimento para algumas espécies. Alterações na região promotora do gene decorrente de mutações podem ser determinantes para o comportamento da sua expressão, uma vez que na região promotora existem sequências motivos de ligação de fatores de transcrição que controlado a regulação da expressão. Com o objetivo de entender melhor a ação do gene *FT* sobre o florescimento de cafeeiro comparamos um fragmento da região promotora entre genótipos representantes das espécies *Coffea arabica*, *Coffea canephora*, *Coffea eugenioides* e *Coffea racemosa*, e encontramos motivos dentro das sequências isoladas descritas como locais importantes para reconhecimento e ligação de fatores de transcrição. Alguns desses motivos direcionam fatores de transcrição ligados ao florescimento que de certa forma atuam sobre o controle da expressão do gene *FT* em cafeeiro. A região promotora dos genes *FT* identificados nas quatro espécies apresentam polimorfismos e o gene *CO* e *FT* foram diferencialmente expresso entre as espécies. Entretanto, não foi encontrada relação entre a expressão do gene *FT* com as características de florescimento específica de cada genótipo estudado.

Palavras-chave: Cafeeiro. Florescimento. Gene *FT*.

ABSTRACT

Brazil is the largest producer and world exporter of coffee and the understanding and improvement of crop management techniques in order to improve product quality and reduce production costs are of great importance for farmers and for the country itself due to its great social and economic importance. One of the biggest problems faced by coffee growers is the sequential and asynchronous flowering of coffee trees, which increases production costs for the production of a good quality product. A better knowledge of the genes involved in flowering can be the key to the understanding the feature of the sequential flowering found in coffee trees. *FLOWERING LOCUS T (FT)* is one of the key genes during the flowering induction in some species. Changes in the promoter gene sequence due to mutations may be crucial for its expression pattern, since in the promoter region are located binding motifs sequences of transcription factors that regulates gene expression. In order to better understand the action of *FT* on coffee flowering, this study aimed to compare the putative polymorphisms of the *FT* promoter region, as well as its expression pattern, in four different coffee species, *Coffea arabica*, *Coffea canephora*, *Coffea eugenoides* and *Coffea racemosa*, with different flowering features. It was possible to find several motifs from cis-regulatory elements in the promoter regions of the identified *FT* sequences. Some of these motifs direct transcription factors related to flowering regulation that act on the control of *FT* gene expression. Although it was not found a relationship between the *FT* expression pattern and the flowering features of the genotypes analyzed in this study, the *FT* promoter region from the four coffee species analyzed in this study show different polymorphisms, which may be linked to the different expression pattern found among the four coffee species analyzed.

Key-Words: Coffee tree. Flowering. *FT* gene.

5. INTRODUÇÃO

O florescimento inicia-se com a indução do estágio vegetativo para o reprodutivo nas plantas (ANDRÉS & COUPLAND, 2012). Esse processo é controlado por vários genes cuja expressão responde a diversos fatores bióticos e abióticos para garantir à época mais favorável a reprodução e, portanto, essa fase tem efeito direto sobre a frutificação. Os estímulos sofridos pela planta durante a indução do florescimento determinam a expressão gênica basicamente pelas rotas de idade, temperatura, vernalização, fotoperíodo e autônoma (BLÜMEL et al., 2015).

Em cafeeiro esse processo fisiológico é bem complexo apresentando características peculiares, tais como, floração sequencial e desuniforme (Ref?). O efeito desta característica influencia na uniformidade do desenvolvimento e no ponto de maturação dos frutos durante a colheita (MELO & SOUSA, 2011). Em consequência disso, é agregado um custo mais elevado que o normal para se produzir café de boa qualidade, por ser necessário mais de uma colheita e fazer a separação dos frutos conforme o estágio de maturação. Isso o torna inacessível à maioria dos consumidores (GOUVEIA, 1981).

Apesar de haver vários estudos com o objetivo de entender o que leva a essas características, pouco se sabe sobre os reais motivos que causam o florescimento desuniforme do cafeeiro. O entendimento melhor da ação de alguns genes durante o processo de indução do florescimento no cafeeiro pode ajudar a encontrar algumas respostas para esse fenômeno. Exemplo de gene como o *FLOWERING LOCUS T (FT)* que está envolvido diretamente na indução e ativação de outros genes de formação dos órgãos florais através da interação no meristema apical da planta. O gene *FT* expresso na folha e sua proteína transportada até o meristema apical onde irá exercer sua função (CORBESIER et al., 2007).

Um dos principais fatores que controlam a transcrição do gene *FT* é proteína traduzida do gene *CONSTANS* (*CO*) que é regulado pelo fotoperíodo (ADRIAN et al., 2010; ZHANG et al., 2015). A ação de fatores de transcrição como o *CO* sobre a expressão do gene pode ser afetada quando mutações no DNA localizadas na região promotora do gene alteram as regiões de motivos de elementos *cis*-reguladores da expressão (DOI: 10.1038/ncomms5558). Essas alterações podem ser cruciais na regulação da expressão gênica, pois podem alterar o tempo e a quantidade dos transcritos ou até mesmo inibir a transcrição (WENKEL et al., 2006).

Levando em consideração a importância de entender melhor a ação dos genes na regulação da indução do florescimento do cafeeiro a fim de ampliar os conhecimentos sobre os motivos que levam ao florescimento sequencial e desuniforme dessa planta, o presente estudo teve como objetivo identificar polimorfismos na região promotora do gene *FT* de quatro espécies do gênero *Coffea*, a fim de entender melhor a influência dessas alterações na expressão do gene entre as espécies e comparar com as características de florescimento dos genótipos usados.

6. MATERIAL E MÉTODOS

6.1 Material vegetal

Do material vegetal usado nas análises, os genótipos Semperflorens (*Coffea arabica*), Robusta (*Coffea canephora*) e Racemosa (*Coffea racemosa*) foram coletados no banco de germoplasma da fazenda experimental da Empresa de Pesquisa Agropecuária de Minas Gerais EPAMIG no município de Patrocínio, Minas Gerais, os genótipos Acauã (*Coffea arabica*) e Eugenioides (*Coffea eugenioides*) foram coletados no campus experimental da Universidade Federal de Lavras, Minas Gerais. Todos foram resfriados em nitrogênio líquido no ato da coleta e crio-preservados.

Os genótipos usados são referentes às quatro espécies estudadas, apenas para a espécie *C. Arabica* houve um exemplar a mais, o genótipo Semperflorens por apresentar florescimento distinto em comparação aos demais da mesma espécie. Foram coletadas as folhas do terço médio da planta, não muito velhas e nem muito novas. Para cada genótipo foram coletadas folhas de três plantas diferentes como repetições biológicas e identificadas como R1, R2 e R3.

6.2 Extrações de DNA e RNA

Inicialmente foi realizada a maceração das folhas com almofariz e pistilo para a quebra mecânica das estruturas e início do processo de lise celular. A extração de RNA foi de acordo com o protocolo de extração de RNA de plantas de uso do reagente Concert (Ivitrogen USA), com poucas alterações (DOYLE, 1990). Foi adicionado 600 µL de Concert ao material macerado e levado ao vórtex para homogeneização por 2 minutos. Os tubos foram colocados na posição horizontal por 10 minutos á temperatura ambiente e posteriormente centrifugados a 14.800 rpm por 10 minutos a 4 °C. Em seguida,

foi coletado 450 μL e 420 μL do sobrenadante e transferido-os para um novo tubo. Adicionou-se 100 μL de NaCl 5M e 300 μL de clorofórmio, em seguida levou-se os tubos ao vórtex por 1 minuto e a centrifugação a 14.800 rpm por 10 minutos a 4 °C. Novamente, coletou-se o sobrenadante nos volumes de 400 μL e 370 μL , transferindo-os para novos tubos. Mais uma vez, adicionou-se 100 μL de NaCl 5M e 300 μL de clorofórmio, repetindo-se as etapas da homogeneização em vórtex e centrifugação citadas anteriormente. Coletou-se 280 μL e 320 μL do sobrenadante, transferindo-os para novos tubos. Foi adicionado Isopropanol na mesma proporção Volume/Volume ao sobrenadante e as amostras foram homogeneizadas por cerca de 15 inversões, em seguida, os tubos foram colocados em freezer -20 °C durante 1 hora para precipitação do RNA. Posteriormente, as amostras foram centrifugadas a 14.800 rpm por 25 minutos a 4 °C. O sobrenadante foi descartado e 600 μL de etanol 75% foram adicionados ao pellet, e as amostras foram centrifugadas a 14.800 rpm por 5 minutos a 4 °C. O etanol foi descartado cuidadosamente e os tubos foram mantidos abertos até o etanol secar completamente. Após a secagem, o RNA foi ressuspensionado com 20 μL de água miliQ autoclavada.

A extração de DNA foi conforme o protocolo de utilização do tampão CTAB (MURRAY & THOMPSON, 1980) com poucas alterações. Foi adicionado 1 ml do tampão de extração CTAB pré aquecido a 65 °C [CTAB 2% (p/v); NaCl 2 M; Tris-HCl 100 mM, pH 8,0; EDTA 25 mM; β -mercaptoetanol 0,2% (v/v)] ao microtubo contendo o tecido pulverizado e, em seguida, a mistura foi homogeneizada em vórtex e incubados por 30 min. a 65 °C. Então, os microtubos foram centrifugados por 10 min. a 11000 rpm. Após a transferência da fase líquida para novos microtubos foi adicionado um volume igual de clorofórmio: álcool isoamílico 24:1 (p/v) para permitir a desnaturação e insolubilização protéica. A solução foi homogeneizada e centrifugada por 5 min. a 11000 rpm. A fase líquida foi novamente transferida para novos microtubos e

adicionado 4 μL de RNase (100 $\mu\text{g}/\text{mL}$). As amostras foram incubadas a 37 $^{\circ}\text{C}$ por 30 min. para a digestão completa dos RNAs presentes. Em seguida, o DNA foi precipitado com a adição de 300 μL de álcool isopropílico e centrifugação a 13000 rpm durante 15 minutos. Após descartar o sobrenadante, foi adicionado 400 μL de etanol absoluto para a remoção do excesso de sais e centrifugado os microtubos em rotação máxima por 3 minutos. As amostras foram secadas durante 3 min. e ressuspensas em 20 μL de água MiliQ autoclavada.

Após a extração do RNA e DNA, as amostras foram analisadas quanto à quantidade e qualidade utilizando-se um espectrofotômetro Nanovue® (NanoVue GE Healthcare). Utilizou-se 1,5 μL de cada amostra para realizar essa análise, observando-se a concentração de ácidos nucleicos e as relações de pureza A260\280 e A260\230, priorizando, preferencialmente, valores entre 1,8 a 2,0. Para avaliar a integridade das amostras, foram submetidas a eletroforese em gel de agarose 0,8% corado com GelRed Nucleic Acid Gel Stain e visualizado em fotodocumentador UV-transilluminator (UVITEC FireReader XS D-77Ls- 20. M). Em seguida as amostras foram estocadas em freezer -20 $^{\circ}\text{C}$.

6.3 Tratamento com DNAase e Síntese de cDNA

As amostras de RNA para a análise de expressão e amplificação das regiões codantes do gen *FT* foram tratadas com o uso do Turbo DNA free™ kit (Ambion) para eliminação do DNA residual. Em tubos de 0,5 mL, 5 μg de RNA foi utilizado, o volume final foi completado para 22 μL com água comercial livre de RNase e demais contaminantes. O mix (tampão + DNase) foi preparado para um volume de reação de 25 μL , contendo 0.1 volume (2,5 μL) de 10x Turbo DNase Buffer e 0.1 volume (5 μL) de Turbo DNase. 3 μL do mix foram adicionados em cada amostra e homogêneos delicadamente sendo, em seguida, incubado em banho-maria a 37 $^{\circ}\text{C}$ por 30 minutos. Na sequência, 0.1

volume (2,5 µL) de DNase Inactivation reagent foi adicionado em cada amostra e as mesmas foram homogeneizadas, mantendo-se por 5 minutos a temperatura ambiente, e finalmente centrifugadas a 10.000 g por 2 minutos. Com muito cuidado, 15 µL do sobrenadante foi coletado e transferido para novos tubos de 0,2 mL. Após, 1,5 µL do RNA tratado foi utilizado para quantificação no Nanovue®. Para avaliar a integridade e pureza do RNA, as amostras foram submetidas à análise por eletroforese em gel de agarose 0,8% corado com GelRed Nucleic Acid Gel Stain e visualizado através de fotodocumentador UV-transilluminator (UVITEC FireReader XS D-77Ls-20.M).

Após o tratamento do RNA com DNase, as amostras apresentando alto grau de integridade e pureza foram utilizadas para a síntese de cDNA utilizando o Kit High-Capacity cDNA Reverse Transcription (Applied Biosystems). Em tubos de 0,2 mL, 1 µg de RNA tratado com DNase foi utilizado, completando-se o volume para 10 µL com água comercial livre de RNase e demais contaminantes. 10 µL do mix, contendo 2 µL de 10X Reverse Transcriptase Buffer, 2 µL de 10x RT random *primers*, 0,8 µL de dNTP mix, 1 µL de MultiScribe™ Reverse Transcriptase e 4,2 µL de água livre de DNase e RNase), foi adicionado em cada reação, resultando em um volume final de reação de 20 µL. Subsequentemente, as amostras foram delicadamente homogeneizadas e transferidas para o termociclador (Multigene Gradient – Labnet), onde ocorreu a síntese do cDNA. A reação ocorreu com as seguintes condições de ciclagem: 10 minutos a 25 °C, 120 minutos a 37 °C e 5 minutos a 85 °C. Ao final da reação as amostras foram armazenadas em freezer -20 °C.

6.4 Análise *in-silico*

Para a identificação da região codante e promotora do gene *FT* em *Coffea*, foram feitas algumas análises de bioinformática. Como o gene *FT* do cafeeiro ainda não foi descrito na literatura, a sua identificação foi feita no banco

CAFEST por trabalhos de bioinformática do LFMP. Como não há informações sobre o genoma de tal espécie, *primers* desenhados a partir da sequência do *FT* de *C. arabica*, foram utilizados para amplificar o gene nas outras espécies. Quanto à região promotora, os *primers* foram desenhados através da região encontrada no genoma da espécie *C. canephora* já sequenciado e disponível para acesso (coffee-genome.org).

As análises das sequências foram feitas com o auxílio de ferramentas de bioinformática. Através da sequência codante do gene *FT* de arabica, com o uso do algoritmo BLAST do banco de dados National Center for Biotchnology Information (NCBI) efetuou-se uma busca por sequências de *FT* de outras famílias. As sequências encontradas foram alinhadas pelo programa ClustalW, para análise de similaridade entre duas ou mais sequências (<http://www.genome.jp/tools/clustalw/>). O resultado do alinhamento foi analisado pelo software GeneDoc (fabricante?). Após a busca pelo gene *FT* de outras espécies, realizou-se uma análise de domínios através da ferramenta Conserved Domeins Database (CDD) presente no banco de dados NCBI a fim de eliminar sequências que não apresentavam o mesmo domínio do gene *FT* (PEBP).

Após as análises para a confirmação da sequência encontrada no CAFEST foi montada uma árvore filogenética com as sequências da proteína do *FT* encontradas no NCBI das famílias Rubiaceae, Fabaceae e Solanaceae com o objetivo de mostrar a distância parental do gene *FT* entre as famílias. A partir de alinhamentos múltiplos, feitos pelo programa ClustalW (THOMPSON et al., 1994), utilizando parâmetros padrão (*default*), a análise de similaridade entre estas sequências de proteínas foi feita pela geração da árvore filogenética pelo programa MEGA versão 5.2 (TAMURA et al., 2007), utilizando o modelo de comparação *Neighbor-joining* (SAITOU & NEI, 1987), com o método de distância-P e supressão *pair-wise*. A validade das árvores quanto à distância

filogenética dos clusters, pôde ser medida pelo teste probabilístico de *bootstraps*, com 1000 repetições (SITNIKOVA et al., 1995).

6.5 Desenho de *primers*

A partir das sequências obtidas em análises *in silico*, os *primers* para amplificação das regiões codante e promotora do gene *FT* foram desenhados através da ferramenta OligoPerfect (<http://tools.thermofisher.com/content.cfm?pageid=9716>). A qualidade dos *primers* desenhados foi analisado por meio do software Oligo Analyzer3.1 (OWCZARZY et al., 2008) no site da empresa IDT (Integrated DNA Technologies).

Para amplificar a região codante do gene *FT* nas quatro espécies foram desenhados na sequência do gene *FT* da espécie *C. arabica* dois *primers* forward e dois reverse. Os *primers* forward foram desenhados no início da sequência (Fw4) e o outro no meio da sequência (Fw5). Os *primers* reverse foram desenhados no meio (Rv3) e no final da sequência (Rv1). Para a amplificação da região promotora foram desenhados quatro *primers*, três forward e um reverse. Os *primers* foram desenhados nas regiões -1072 pb (Fw1), -984 pb (Fw2) e -580 pb (Fw3) anterior ao códon ATG de início da sequência codante. O *primer* reverse (Rv2) para amplificar a região promotora foi desenhado no início da sequência codante do gene. Um esquema das regiões de desenho dos *primers* esta disponível na fig. 5 e mostra o local de cada *primer* e o tamanho dos amplicons formados. A tabela 2 apresenta a sequência dos *primers* utilizados nas análises.

FIGURA 5 – Esquema demonstrativo dos amplicons gerados pelos *primers* desenhados para a amplificação da região promotora e codante do gene *FT* em *C. arábica* e *C. canephora*. As setas indicam o sentido da sequência amplificada pelo *primer* e as barras verdes o tamanho do amplicon formado por cada *primer* descrito acima.

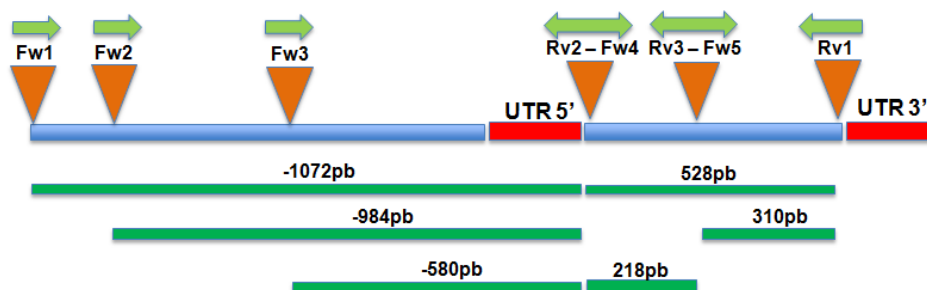


TABELA 2 – Sequências dos *primers* desenhados para a amplificação da região codante e promotora do gene *FT* dos cinco genótipos usados representante das quatro espécies estudadas. Na primeira coluna estão as identificações de cada *primer* conforme a figura 5, na segunda coluna estão as sequências de cada *primer*. As bases sublinhadas (CACC) foram acrescentadas para complementaridade do vetor Pentr (Fabricante ou id no NCBI).

Identificação	Sequência
Fw1	<u>CACCT</u> GGCTCACGGTTCAAT
Fw2	<u>CACCCTT</u> CCTCGTACTTCTTGC
Fw3	<u>CACCCACTA</u> ATTGATACTCTGC
Fw4	<u>CACCATG</u> TATAGGGGGGTAGT
Fw5	<u>CACCTCCA</u> AGTCCCAGTAATCC
Rv1	TCATGATGATCTTCTGCCTCCA
Rv2	ATCTACCCCCCTATACATCGTC
Rv3	GGATTACTGGGACTTGGAGC

O desenho dos *primers* para as análises de PCR em tempo real seguiu o mesmo protocolo utilizado no desenho dos *primers* para amplificar as regiões

codantes e promotora do gene, porém seguindo os parâmetros normalmente utilizados para amplificação de fragmentos pela técnica de RT-qPCR, que são o menor tamanho do amplicon (80-150 pb), a utilização de primers na região conservada do gene e um maior rigor na qualidade dos primers desenhados. Os primers foram sintetizados pela empresa ThermoFisher Scientific.

6.6 Isolamento da região promotora e codante do gene *FT*.

As sequências foram isoladas através da PCR (Reação em Cadeia da Polimerase) seguindo o protocolo descrito para a enzima utilizada Phusion Hi-fi (fabricante e número de catálogo). Os fragmentos gerados através da PCR foram visualizados por Eletroforese em Gel de Agarose. Após o isolamento das sequências foi realizada a purificação dos fragmentos para eliminação dos reagentes utilizados durante a PCR com o Kit especificar o nome seguindo as especificações do fabricante (BioLab, número de catálogo). As amostras foram armazenadas em ultra-freezer -20 °C do LFMP para serem sequenciadas.

6.7 Sequenciamento

As amostras foram quantificadas no aparelho Nanovue® e a concentração foi ajustada para 20 a 30 ng/ul para cada 100 pb a ser sequenciado conforme as exigências da empresa Myleus (www.myleus.com). Para a região promotora efetuou-se o sequenciamento em partes, devido ao tamanho da sequência ultrapassar os padrões de confiança da técnica (600 pb). Já para a região codante, o fragmento gerado foi menor do que o tamanho limite dos fragmentos normalmente gerados pela técnica, não sendo necessário mais de uma reação para ter o fragmento completo. Após o sequenciamento, os fragmentos foram unidos através da identificação da região de encontro na sobreposição das sequências senso e antisenso. Para cada 600 pb ou cada reação

foi usado um *primer*. Na região codante foi necessário apenas um *primer* para obtenção da sequência completa, porém foram desenhados mais dois *primers* no meio da sequência (Fig. 5) para ter sobreposição das sequências e aumentar a confiabilidade do resultado.

Na região promotora utilizou-se um *primer* (Rv2) em sentido antisenso no início da sequência codante do gene e o outro (Fw1) na região de -1072 pb no sentido senso (5' – 3'). Com o resultado do sequenciamento dos dois fragmentos foi possível sobrepor a região de encontro com eficiência e unir as duas sequências para obtenção da sequência promotora do gene *FT* dos cinco genótipos avaliados.

6.8 Análise de polimorfismo na região promotora do gene *FT*

As sequências da região promotora do gene *FT* obtidas através de sequenciamento foram alinhadas através da ferramenta ClustalW (<http://www.genome.jp/tools/clustalw/>) e analisadas pelo software GeneDoc (NICHOLAS et al., 1997) para a identificação de possíveis polimorfismos entre os genótipos avaliados. A presença de sequências *cis*-reguladoras do DNA na região promotora do gene dos genótipos efetuou-se através de buscas no banco de dados PlantCARE (Plant Cis-Acting Regulatory Elements) e PlantPAN (Plant Promoter Analysis Navigator).

6.9 Análise da expressão gênica por RT-qPCR

A análise da expressão dos genes *FT* e *CO* relacionados ao florescimento foram conduzidas utilizando-se o termociclador Rotor-Gene Q Real-Time PCR (Qiagen), utilizando o sistema de detecção SYBR® Green. O volume final da reação para cada amostra foi de 15 µL: 7,5 µL de SYBR Green

(QuantiFast SYBR Green PCR Kit – Qiagen), 3 μL de *primers* específicos para cada gene alvo (1,5 μL *primer* forward + 1,5 μL *primer* reverse), 1,5 μL de cDNA a 10 ng/ μL e 3 μL de água livre de RNase e demais contaminantes.

Foram utilizadas três repetições biológicas de cada genótipo (Semperflorens, Robusta, Racemosa, Acauã e Eugenioídes), sendo utilizadas triplicadas técnicas para cada uma das amostras. As condições de ciclagem foram: 5 min a 95 °C para ativação da enzima, seguidos de 40 ciclos de 5 s a 95 °C e 10 s a 60 °C, e completado com uma curva de melting para avaliar a especificidade da reação, aumentando-se a temperatura de 60 °C para 95 °C, com o aumento de 1 °C a cada 5 s. A eficiência de amplificação dos *primers* dos genes alvos e de referência foi determinada por meio de curva de diluição. A análise da expressão relativa foi calculada através do método $\Delta\Delta\text{CT}$, de acordo com (PFAFFL, 2001), utilizando os genes de referência UBQ - Ubiquitina e RPL39 - Proteína Ribossomal L39, (BARRETO, resultados não publicados). A diferença relativa da expressão de cada gene foi feita com base em uma amostra calibradora, que foi escolhida de acordo com o menor nível de expressão para cada um dos genes nos genótipos analisados.

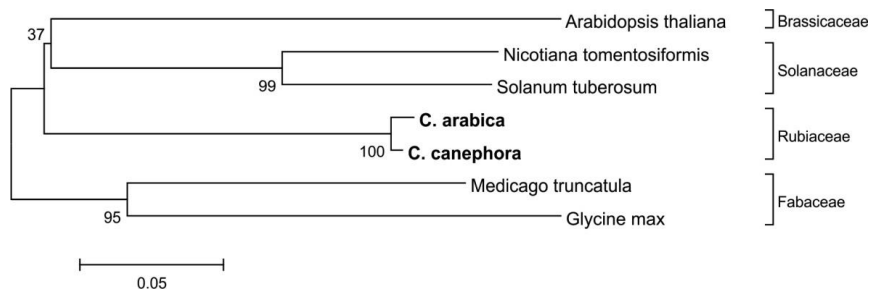
7. RESULTADOS E DISCUSSÃO

7.1 Análise *in silico*

7.1.1 Identificação da sequência codante do gene *FT*

Buscas em bancos de dados disponíveis para as espécies *C. arabica* e *C. canephora* revelaram a região codante do gene *FT* de café para essas espécies. A sequência de *C. arabica* foi encontrada no banco de dados de transcritos da espécie (CAFEST) e de *C. canephora* no banco de dados Coffee Genome Hub. O resultado do alinhamento entre as sequências da proteína do *FT* de *C. arabica* e *C. canephora* mostra um alto grau de similaridade e conservação entre elas em comparação com outras sequências de famílias diferentes, possibilitando o uso do *primer* desenhado na sequência de *C. arabica* para amplificação do gene nas outras espécies do gênero *Coffea*.

FIGURA 6 – Árvore filogenética contendo sete sequências de proteínas traduzidas a partir do gene *FT* de diferentes famílias e construída pelo software MEGA 5.2. As sequências estão separadas por clados e agrupadas de acordo com a similaridade através de alinhamento pelo programa ClustalW submetido a bootstrap de 1000 como parâmetro estatístico. As numerações nos clados indicam a porcentagem em que as sequências foram agrupadas em conjunto em todas as vezes que a árvore foi gerada (bootstrap de 1000 repetições). *Arabidopsis thaliana* (NM_105222.2), *Nicotiana tomentosiformis* (KC485968.1), *Solanum tuberosum* (NM_001287968.1), *Coffea arabica* (CAFEST), *Coffea canephora* (chr10:7950878..7951201), *Medicago truncatula* (XM_013596135.1) e *Glycine max* (AB550125.1).



A análise de domínio mostrou que todas as sequências buscadas apresentaram o domínio PEBP (Phosphoethanolamine Binding Protein). A proteína PEBP é uma pequena proteína que agrupa o gene *FT* na mesma superfamília (LI et al., 2012; ZHENG et al., 2016). A proteína PEBD está presente na sinalização de rotas do florescimento e de crescimento (CHAUTARD et al., 2004). A presença do domínio PEBP na sequência do *FT* de *C. arabica* é um indicativo de que o gene codifica para a proteína FT uma vez que esse domínio está presente em proteína FT de outras espécies (ZHENG et al., 2016).

FIGURA 7 – Comparação das sequências do gene *FT* através da análise de domínio pela ferramenta CDD (Conserved Domain Database) no banco de dados NCBI. **A** – *Coffea arabica* (Rubiaceae); **B** – *Coffea canephora* (Rubiaceae); **C** – *Arabidopsis Thaliana* (Brassicaceae); **D** – *Nicotiana tomentosiformis* (Solanaceae); **E** – *Solanum tuberosum* (Solanaceae); **F** – *Medicago truncalata* (Fabaceae); **G** – *Glycine Max* (Fabaceae). A região do domínio na sequência é mostrada através da barra com o nome do domínio ao lado da sequência.

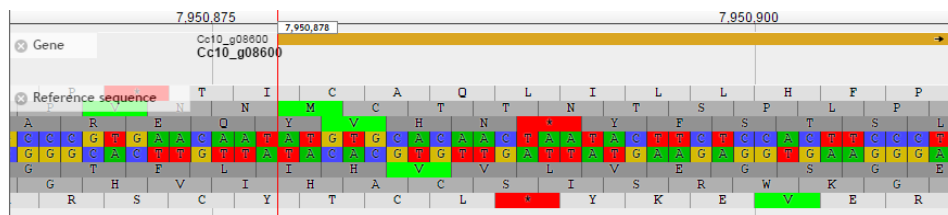


7.1.2 Busca da região promotora do gene *FT*

Após buscas da região promotora do gene *FT* usando a ferramenta Blastn no Coffee Genome Hub, a sequência do *FT* de *C. arabica* foi alinhada com a sequência do *FT* de *C. canephora* no cromossomo 10 do genoma. A partir do início da sequência alinhada com o DNA foi possível selecionar a região upstream do gene. A região separada para desenho dos *primers* foi de -1080 pb contados a partir do ATG da região codante do gene usado no alinhamento com o DNA de *C. canephora*. Com o alinhamento no genoma de *C. canephora*,

foi possível separar parte da sequência promotora do gene *FT* (-1080 pb) e usar como molde no desenho dos *primers* para as demais espécies.

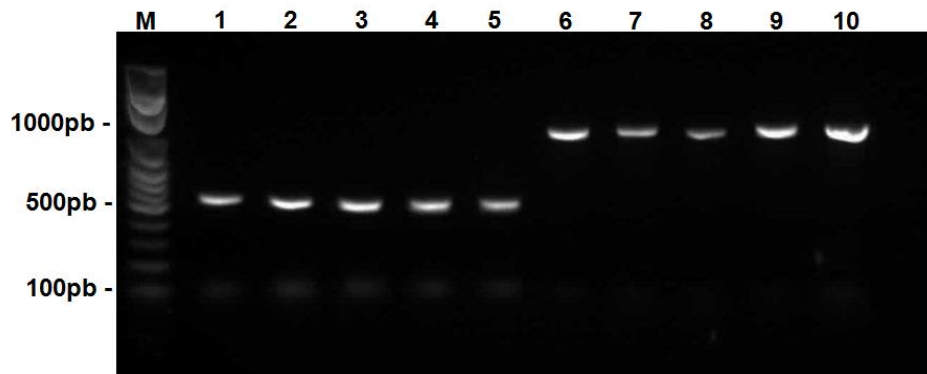
FIGURA 8 – Esquema da região codante e promotora do gene *FT* através do alinhamento da sequência do gene de *C. arabica* no genoma da espécie *C. Canephora* obtida pela ferramenta (colocar o programa que gerou a figura). A barra amarela mostra o sentido e o alinhamento do gene com o genoma e a linha vermelha cortando a sequência mostra o ponto da divisão entre o ATG da região codante e a região upstreame (-1) no genoma.



7.2 Isolamento das sequências codante e promotora do gene *FT*

Através de eletroforese com gel de Agarose realizada com o produto da PCR usando o cDNA e DNA do material genômico representante das quatro espécies e os *primers* desenhados no início e fim da região codante e promotora do gene *FT* de *C. arabica* e *C. canephora*, verificou-se que houve amplificação dos fragmentos em todos os genótipos avaliados. Com base no marcador de peso molecular, as bandas formadas ficaram próximas de 500 pb para a região codante e 1000 pb para a região promotora em todos os genótipos, como mostra a imagem do gel (Fig. 9).

FIGURA 9 – Eletroforese em gel de agarose mostrando os fragmentos gerados em PCR com os respectivos *primers* da região codante (Fw4 e Rv1) e os *primers* da região promotora (Fw1 – Rv2) do gene *FT* nos cinco genótipos. As bandas de um a cinco correspondem à região codante do gene, as bandas de seis a dez correspondem à região promotora do gene *FT*. O tamanho dos fragmentos corresponde ao esperado pela predição das sequências conforme indicado pelo marcador, ou seja, 500 pb para a região codante e 1000 pb para o promotor. M – Marcador de peso molecular (100 pb); 1 e 6 = 37 (Semperflorens *C. arabica*); 2 e 7 = 39 Robusta *C. canephora*); 3 e 8 = 57 (Racemosa *C. racemosa*); 4 e 9 = ACA (Acauã *C. arabica*); 5 e 10 = EUG (Eugenióides *C. eugenioides*).



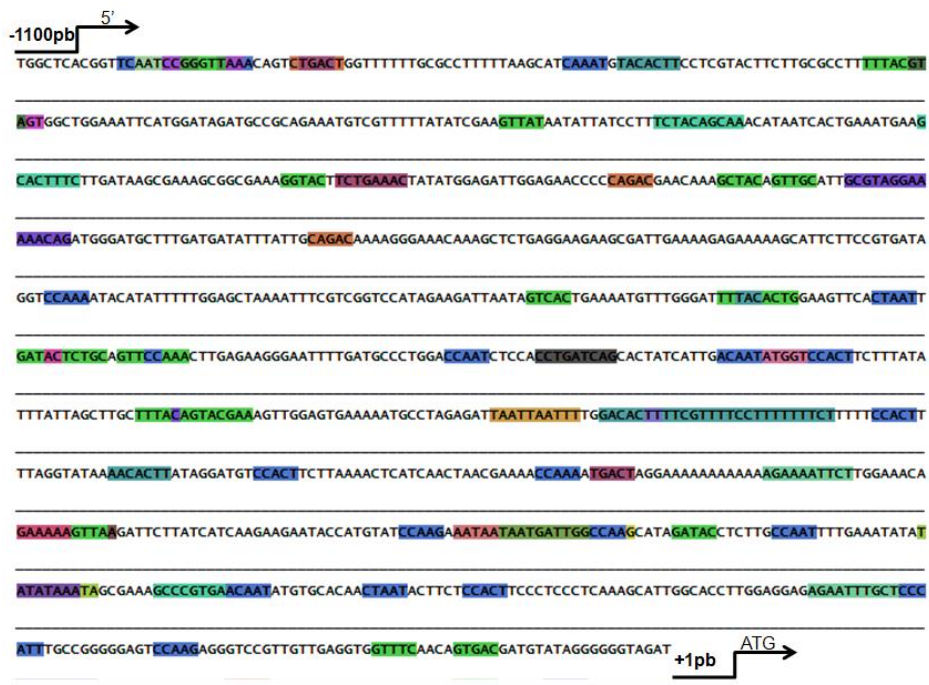
A amplificação dos fragmentos da região promotora e codante do gene *FT* entre diferentes espécies mostra que os *primers* de uma espécie são capazes de amplificar fragmentos em outra e, mais importante que isso, indica que o tamanho do gene não varia entre as espécies. Esse resultado comprova o esperado, a partir das análises obtidas *in silico* mostrando aproximação entre as espécies de *Coffea*.

7.3 Sequências *cis*-reguladoras

Com as buscas realizadas no site PlantCARE e PlantPAN através das regiões promotoras sequenciadas, foi encontrada várias sequências *cis*-

reguladoras de fatores de transcrição além da região TATA BOX responsável por direcionar as proteínas RNA polimerase II fundamentais para o início da transcrição do gene (BREATHNACH & CHAMBON, 1981). Os motivos encontrados nas sequências (Fig. 7) estão relacionados a diversos fatores de transcrição coordenados por diferentes estímulos.

FIGURA 10- Regiões *cis*-reguladoras encontradas na sequência promotora do gene *FT* do genoma de *C. canephora* através de busca no banco de dados PlantPAN 2.0. Cada região marcada representa um motivo. Cada grupo de motivos indicados por uma cor está relacionado a um tipo de fator de transcrição. No início da sequência é indicado o sentido (5' para o 3') e o tamanho da sequência (-1100 pb). O final da sequência é indicado pelo nucleotídeo +1 e a indicação do sentido do início da transcrição da região codante (ATG).



Dentre os vários elementos encontrados existem alguns que estão ligados ao controle de genes relacionados ao florescimento como o *FT*. Foram encontrados elementos *cis* CCAAT-Box que agrupam motivos de ligação a fatores de transcrição mediadores da ação da proteína CO na ativação da expressão do gene *FT* (ROBSON et al., 2001; WIGGE et al., 2005; TIWARI et al., 2010), alguns motivos comuns em eucariotos como o GATA-motivo e o grupo I-Box presente em genes regulados pela luz em monocotiledonias e eudicotiledonias (TERZAGHI & CASHMORE, 1995). Além de outros

relacionados à resposta ao estresse biótico e abiótico e a hormônio como o etileno

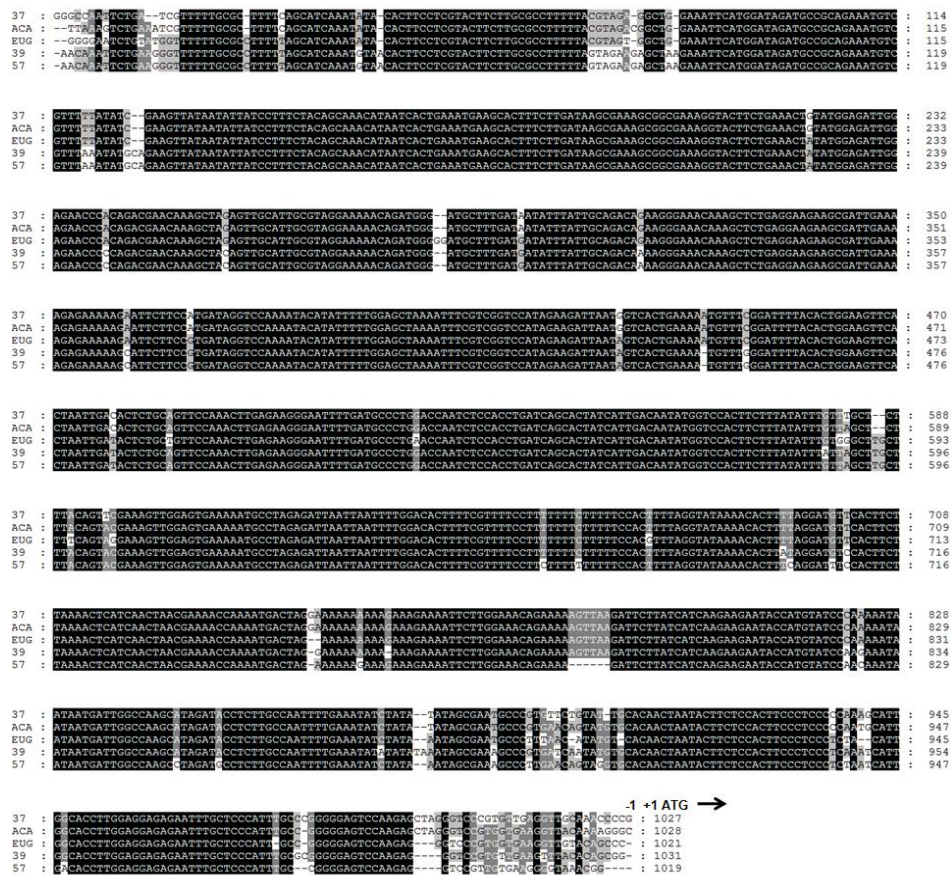
TABELA 3 – Fatores de transcrição que se ligam a motivos encontrados na região promotora dos cinco genótipos representante das quatro espécies de *Coffea*, através da busca de motivos no PlantPAN 2.0. Os fatores de transcrição descritos apresentam relações com o controle da expressão de genes do florescimento. A primeira coluna mostra os fatores encontrados e a segunda coluna descreve o envolvimento de cada um dos fatores na regulação da expressão gênica.

Cis-elementos (Fator de transcrição)	Envolvimento
B3	Envolvido na rota do florescimento (FRANCO-ZORRILLA <i>et al.</i> , 2014).
VOZ	Atua como repressor da expressão do gene <i>FLC</i> e ativador da expressão do gene <i>FT</i> (MITSUDA <i>et al.</i> , 2004).
ZF-HD	Relacionado a genes reguladores da arquitetura floral e desenvolvimento foliar (WINDHÖVEL <i>et al.</i> , 2001).
MADS-box; MIKC	Envolvidos no controle do tempo de florescimento através das rotas de vernalização, fotoperíodo e autônomas (MATHELIER <i>et al.</i> , 2013).
NF-Y	O complexo NF-Y é responsável pelo estímulo da expressão de vários genes através da ligação em motivos do grupo CCAAT-box (MANTOVANI, 1999).
SRS	Relacionado à expressão de genes do desenvolvimento de estames e expansão celular bem como no controle do tempo de florescimento (KUUSK <i>et al.</i> , 2006).

7.4 Polimorfismo na região promotora do gene *FT*

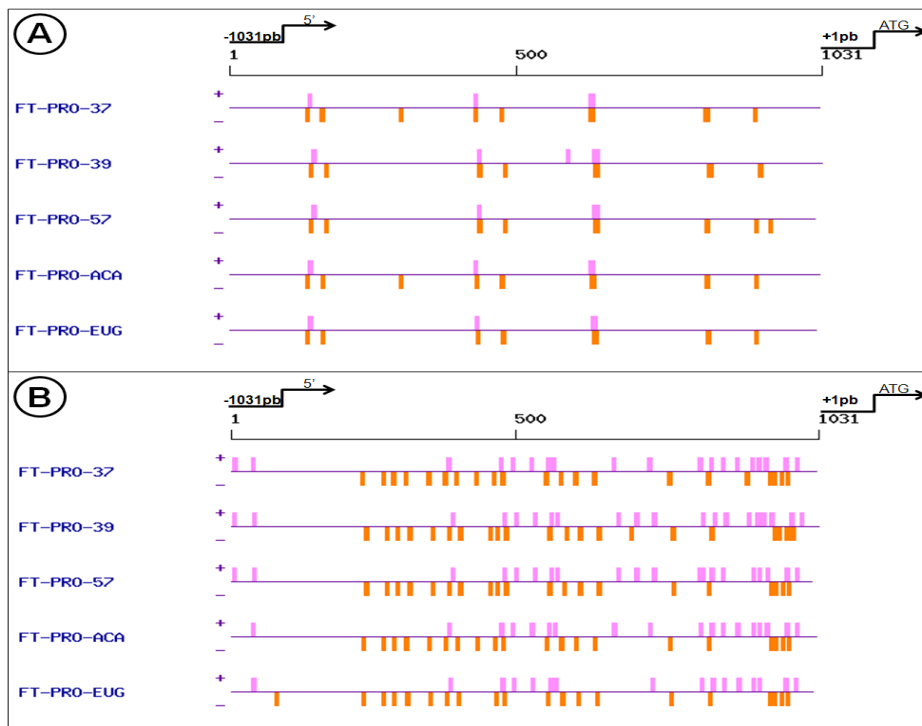
Ao alinhar as sequências promotoras, nota-se que existe uma grande similaridade entre as quatro espécies, porém, as alterações de bases existentes entre uma e outra estão associados a motivos de ligação de fatores de transcrição. Essas alterações podem estar relacionadas ao controle da expressão do gene *FT*, em consequência disso, levar a diferenças na expressão do gene quando comparado entre as espécies.

FIGURA 11 – Alinhamento das sequências (primeiros -1000 pb upstream) da região promotora dos cinco genótipos. O alinhamento foi gerado pelo programa ClustalW e exibido com o GeneDoc (Fabricante). O início da região promotora é indicado na parte inferior da figura pelo +1 pb ao lado da seta de indicação do sentido do início da transcrição do mRNA (ATG), a esquerda das sequências estão as identificações de cada genótipo (**37** – Semperflorens (*C. arabica*); **39** – Robusta (*C. canephora*); **57** – Racemosa (*C. racemosa*); **ACA** – Acauã (*C. arabica*) e **EUG** – Eugenióides (*C. eugenoides*)) e a direita estão os números de nucleotídeos. Os traços na horizontal indicam lacunas inseridas para a otimização do alinhamento. As regiões mais escuras indicam locais de similaridade entre as sequências, as regiões mais claras e brancas indicam pontos de diferença entre duas ou mais sequência.



Ao comparar as sequências através do PlantPAN 2.0 nota-se que para alguns dos elementos citados na tabela 2, existe diferenças no número de motivos entre algumas das sequências. Essa alteração está associada aos polimorfismos encontrados nas sequências. Alterações nessas sequências impedem à ligação dos fatores de transcrição e podendo influenciar sobre a expressão do gene (KORKUĆ et al., 2014).

FIGURA 12- Comparação da diferença do número de motivos dos fatores de transcrição ZF-HD (A) e NF-Y (B) nas sequências da região promotora dos cinco genótipos através do PlantPLAN 2.0. O local do motivo na linha de representação da sequência está identificado com uma barra lilás no sentido 5' – 3' e uma barra laranja na sequência complementar 3'-5' abaixo. **FT-PRO-37** = Semperflorens (*C. arabica*); **FT-PRO-39** = Robusta (*C. canephora*); **FT-PRO-57** = Racemosa (*C. racemosa*); **FT-PRO-ACA** = Acauã (*C. arabica*) e **FT-PRO-EUG** = Eugenióides (*C. eugenioides*).



Através dos alinhamentos das sequências promotoras, observando a sequência 5'-3' é possível observar no alinhamento A que representa o fator de transcrição ZF-HD, que o genótipo 39 possui um motivo a mais em comparação aos demais genótipos. Já no alinhamento B referente ao fator de transcrição NF-Y é possível observar que os genótipos 39 e 57 possuem 19 motivos, os genótipos 37 e ACA possuem 18 motivos e o EUG 17 motivos.

O fator de transcrição ZF-HD está envolvido na regulação de genes de formação dos órgãos florais e desenvolvimento foliar (MANTOVANI, 1999). Já o fator de transcrição NF-Y faz parte de um complexo envolvido na regulação da transcrição de vários genes. Os motivos de identificação desse complexo são os CCAAT-box, através desse complexo ele media a ação de fatores de transcrição como o *CONSTANS* (WIGGE et al., 2005; TIWARI et al., 2010).

7.5 Sequências codantes

Após o sequenciamento e alinhamento das regiões codantes do gene *FT* dos cinco genótipos (Fig. 13?), é possível notar que, apesar de apresentarem uma alta identidade entre si (93% a 96%) diferindo também no tamanho total da sequência. Os locais de maior diferença entre os genótipos estão localizados na porção central das sequências, entre +170 pb e +220 pb e +235 pb e +280 pb. As alterações observadas são de SNPs, por deleção e adição quando comparado entre os genótipos.

É normal que haja diferença nas sequências do gene entre os genótipos, pois são de espécies diferentes. Isso pode estar associado a diversos fatores do processo evolutivo das espécies, mas não implica que o gene venha a ter função diferente, pois, o código genético é degenerado e talvez uma alteração em um nucleotídeo pode não necessariamente alterar o aminoácido na proteína e mesmo se alterar por um aminoácido semelhante, por exemplo, um aminoácido

básico por outro básico pode ser que não cause grande alteração na função da proteína (MATIOLI & FERNANDES, 2012).

FIGURA 13 - Alinhamento das seqüências codante dos cinco genótipos. O alinhamento foi gerado pelo programa Clustalw e exibido com o GeneDoc. A esquerda das seqüências estão as identificações de cada genótipo, sendo: **57** – Racemosa (*C. racemosa*); **EUG** – Eugenióides (*C. eugenoides*); **ACA** – Acauã (*C. arabica*); **37** – Semperflorens (*C. arabica*) e **39** – Robusta (*C. canephora*). À direita estão a faixa de nucleotídeos de cada linha. Os traços na horizontal indicam lacunas inseridas para a otimização do alinhamento. As regiões mais escuras indicam locais de similaridade entre as seqüências, as regiões mais claras e brancas indicam pontos de diferença entre duas ou mais seqüência.

```

FT-RNA-57 : TCACCATGTATAGGGGGTAGATCCTCTAGTGGTTGGACGTTGATAGCAGATGTTTTGGACCCCTTCAACAGGTCCTA : 80
FT-RNA-EUG : TCACCATGTATAGGGGGTAGATCCTCTAGTGGTTGGACGTTGATAGCAGATGTTTTGGACCCCTTCAACAGGTCCTA : 80
FT-RNA-ACA : TCACCATGTATAGGGGGTAGATCCTCTAGTGGTTGGACGTTGATAGCAGATGTTTTGGACCCCTTCAACAGGTCCTA : 80
FT-RNA-37 : TCACCATGTATAGGGGGTAGATCCTCTAGTGGTTGGACGTTGATAGCAGATGTTTTGGACCCCTTCAACAGGTCCTA : 80
FT-RNA-39 : TCACCATGTATAGGGGGTAGATCCTCTAGTGGTTGGACGTTGATAGCAGATGTTTTGGACCCCTTCAACAGGTCCTA : 80

FT-RNA-57 : CCCCTTAGAGTTATCCATGGCAACAGAGACGTGATTAATGGTTGTGAGTTTAC-ACCCTCTCAGGTCGTTAGCCAGCCTA : 159
FT-RNA-EUG : CCCCTTAGAGTTATCCATGGCAACAGAGACGTGATTAATGGTTGTGAGTTTAC-ACCCTCTCAGGTCGTTAGCCAGCCTA : 160
FT-RNA-ACA : CCCCTTAGAGTTATCCATGGCAACAGAGACGTGATTAATGGTTGTGAGTTTAC-ACCCTCTCAGGTCGTTAGCCAGCCTA : 159
FT-RNA-37 : CCCCTTAGAGTTATCCATGGCAACAGAGACGTGATTAATGGTTGTGAGTTTAC-ACCCTCTCAGGTCGTTAGCCAGCCTA : 159
FT-RNA-39 : CCCCTTAGAGTTATCCATGGCAACAGAGACGTGATTAATGGTTGTGAGTTTAC-ACCCTCTCAGGTCGTTAGCCAGCCTA : 159

FT-RNA-57 : GGGTTGAGATTGG---AGCGATGATCTCGTAAGCTTCT-AACTCTTACCTCAGTGGTGGCTCCAAGTCCCAGTAAT : 235
FT-RNA-EUG : GGGTTGAGATTGGGAAGCTGGATGATCTCGTAAGCTTACGG-ACTCTTACCTCAGTGGTGGCTCCAAGTCCCAGTAAT : 239
FT-RNA-ACA : GGGTTGAGATTGGG---AGCGATGATCTCGTAAGCTTACGG-ACTCTTACCTCAGTGGTGGCTCCAAGTCCCAGTAAT : 236
FT-RNA-37 : GGGTTGAGATTGGAGGCTGGAT---TCGTAG---CTCTACTCTCTGGTGGTGGCTCCAAGTCCCAGTAAT : 229
FT-RNA-39 : GGGTTGAGATTGGAGGCTGGAT---TCGTAG---CTCTACTCTCTGGTGGTGGCTCCAAGTCCCAGTAAT : 228

FT-RNA-57 : CCCATT---TGGGATTTTGGCCCTGAGTTGC-TCATGATATCCAGCCACCAGGAGCATCATTCGGCCAAGAAATT : 312
FT-RNA-EUG : CCCATG---CCGAGGTTTGGCACTGGTTGGCACTGATATCCAGCCACCAGGAGCATCATTCGGCCAAGAAATT : 317
FT-RNA-ACA : CCCATG---CCAGGATTTGGCACTGGTTGG-TCATGATATCCAGCCACCAGGAGCATCATTCGGCCAAGAAATT : 313
FT-RNA-37 : CCCATGGGCGGATATTTGGCCCTGGTTGG-GTCTGATATCCAGCCACCAGGAGCATCATTCGGCCAAGAAATT : 308
FT-RNA-39 : CCCATG---TGGGCAATTTGGCACTGGTTGC-TCATGATATCCAGCCACCAGGAGCATCATTCGGCCAAGAAATT : 305

FT-RNA-57 : GGTGCTACAGAGTCCACGGCCATCAATGGGAATCCACCGCTTCGTCTTGGCGTTTCCGACAGCTTGGCGGCAGAC : 392
FT-RNA-EUG : GGTGCTACAGAGTCCACGGCCATCAATGGGAATCCACCGCTTCGTCTTGGCGTTTCCGACAGCTTGGCGGCAGAC : 397
FT-RNA-ACA : GGTGCTACAGAGTCCACGGCCATCAATGGGAATCCACCGCTTCGTCTTGGCGTTTCCGACAGCTTGGCGGCAGAC : 393
FT-RNA-37 : GGTGCTAG--AGAGTCCACGGCCATCAATGGGAATCCACCGCTTCGTCTTGGCGTTTCCGACAGCTTGGCGGCAGAC : 386
FT-RNA-39 : GGTGCTACAGAGTCCACGGCCATCAATGGGAATCCACCGCTTCGTCTTGGCGTTTCCGACAGCTTGGCGGCAGAC : 385

FT-RNA-57 : AGTCTATGCTCCGGGGTGGCGCCAGAACTTCAGCACCAGGGACTTTGCTGAACCTTACAATTTAGGTTCCGCGTTGCAG : 472
FT-RNA-EUG : AGTCTATGCTCCGGGGTGGCGCCAGAACTTCAGCACCAGGGACTTTGCTGAACCTTACAATTTAGGTTCCGCGTTGCAG : 477
FT-RNA-ACA : AGTCTATGCTCCGGGGTGGCGCCAGAACTTCAGCACCAGGGACTTTGCTGAACCTTACAATTTAGGTTCCGCGTTGCAG : 473
FT-RNA-37 : AGTCTATGCTCCGGGGTGGCGCCAGAACTTCAGCACCAGGGACTTTGCTGAACCTTACAATTTAGGTTCCGCGTTGCAG : 466
FT-RNA-39 : AGTCTATGCTCCGGGGTGGCGCCAGAACTTCAGCACCAGGGACTTTGCTGAACCTTACAATTTAGGTTCCGCGTTGCAG : 465

FT-RNA-57 : CAATTTACTACAATTGCCAGAGGGAAGGTGGATCTGGAGGCAGAAGATCATCATGA- : 528
FT-RNA-EUG : CAATTTACTACAATTGCCAGAGGGAAGGTGGATCTGGAGGCAGAAGATCATCATGA : 534
FT-RNA-ACA : CAATTTACTACAATTGCCAGAGGGAAGGTGGATCTGGAGGCAGAAGATCATCATGA : 530
FT-RNA-37 : CAATTTACTACAATTGCCAGAGGGAAGGTGGATCT----- : 501
FT-RNA-39 : CAATTTACTACAATTGCCAGAGGGAAGGTGGATCTGGAGGCAGAAGATCATCATGA : 522

```

7.6 Análise da expressão gênica

A análise da expressão gênica mostrou que houve diferenças significativas na expressão dos genes *FT* e *CO* entre os cinco genótipos estudados (Figura 14). O gene *CO* apresentou maiores níveis de expressão nos genótipos 37, 39, e EUG, não diferindo entre os mesmos, sendo relativamente menos expressos nos genótipos ACA e 57, onde se mostrou até 18 vezes menos expressos no caso do genótipo 57, comparado com seu nível de expressão em EUG (Figura 14). Já para o gene *FT*, pode ser observado que este gene apresentou o maior nível de expressão no genótipo EUG, sendo duas vezes mais expresso que no genótipo ACA e até 42 vezes mais expresso quando comparados com os genótipos 37, 39 e 57, os quais apresentaram os níveis expressão semelhantes para esse gene (Figura 14).

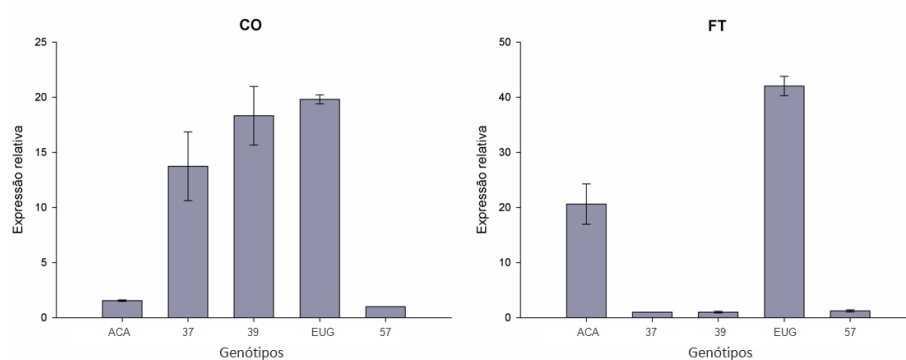
Alguns estudos mostram que o gene *CO* está diretamente ligado à ativação do gene *FT* (SAMACH et al., 2000; PUTTERILL et al., 2004; ABE et al., 2005). Apesar de ter sido observado que níveis relativamente altos e baixos de expressão do gene *CO* estarem associados com níveis relativamente altos e baixos do gene *FT* nos genótipos EUG e 57, respectivamente, esse perfil de resposta não foi observado para os outros genótipos (Figura 14). Para os genótipos 37 e 39, níveis de expressão relativamente mais baixos do gene *FT* foram encontrados, apesar dos níveis relativamente mais altos do gene *CO* (Figura 14). Já para o genótipo ACA, mesmo com níveis relativamente mais baixos de expressão para o gene *CO*, um nível relativamente alto de expressão para o gene *FT* pode ser encontrado (Figura 14).

Isso pode estar associado ao fato de que, além do *CO*, a regulação do gene *FT* está associada a outros fatores de transcrição, como observado pela detecção dos vários motivos de elementos *cis*-regulatórios encontrados na região promotora dos genes *FT* (Figura 10), e ao fato de que polimorfismos na sua

região promotora do gene *FT*, como os encontrados nesse trabalho (Figura 11), podem também estar associados a variações no perfil de expressão desse gene entre os diferentes genótipos analisados. A Região proximal do sítio de início da transcrição, aproximadamente de -1Kb, está diretamente ligada à ação de elementos relacionados ao controle da expressão do gene *FT* em *Arabidopsis thaliana* (ADRIAN et al., 2010; LIU et al., 2014). Além disso, a regulação da expressão dos genes *CO* e *FT* está associada a diversos fatores como fotoperíodo, vernalização, idade, além de outros que não se sabe ainda (ALBANI & COUPLAND, 2010; PUTTERILL et al., 2013) que podem ter influenciado no resultado obtido.

Essa relação pode estar associada à ação do *CO* na ativação da expressão do *FT* em que, a alta expressão do *CO* coincide com a alta expressão do *FT*. Porém ao associar a expressão gênica às características de florescimento dos genótipos, não houve correlação.

FIGURA 14 –Perfil da expressão relativa por RT-qPCR dos genes *CO* e *FT* em folhas de cinco genótipos diferentes do cafeeiro. As barras representam a diferença na expressão em cada um dos cinco genótipos em relação a uma amostra normalizadora (genótipo 57 para ambos os genes). Os valores de expressão de cada genótipo foram obtidos a partir de três repetições biológicas e as barras de erro representam o erro padrão entre elas. Os valores de expressão foram normalizados pelos genes de referência *RPL39* e *Ubiquitina*. 37 – Semperflorens (*C. arabica*); 39 – Robusta (*C. canephora*); 57 – Racemosa (*C. racemosa*); ACA – Acauã (*C. arabica*) e EUG – Eugenióides (*C. eugenioides*).



Mesmo no genótipo Semperflorens, o qual apresenta um padrão de florescimento diferente das outras espécies, não foi observados padrões de expressão muito diferentes quando comparado com os outros genótipos.

8. CONCLUSÃO

A região promotora dos genes *FT* identificados nas quatro espécies (*Coffea arabica*, *Coffea canephora*, *Coffea eugenioides* e *Coffea racemosa*) e cinco genótipos (Semperflorens, Robusta, Acauã, Racemosa e Eugenióides) do cafeeiro analisados nesse trabalho apresentam polimorfismos do tipo SNP, deleção e adição de nucleotídeos.

O gene *CO* é diferencialmente expresso entre as espécies de cafeeiro, apresentando níveis de expressão mais elevados nos genótipos Semperflorens, Robusta, e Eugenióides, em comparação aos genótipos Acauã e Racemosa.

O gene *FT* é diferencialmente expresso entre as espécies de cafeeiro, apresentando níveis de expressão mais elevados nos genótipos Eugenióides e Acauã, em comparação aos genótipos Semperflorens, Robusta, e Racemosa. Isso pode explicar diferenças observadas durante o desenvolvimento reprodutivo das espécies. Concluir com futuras análises.

9. CONSIDERAÇÕES FINAIS

Este estudo identificou um fragmento de 1kb da região promotora do gene *FT*, bem como da região codante de quatro espécies pertencentes ao gênero *Coffea* (*Coffea arabica*, *Coffea canephora*, *Coffea eugenioides* e *Coffea racemosa*) presentes em cinco genótipos (Semperflorens, Robusta, Acauã, Racemosa e Eugenióides). As análises dessas regiões identificaram elementos *cis*-regulatórios os quais podem estar diretamente relacionados na regulação da expressão do gene *FT*. Além disso, o trabalho e permitiu a identificação de polimorfismos nessas regiões e entre os diferentes genótipos que podem estar envolvidos na regulação dos diferentes perfis de expressão do *FT*. De acordo como isso, demonstramos que os perfis de expressão do gene *FT* variam entre as espécies e isso poderia estar relacionado às diferenças quanto ao período de florescimento observados em cada espécie.

Estudos futuros envolvendo análises de expressão do gene *FT* em diferentes estádios do desenvolvimento do cafeeiro, e estudos de análise funcional desse gene em espécies modelo e no cafeeiro são necessários para uma melhor compreensão do papel desse gene no controle do florescimento do cafeeiro e nas suas diferentes espécies e genótipos.

10. REFERÊNCIAS

ABE, M. et al. A bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. **Science**, v. 309, n. 5737, p. 1052-1056, 2005. ISSN 0036-8075.

ALBANI, M. C.; COUPLAND, G. Chapter eleven-comparative analysis of flowering in annual and perennial plants. **Current topics in developmental biology**, v. 91, p. 323-348, 2010. ISSN 0070-2153.

ANDRÉS, F.; COUPLAND, G. The genetic basis of flowering responses to seasonal cues. **Nature Reviews Genetics**, v. 13, n. 9, p. 627-639, 2012. ISSN 1471-0056.

BLÜMEL, M.; DALLY, N.; JUNG, C. Flowering time regulation in crops—what did we learn from Arabidopsis? **Current opinion in biotechnology**, v. 32, p. 121-129, 2015. ISSN 0958-1669.

BREATHNACH, R.; CHAMBON, P. Organization and expression of eucaryotic split genes coding for proteins. **Annual review of biochemistry**, v. 50, n. 1, p. 349-383, 1981. ISSN 0066-4154.

CHAUTARD, H. et al. Tfs1p, a member of the PEBP family, inhibits the Ira2p but not the Ira1p Ras GTPase-activating protein in *Saccharomyces cerevisiae*. **Eukaryotic Cell**, v. 3, n. 2, p. 459-470, 2004. ISSN 1535-9778.

CORBESIER, L. et al. FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. **science**, v. 316, n. 5827, p. 1030-1033, 2007. ISSN 0036-8075.

DOYLE, J. J. Isolation of plant DNA from fresh tissue. **Focus**, v. 12, p. 13-15, 1990.

FRANCO-ZORRILLA, J. M. et al. DNA-binding specificities of plant transcription factors and their potential to define target genes. **Proceedings of the National Academy of Sciences**, v. 111, n. 6, p. 2367-2372, 2014. ISSN 0027-8424.

GOUVEIA, N. Florescimento e maturação em *Coffea arabica* cv. Catuai vermelho. 9. Congresso Brasileiro de Pesquisas Cafeeiras 27-30 Oct 1981 São Lourenço, MG (Brasil), 1981, Instituto Brasileiro do Café, Rio de Janeiro (Brasil).

KORKUĆ, P.; SCHIPPERS, J. H.; WALTHER, D. Characterization and identification of cis-regulatory elements in *Arabidopsis* based on single-nucleotide polymorphism information. **Plant physiology**, v. 164, n. 1, p. 181-200, 2014. ISSN 1532-2548.

KUUSK, S.; SOHLBERG, J. J.; MAGNUS EKLUND, D.; SUNDBERG, E. Functionally redundant SHI family genes regulate *Arabidopsis* gynoecium development in a dose-dependent manner. **The Plant Journal**, v. 47, n. 1, p. 99-111, 2006. ISSN 1365-313X.

MANTOVANI, R. The molecular biology of the CCAAT-binding factor NF-Y. **Gene**, v. 239, n. 1, p. 15-27, 1999. ISSN 0378-1119.

MATHELIER, A. et al. An extensively expanded and updated open-access database of transcription factor binding profiles. **Nucleic acids research**, p. gkt997, 2013. ISSN 0305-1048.

MATIOLI, S. R.; FERNANDES, F. M. D. C. Biologia molecular e evolução. In: (Ed.). **Biologia molecular e evolução**: Holos/Sociedade Brasileira de Genética, 2012.

MELO, B.; SOUSA, L. B. BIOLOGIA DA REPRODUÇÃO DE *Coffea arabica* L. e *Coffea canephora* Pierre. **Revista Verde de Agroecologia e Desenvolvimento Sustentável**, v. 6, n. 2, 2011. ISSN 1981-8203.

MITSUDA, N. et al. isolation and characterization of novel vascular plant transcription factors with a one-zinc finger from *Arabidopsis thaliana*. **Plant and Cell Physiology**, v. 45, n. 7, p. 845-854, 2004. ISSN 0032-0781.

MURRAY, M.; THOMPSON, W. F. Rapid isolation of high molecular weight plant DNA. **Nucleic acids research**, v. 8, n. 19, p. 4321-4326, 1980. ISSN 0305-1048.

NICHOLAS, K. B.; NICHOLAS, H.; DEERFIELD, D. GeneDoc: analysis and visualization of genetic variation. **Embnew. news**, v. 4, n. 1, 1997.

OWCZARZY, R. et al. IDT SciTools: a suite for analysis and design of nucleic acid oligomers. **Nucleic acids research**, v. 36, n. suppl 2, p. W163-W169, 2008. ISSN 0305-1048.

PFAFFL, M. W. A new mathematical model for relative quantification in real-time RT-PCR. **Nucleic acids research**, v. 29, n. 9, p. e45-e45, 2001. ISSN 0305-1048.

PUTTERILL, J.; LAURIE, R.; MACKNIGHT, R. It's time to flower: the genetic control of flowering time. **Bioessays**, v. 26, n. 4, p. 363-373, 2004. ISSN 1521-1878.

PUTTERILL, J. et al. FT genes and regulation of flowering in the legume *Medicago truncatula*. **Functional Plant Biology**, v. 40, n. 12, p. 1199-1207, 2013. ISSN 1445-4416.

SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular biology and evolution**, v. 4, n. 4, p. 406-425, 1987. ISSN 0737-4038.

SAMACH, A. et al. Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. **Science**, v. 288, n. 5471, p. 1613-1616, 2000. ISSN 0036-8075.

SITNIKOVA, T.; RZHETSKY, A.; NEI, M. Interior-branch and bootstrap tests of phylogenetic trees. **Molecular Biology and Evolution**, v. 12, n. 2, p. 319-333, 1995. ISSN 0737-4038.

TAMURA, K. et al. Molecular evolutionary genetics analysis (MEGA) software version 4.0. **Molecular biology and evolution**, v. 24, n. 8, p. 1596-1599, 2007. ISSN 0737-4038.

TERZAGHI, W. B.; CASHMORE, A. R. Light-regulated transcription. **Annual review of plant biology**, v. 46, n. 1, p. 445-474, 1995. ISSN 1040-2519.

THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic acids research**, v. 22, n. 22, p. 4673-4680, 1994. ISSN 0305-1048.

WENKEL, S. et al. CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of Arabidopsis. **The Plant Cell**, v. 18, n. 11, p. 2971-2984, 2006. ISSN 1532-298X.

WINDHÖVEL, A. et al. Characterization of a novel class of plant homeodomain proteins that bind to the C4 phosphoenolpyruvate carboxylase gene of Flaveria trinervia. **Plant molecular biology**, v. 45, n. 2, p. 201-214, 2001. ISSN 0167-4412.