



SAMUEL CHAVES SILVA

**CARACTERIZAÇÃO *in silico* E ANÁLISE DA
EXPRESSÃO DE microRNAs E GENES ALVO
EM DIFERENTES ESPÉCIES DE *Coffea*.**

**Lavras – MG
2012**

SAMUEL CHAVES SILVA

**CARACTERIZAÇÃO *in silico* E ANÁLISE DA EXPRESSÃO DE
microRNAs E GENES ALVO EM DIFERENTES ESPÉCIES DE *Coffea*.**

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

Orientador

Antonio Chalfun Júnior

**LAVRAS – MG
2012**

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

Silva, Samuel Chaves.

Caracterização *in silico* e análise da expressão de microRNAs e genes alvo em diferentes espécies de *Coffea* / Samuel Chaves Silva.
– Lavras : UFLA, 2012.

104 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2012.

Orientador: Antonio Chalfun Júnior

Bibliografia.

1. Café. 2. Microtranscriptoma. 3. qPCR. 4. RACE. 5. Arf. 8. I.
Universidade Federal de Lavras. II. Título.

CDD – 583.5204873283

SAMUEL CHAVES SILVA

**CARACTERIZAÇÃO *in silico* E ANÁLISE DA EXPRESSÃO DE
microRNAs E GENES ALVO EM DIFERENTES ESPÉCIES DE *Coffea*.**

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

APROVADA em 24 de fevereiro de 2012.

Alan Carvalho Andrade, PhD – (Embrapa Cenargen)

Dr. Fabio Tebaldi Silveira Nogueira – (IBB/UNESP)

Antonio Chalfun Júnior, PhD
Orientador

**LAVRAS – MG
2012**

*Às pessoas que tanto me apoiaram e me
ergueram nos momentos difíceis: meus avós
Ana e Antonio, minha mãe Klênia e tia Neila.
Obrigado por estarem sempre comigo, vocês
fazem parte desta conquista!*

DEDICO

“Everybody is a genius. But if you judge a fish by its ability to climb a tree, it will live its whole life believing that it is stupid.”

Albert Einstein

RESUMO

O fenômeno de RNA de interferência mediado por microRNAs, é uma via importante para a estabilidade e regulação da tradução de RNA mensageiros. Muitas das vias reguladas são fundamentais para o desenvolvimento do organismo e para a manutenção da integridade do genoma. Alta homologia em sequência e estrutura são características interessantes dos genes que codificam estas moléculas. Estas particularidades permitem que, a partir de sequências de RNAs silenciadores conhecidos em outras espécies, microRNAs homólogos sejam descritos em espécies filogeneticamente relacionadas. A partir das informações geradas pelo sequenciamento do café, durante o projeto CAFEST, foi possível, pela primeira vez, identificar microRNAs (miRNAs) envolvidos na modelagem do transcriptoma das espécies *Coffea arabica*, *Coffea canephora*, *Coffea congensis* e *Coffea eugenioides*. Neste estudo são apresentadas as sequências, estruturas e possíveis funções dos microRNAs no cafeeiro. Além da caracterização estrutural, também foram apresentados resultados de qRT-PCR que correlacionam a expressão dos genes MIR e seus genes alvos. As diversas famílias de miRNAs descritas para o cafeeiro, estão correlacionadas a diferentes processos biológicos que incluem, padrões de crescimento vegetal, processos metabólicos, respostas hormonais, degradação de proteínas, respostas a estresses e rotas de sinalização celular. A grande variedade de genes alvo indica a plasticidade funcional destas moléculas e reforça a importância do entendimento das vias de silenciamento gênico, para futuros trabalhos envolvendo a manipulação e o desenvolvimento da cultura do café.

Palavras-chave: Café. Microtranscriptoma. Arf 8. RACE. qPCR

ABSTRACT

The phenomenon of interference RNA mediated by microRNAs is an important pathway for the stability and regulation of messenger RNA translation. Many of the regulated routes are fundamental to the organism development and to the genome integrity maintenance. High homology in sequence and structure are interesting features of the genes that encode these molecules. Those particularities allow that, from sequences of known silencing RNAs in other species, homologous microRNAs be described in phylogenetically related species. Using the information generated by the coffee sequencing, it was possible, for the first time, to identify microRNAs (miRNAs) involved in the transcriptome modeling of the species *Coffea arabica*, *Coffea canephora*, *Coffea congensis* and *Coffea eugenioides*. In this study, the sequences, structures and possible functions of microRNAs in coffee are presented. Besides the structural characterization, we also show the qRT-PCR analysis that correlate the expression of MIR genes and their target genes. The various miRNA families described for coffee are correlated to different biological processes that include plant growth patterns, metabolic processes, hormone responses, protein degradation, responses to stress and cell signaling routes. The great variety of target genes indicates the functional plasticity of these molecules and reinforces the importance gene silencing pathways understanding, for future studies involving the manipulation and development of coffee cropping.

Keywords: Coffee. Microtranscriptome. Arf 8. RACE. qPCR

LISTA DE ILUSTRAÇÕES

Figura 1.1 Via de biogênese dos miRNAs.....	18
Figura 1.2 Árvore filogenética dos genes WRKY1 em espécies do gênero <i>Coffea</i>	25
Figura 1.3 Modelo para representar variações interspecíficas da acumulação de miRNAs.....	29
Figura 2.1 Fluxograma apresentando a metodologia utilizada na discriminação dos possíveis miRNAs do café.....	45
Figura 3.1 Resumo do processo de amplificação de miRNAs em <i>stem-loop</i> RT- PCR.....	69
Figura 3.2 Eletroforese em gel de agarose 3% do produto de PCR de diferentes famílias MIR.....	75
Figura 3.3 Representação dos ciclos <i>threshold</i> dos genes miR172 e GAPDH nas espécies <i>Coffea canephora</i> e <i>Coffea eugenioides</i>	77
Figura 3.4 Perfil de expressão relativa de miR172 em folhas jovens de <i>Coffea</i> <i>arabica</i> , <i>Coffea canephora</i> , <i>Coffea congensis</i> e <i>Coffea</i> <i>eugenioides</i>	78
Figura 3.5 Perfil de expressão relativa de miR167 em folhas jovens de <i>Coffea</i> <i>arabica</i> , <i>Coffea canephora</i> e <i>Coffea congensis</i>	82
Figura 3.6 Perfil de expressão relativa de miR171 em folhas jovens de <i>Coffea</i> <i>arabica</i> , <i>Coffea canephora</i> e <i>Coffea congensis</i>	83
Figura 3.7 Perfil de expressão relativa de miR172 em folhas jovens de <i>Coffea</i> <i>arabica</i> , <i>Coffea canephora</i> e <i>Coffea congensis</i>	84
Figura 3.8 Perfil de expressão relativa de miR390 em folhas jovens de <i>Coffea</i> <i>arabica</i> , <i>Coffea canephora</i> e <i>Coffea congensis</i>	85
Figura 3.9 Sítio do transcrito do gene ARF8 complementar ao microRNA miR 167.....	88
Figura 3.10 Perfil de expressão relativa de miR167 e ARF8 em folhas jovens de <i>Coffea arabica</i> e <i>Coffea canephora</i>	89

Figura 3.11 Possível sequência do Fator Responsivo a Auxina 8 de café.....90

Figura 3.12 Alinhamento entre trecho das sequências ARF8 de *Coffea arabica* (A) e *Coffea canephora* (B) inseridas em cinco clones bacterianos.....91

LISTA DE TABELAS

Tabela 2.1 Lista de miRNAs homólogos identificados no café.....	49
Tabela 2.2 Possíveis genes regulados pelos miRNAs descritos.....	51
Tabela 3.1 <i>Primers</i> utilizados no experimento qRT-PCR.....	72
Tabela 3.2 <i>Primers</i> utilizados nos experimentos 5' RACE e qRT-PCR.....	74

SUMÁRIO

CAPÍTULO 1 Os miRNAs e o café.....	10
1 INTRODUÇÃO.....	10
2 REFERENCIAL TEÓRICO.....	11
2.1 Projeto Genoma Brasileiro Café.....	11
2.2 Silenciamento gênico pós-transcricional.....	14
2.3 O alotetraploide <i>Coffea arabica</i> e seus possíveis progenitores <i>Coffea canephora</i> , <i>Coffea congensis</i> e <i>Coffea eugenioides</i>	19
2.4 miRNAs em poliploides, padrões de expressão aditivo e não-aditivo.....	26
3 OBJETIVO.....	32
REFERÊNCIAS.....	33
CAPÍTULO 2 Caracterização <i>in silico</i> dos possíveis microRNAs contidos no banco de ESTs do café.....	41
1 INTRODUÇÃO.....	41
2 MATERIAIS E MÉTODOS.....	42
3 RESULTADOS E DISCUSSÃO.....	46
4 CONCLUSÕES.....	57
REFERÊNCIAS.....	58
CAPÍTULO 3 Análise da expressão de microRNAs em <i>Coffea arabica</i>, <i>Coffea canephora</i>, <i>Coffea congensis</i> e <i>Coffea eugenioides</i>.....	64
1 INTRODUÇÃO.....	64
2 MATERIAIS E MÉTODOS.....	65
2.1 Coleta do material vegetal.....	65
2.2 Extração de RNA.....	66
2.3 Tratamento com DNase.....	67
2.4 Síntese de cDNA.....	68
2.5 Expressão dos genes MIR por <i>stem-loop</i> RT-PCR.....	70
2.6 Seleção dos genes alvos.....	73
2.7 Análise de 5' RACE.....	73
3 RESULTADOS E DISCUSSÃO.....	74
3.1 <i>stem-loop</i> RT-PCR.....	74
3.2 5' RACE.....	86
4 CONCLUSÕES.....	92
5 PERSPECTIVAS.....	93
REFERÊNCIAS.....	95
ANEXO A.....	99

CAPÍTULO 1 Os miRNAs e o café

1 INTRODUÇÃO

A capacidade de formar estruturas que promovem a criação de sítios catalíticos e a interação com outros componentes celulares têm revelado a importância do RNA nos processos de controle da expressão gênica. O fenômeno de RNA de interferência é uma via importante para a estabilidade e regulação da tradução de RNAs mensageiros, sua atuação influencia de forma crucial a dinâmica celular. Nas plantas a maquinaria de regulação pós-transcricional é mediada, principalmente, por moléculas conhecidas como microRNAs (miRNAs).

Os miRNAs modelam a expressão gênica por meio do pareamento de bases específicas a um mRNA alvo, induzindo a clivagem ou o impedimento de tradução deste mRNA. Muitas das vias reguladas são fundamentais para o desenvolvimento do organismo e para a manutenção da integridade do genoma. Processos fundamentais do desenvolvimento vegetal como a determinação da identidade de células meristemáticas, da polaridade dos órgãos foliares e da transição entre os estágios vegetativo e reprodutivo, são dependentes dos mecanismos de silenciamento promovidos pelos miRNAs. Além disso, os pequenos RNAs são os principais moduladores das respostas epigenéticas à hibridação de genomas influenciando os principais processos celulares envolvidos com a funcionabilidade e estabilidade de núcleos que contenham dois ou mais genomas diferentes.

A criação do banco de ESTs (*Expressed Sequence Tags*) CAFEST, trouxe a oportunidade de conhecer mais profundamente diversos processos biológicos. Dentre as várias opções, a caracterização das vias de silenciamento gênico pós-transcricional (PTGS), mostra-se como uma temática bastante

atrativa uma vez que o mecanismo atua nas mais diversas rotas metabólicas e permanece, até o momento, desconhecido para o café.

Neste trabalho são apresentadas as sequências, estruturas e possíveis funções dos microRNAs no cafeeiro. Além disso, com os ensaios de RT-PCR e de *Rapid Amplification of cDNA Ends* (RACE) foi possível conhecer o padrão de expressão apresentado por estes RNAs, verificar a presença e a relação destes em espécies com diferentes níveis de ploidia e confirmar a regulação exercida por miR167 sobre os genes auxina responsivos ARF8.

2 REFERENCIAL TEÓRICO

2.1 Projeto Genoma Brasileiro Café

O mercado brasileiro de café é o mais desenvolvido dentre os produtos agrícolas nacionais, apresentando o maior volume de transações entre os contratos negociados na Bolsa de Mercadorias e Futuros (BM&FBOVESPA, 2010), o que demonstra a influência econômica do produto no agronegócio do país. A disposição espacial diversificada da produção brasileira amplia a capacidade em atender às mais diferentes exigências mundiais no que se refere a paladares e preços e coloca o país em posição de líder mundial, não só em quantidade, mas também em qualidade.

Esforços convencionais puderam cumprir com a maioria das necessidades da agricultura para a produção de café. Estratégias de melhoramento clássico garantiram sucesso notável na seleção e produção de cultivares resistentes a diferentes doenças, tais como a ferrugem alaranjada das folhas e a antracnose dos frutos (SILVA et al., 2006). Além da obtenção de variedades com alterações na arquitetura da planta e no período de maturação dos frutos (SERA, 2001). Entretanto, quando apenas técnicas “tradicionalistas” são utilizadas, a identificação de problemas e a introdução de um novo traço em uma

variedade de café, passam a ser processos longos que podem perfazer de 20 a 35 anos antes da liberação de um novo cultivar (RIBAS; PEREIRA; VIEIRA, 2006). As técnicas de engenharia genética trazem a possibilidade de obtenção destes genótipos melhorados, em tempo reduzido e de forma mais pontual.

Por analogia ao termo genoma, que representa o conjunto completo das sequências de DNA de um organismo, criou-se também os termos transcriptoma e proteoma que representam, respectivamente, o conjunto completo dos transcritos (RNAs) e das proteínas de um organismo. Em projetos transcriptoma são utilizadas informações oriundas do mRNA, assim apenas regiões gênicas são sequenciadas. O processo é realizado por meio da técnica de Reação em Cadeia da Polimerase com transcrição reversa (RT-PCR). Com esta técnica constrói-se uma biblioteca de cDNA que representa todo o conjunto de mRNAs existentes na célula ou tecido no momento da extração.

Uma vez que os sequenciadores atuais apenas conseguem ler sequências curtas de DNA, é preciso fragmentar o cDNA clonado em pequenos pedaços antes do sequenciamento. Posteriormente, a informação original será reconstruída com o auxílio de ferramentas de bioinformática que montarão os fragmentos sequenciados, chamados *reads*. Estes pequenos trechos sequenciados funcionam como etiquetas identificando genes expressos, por esse motivo eles são chamados de ESTs (*Expressed Sequence Tags*). Assim o transcriptoma é o reflexo direto da expressão gênica e avalia o conjunto de transcritos de um organismo em determinadas situações celulares (MIR, 2005).

As coleções de ESTs são muito importantes para os estudos de genômica funcional por permitirem a construção de perfis quali-quantitativos da expressão gênica em diferentes tecidos, tipos celulares e/ou estágios de desenvolvimento (WU et al., 2011). A metodologia tem sido largamente utilizada para identificação, anotação e mapeamento de novos genes. Em plantas o sequenciamento de clones de cDNA foi feito primeiramente para as espécies

modelos *Arabidopsis thaliana* (NEWMAN et al., 1994) e *Oriza sativa* (YAMAMOTO; SASAKI, 1997). Devido à similaridade em nível genômico entre o arroz e outras culturas importantes, o sequenciamento deste genoma teve um impacto significativo na biotecnologia vegetal e na bioinformática. Posteriormente, o transcriptoma de uma série de outras espécies, como *Vitis vinifera* (ABLETT et al., 2000), *Saccharum spp.* (FIGUEIREDO et al., 2001), *Solanum lycopersicum* (VAN DER HOEVEN et al., 2002), *Solanum tuberosum* (RONNING et al., 2003) e *Triticum vulgare* (OGIHARA et al., 2003) também foram sequenciadas.

O Projeto Genoma Brasileiro do Café foi desenvolvido em parceria entre CBP & D Café (Consórcio Brasileiro de Pesquisa e Desenvolvimento do Café), FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), Embrapa Recursos Genéticos e Biotecnologia além de outras instituições Federais, com o objetivo de elucidar detalhadamente a função dos milhares de genes ligados aos diferentes processos metabólicos e de desenvolvimento das plantas de café. Para isso, 214.964 clones foram aleatoriamente escolhidos de 37 bibliotecas de cDNA de *C. arabica*, *C. canephora* e *C. racemosa* e então sequenciados. O sequenciamento resultou na anotação de 153.739 sequências. Os ESTs foram agrupados em 17.982 *contigs* e em 32.155 *singlets* (VIEIRA et al., 2006). A criação desta plataforma de dados permitiu a identificação de cerca de 33.000 diferentes unigenes e viabilizou ainda, a construção de mapas genéticos, que potencialmente aumentarão a densidade de marcadores moleculares e auxiliarão no melhoramento de novas variedades. De forma geral, o CAFEST fortaleceu a capacidade de pesquisa do Consórcio nas áreas de genômica (BARRETO et al., 2011), bioinformática (LIMA et al., 2011) e biologia estrutural (VIEIRA et al., 2006).

2.2 Silenciamento gênico pós-transcricional

O dogma central da biologia molecular que postula a função do RNA como um carreador intermediário de informação gênica tem sido questionado há muito tempo. Já na década de 60, Britten e Davidson (1969) propuseram que RNAs “ativadores”, transcritos de regiões genômicas redundantes, poderiam formar dímeros com sequências regulatórias no DNA e ativar uma grande quantidade de genes. Este primeiro esboço de regulação gênica desempenhada por RNAs alertou para que estes provavelmente desempenhassem funções muito mais relevantes do que simples mensageiros ou fábricas de proteínas e que as células dispõem de diversos mecanismos para controlar seu metabolismo.

Napoli, Lemieux e Jorgensen (1990) e Van der Krol (1990) obtiveram resultados inusitados, e difíceis de explicar em seus trabalhos de transformação de petúnias. Os pesquisadores tentavam produzir petúnias com uma pigmentação violeta mais intensa, inserindo por cisgenia uma cópia extra do gene que produz a enzima responsável pelo pigmento floral. Contrariamente ao esperado, em vez de roxas, as flores exibiam pétalas brancas. A cópia extra do gene para pigmentação não apenas não resultava em intensificação da cor como suprimia, inclusive, a ação do gene endógeno. O fenômeno evidenciava a existência de algum mecanismo que impedia tanto a expressão dos cisgenes como dos genes originalmente presentes na planta e por isso ficou conhecido como cossupressão (BARBOSA; LIN, 2004).

Três anos depois do trabalho com as petúnias, Lee e colaboradores publicaram o primeiro trabalho demonstrando o silenciamento de um gene pela ação de um RNA endógeno (LEE; FEINBAUM; AMBROS, 1993). No estudo, realizado em *Caenorhabditis elegans*, foi constatado que uma mutação em lin-4 causava um aumento na tradução do RNA de outro gene específico, lin-14. Ao

clonar *lin-4* e administrá-lo a mutantes *Caenorhabditis*, a expressão de *lin-14*, característica dos fenótipos selvagens podia ser recuperada. Nesse trabalho os autores verificaram que *lin-4* não codificava proteínas e que o sequenciamento dos dois genes revelava um sítio complementar a *lin-4* na região 3' UTR de *lin-14*. Este seria o motivo pelo qual o transcrito de *lin-4* era capaz de se ligar a região 3' UTR de *lin-14* e regular negativamente sua tradução.

Somente alguns anos mais tarde, com experimentos realizados por Fire et al. (1998), foi que as bases moleculares do silenciamento foram elucidadas e o conceito de PTGS (*Post-transcriptional gene silencing*) ou RNAi (RNA interferente) foi realmente consolidado. A inoculação de RNAs de duplo filamento (dsRNA - *Double Strand RNA*) contendo as fitas sense e antisense do RNA alvo, causou a degradação de RNAs mensageiros específicos em *Caenorhabditis elegans*. As moléculas de RNA inseridas correspondiam ao gene *unc-22*, importante para a função muscular, sua ação resultou em abolição da proteína codificada por este gene e consequentes deformações nos nematóides (FIRE et al., 1998). Os pesquisadores denominaram o fenômeno de “RNA de interferência” para transmitir a ideia do papel chave que tem o RNA de duplo filamento na deflagração e censura de genes correspondentes. O que houve, tanto para as petúnias como para o nematódeo, foi a ativação do aparato de RNAi, sistema que atua como um sensor, reprimindo a expressão gênica pela interceptação específica de RNAs mensageiros (LAU; BARTEL, 2001). Em outras palavras, intermediários de RNA dupla-fita silenciaram a expressão dos genes que apresentavam sequência especificamente similar às suas, acarretando a censura de mRNAs homólogos e assim o aparecimento de flores menos violáceas e de vermes deformados.

Os sistemas de identificação e degradação de moléculas de RNA, provavelmente iniciaram-se como parte integrante do sistema de defesa natural das plantas a replicações virais e mobilização de elementos transponíveis

(DENLI; HANNON, 2003). Hoje se sabe que este mecanismo, eficiente de regulação gênica inata, atua fortemente no controle de genes envolvidos no desenvolvimento do organismo, na manutenção da integridade de genomas, na metilação de DNA e formação de heterocromatina.

Os siRNAs e os miRNAs compartilham funções essencialmente parecidas, no entanto, eles diferem quanto à origem (BARTEL, 2004). Enquanto os siRNAs provêm dos mesmos tipos de genes ou regiões genômicas que se tornam mais tarde silenciadas. Os miRNAs são oriundos de genes chamados MIR, cuja única missão é produzir esses diminutos RNAs regulatórios que impedirão a tradução do mRNA de outros genes (HE; HANNON, 2004).

Os genes dos miRNAs são evolutivamente conservados e podem estar localizados em íntrons ou éxons de genes codificadores de proteínas ou em regiões intergênicas (RODRIGUEZ et al., 2004). Seus transcritos são RNAs não codantes de sequências curtas (20 - 22 nucleotídeos) que regulam a expressão gênica negativamente por meio do pareamento de bases específicas a um mRNA alvo. Este pareamento resulta na clivagem do mRNA ou na repressão de sua tradução (BARTEL, 2004). Sua biossíntese inicia-se pela transcrição dos genes MIR pela enzima *RNA-polimerase III*. O transcrito primário (pri-miRNA) é uma longa molécula de RNA de fita simples contendo um trecho rico em adeninas, responsável pela conformação estrutural característica destes RNAs, conhecida como estrutura *hairpin*, (a estrutura lembra a forma de um grampo de cabelo). O pri-miRNA sofre a ação de diversas enzimas, principalmente da RNase III DICER-LIKE1 (DCL1), e produz então um precursor intermediário chamado pre-miRNA (VAUCHERET, 2006). Os pre-miRNAs são transportados para o citoplasma pela enzima HASTY (HST) onde serão processados por DCL1 e darão origem a um produto fita dupla (20 – 22 nucleotídeos), contendo em um de seus braços o miRNA maduro e no outro uma sequência adicional conhecida como miRNA* (SCHWARZ et al., 2003; LI et al., 2005).

As ribonucleases da família gênica *DICER-LIKE* (DCLs) clivam seus produtos em posições ligeiramente irregulares, de forma que cada miRNA formado possua 2 nucleotídeos projetados de cada lado (BERNSTEIN et al., 2001). A proteína associada à *DICER* conhecida como R2D2 percebe a energia termodinâmica livre na extremidade 5' das fitas ligando-se a mais estável (TOMARI; MATRANGA, 2004). O dímero é então desenrolado e um dos filamentos, o miRNA maduro, é transportado ao complexo proteico RISC (*RNA Induced Silence Complex*) (HAMMOND et al., 2000) enquanto a sequência que compreende o miRNA* é degradada. A seleção da fita ativa depende da estrutura do dúplice, dos nucleotídeos na ponta 5' e da assimetria termodinâmica (KAWAMATA; TOMARI, 2010).

O RISC contém enzimas conhecidas como SLICER (AGO1), estas enzimas são “guiadas” pelo miRNA maduro e promovem a repressão da tradução de genes alvos ou atuam clivando mRNAs complementares em dois filamentos que, tornam-se então, incapazes de traduzir uma proteína funcional (VOINNET, 2009) (Figura 1.1).

Tanto a clivagem quanto o mecanismo de repressão de tradução, requerem pares de base complementares entre o miRNA e o mRNA controlado. Em plantas a interferência ocorre, geralmente, pela clivagem do mRNA (BENTWICH, 2005). Vários experimentos utilizando a técnica modificada de RACE 5' (*rapid amplification of cloned ends*) mostraram que essa clivagem geralmente acontece no centro da complementaridade entre o miRNA e o mRNA, usualmente no resíduo que parecia com o décimo ou 11º nucleotídeo do miRNA (YANG; XUE; AN, 2007).

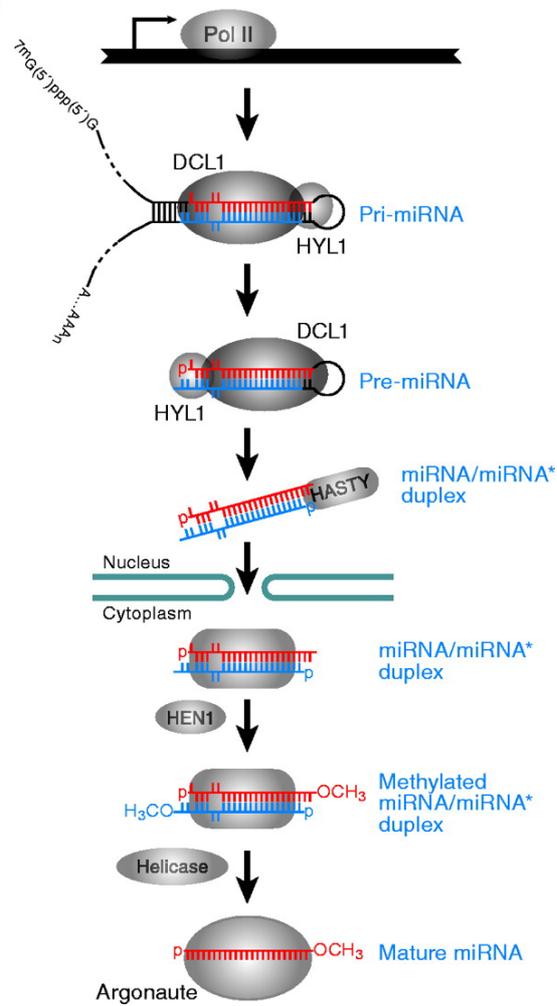


Figura 1.1 Modelo da biogênese de miRNAs e mecanismo de ação em plantas. Adaptado de Du & Zamore (2005).

2.3 O alotetraploide *Coffea arabica* e seus possíveis progenitores *Coffea canephora*, *Coffea congensis* e *Coffea eugenioides*.

A família Rubiaceae é considerada uma das maiores famílias de angiospermas distribuída em todo o mundo, inclui aproximadamente 637 gêneros e cerca de 10.700 espécies. A família foi primeiramente descrita em 1789, por Antoine Laurent de Jussieu, e reúne um considerável número de espécies com importância econômica, principalmente como ornamental, medicinal ou alimentícia (MENDOZA et al., 2004).

O gênero *Coffea* conta atualmente com 103 espécies identificadas (DAVIS et al., 2006). Neste gênero estão incluídas as espécies de maior interesse econômico da família Rubiaceae, *Coffea arabica* e *Coffea canephora*. Com exceção de *C. arabica* que é alotetraploide ($2n = 4x = 44$) e apresenta taxa de autofecundação em torno de 90%, todas as outras espécies conhecidas do gênero são diploides ($2n = 2x = 22$ cromossomos) e munidas de sistemas de autoincompatibilidade reprodutiva, multiplicando-se exclusivamente por fecundação cruzada (FAZUOLI et al., 1999).

O café arábica é produzido em mais de sessenta países e classificado entre as cinco exportações agrícolas mais valorizadas de países em desenvolvimento (FOOD AND AGRICULTURE ORGANIZATION, 2011). O comércio internacional dessa *commodity* tem importância expressiva em todos os setores da economia. Seu cultivo, processamento, comercialização, transporte e mercado proporcionam milhões de empregos em todo o mundo. Apesar de sua importância comercial, algumas dúvidas ainda permanecem sobre o processo de especiação deste alotetraploide, sobre quais espécies diploides se combinaram para originar esse genoma.

O conhecimento das relações e da diversidade genética das espécies que compõe este grupo não é apenas importante para a resolução de questões

taxonômicas, este conhecimento é indispensável também para determinação dos melhores cruzamentos e combinações híbridas, para o planejamento de processos de transgenia que favoreçam a integração e expressão correta da cópia gênica a ser inserida e para o bom uso de bancos de germoplasma. Espécies que não são comercialmente cultivadas como *Coffea congensis*, *Coffea racemosa* e *Coffea liberia* dispõem de grande diversidade genética, possuem características agrônomicas desejáveis e apresentam baixa suscetibilidade a várias pragas e doenças (PETITOT; LECOULS; FERNANDEZ, 2008). O conhecimento acurado sobre os subgenomas que compõem o material genético tetraploide de *Coffea arabica*, pode otimizar o desenho de projetos de transferência gênica e garantir a possibilidade do uso eficiente do vasto acervo de características adaptativas encontrado nas espécies diploides.

Os primeiros trabalhos citológicos e com certa base molecular sobre as relações genéticas entre o cafeeiro e seus possíveis genitores diploides, foram publicados no fim da década de 90. No ano de 1998, Cros e colaboradores trabalharam com DNA plastidial de 23 espécies de café. Os autores traçaram relações filogenéticas entre essas espécies baseados nos padrões de variação de regiões não codificadoras destes DNAs. Analisando a sequência do espaçador intergênico trnL-trnF do cloroplasto, os pesquisadores encontraram alta similaridade entre o DNA plastidial de *Coffea arabica* e de *Coffea eugenioides* (CROS et al., 1998). Na maioria das angiospermas, incluindo espécies cultivadas, como milho, arroz, trigo, algodão e também no café, os cloroplastos não apresentam segregação mendeliana, seus genes são de herança maternal e por isso não são transmitidos pelo pólen (CORRIVEAU; COLEMAN, 1998). Assim, as informações geradas por Cros e seu grupo mostraram que *Coffea arabica* seria um anfídiploide formado por duas espécies diploides, e que no processo de hibridização *Coffea eugenioides* teria participado como progenitor

materno, e um membro do grupo caneporoide, que poderia ser *Coffea congensis* ou *Coffea Canephora*, seria o progenitor paterno.

No mesmo ano de 1998, Raina, Mukai e Yamamoto utilizaram hibridização genômica *in situ* (GISH) e hibridização *in situ* por fluorescência (FISH) para estudar a organização e evolução do genoma de *Coffea arabica* e propor, pela primeira vez, a identidade das duas espécies ancestrais do alotetraploide. Nesse trabalho os autores utilizaram DNA marcado das espécies *Coffea eugenioides*, *Coffea congensis*, *Coffea canephora* e *Coffea liberica*. Estes DNAs serviram como sonda para discriminar cromossomos no café arábica. Como resultado eles obtiveram que as sondas construídas com o DNA de *Coffea eugenioides* marcaram fortemente 22 cromossomos de *Coffea arabica* enquanto que os outros 22 cromossomos foram preferencialmente marcados por sondas de *Coffea congensis*, mas também apresentaram hibridização com sondas construídas a partir do material genético de *Coffea canephora*. Os autores discutem que, apesar de haver hibridização para ambas as espécies caneporoides, a intensidade dos sinais de hibridização entre os cromossomos de *Coffea arabica* e as sondas de *Coffea canephora* era fraca e fragmentada quando comparada aos sinais emitidos pelo pareamento com as sondas de *Coffea congensis*. O padrão de similaridades observado entre *Coffea arabica* e as espécies diploides foi confirmado pela técnica de FISH, usando os genes ribossômicos pTa 71 e pTa 794. Baseados nestes resultados o grupo japonês afirmou que não restavam mais dúvidas sobre a ancestralidade de *Coffea arabica*, o trabalho apresentava então *Coffea congensis* como o progenitor masculino do híbrido (RAINA; MUKAI; YAMAMOTO, 1998).

No entanto, um ano mais tarde o grupo de pesquisa coordenado por Philippe Lashermes publicou um trabalho que combinava as técnicas Polimorfismo dos Comprimentos dos Fragmentos de Restrição (RFLPs) e Hibridização Genômica *in situ* (GISH), para investigar a relação entre as

mesmas espécies de café. Empregando 23 sondas nucleares de cópias únicas para comparar os padrões de RFLP das possíveis espécies progenitoras com os padrões de *Coffea arabica*, o grupo de Lashermes habilitou-se a afirmar que *Coffea arabica* é claramente um híbrido formado pela união dos genomas de *Coffea eugenioides* e *Coffea canephora*, ou de ecótipos relacionados a estas espécies diploides (LASHERMES et al., 1999). Os resultados obtidos com a técnica de GISH apontam para a mesma hipótese, utilizando o DNA genômico total de cada uma das espécies diploides em estudo como sondas, os autores puderam discriminar dois conjuntos de 22 cromossomos cada. Um grupo mostrando uma afinidade mais marcante para a sonda construída com o material genômico de *Coffea eugenioides* e o segundo grupo, diferente do resultado encontrado por Raina et al. no ano anterior, preferencialmente se hibridizando com a sonda de *Coffea canephora*.

Somente alguns anos mais tarde, em 2008, outro trabalho abordando a questão dos subgenomas que constituem o genoma de *Coffea arabica* foi publicado. O trabalho de um grupo francês descreveu naquele ano a primeira identificação e avaliação funcional de cópias de gene homólogos em *Coffea arabica*. O gene estudado era o CaWRKY1, gene que codifica um fator de transcrição da superfamília WRKY, associada a respostas de defesa das plantas a ataques de patógenos (PETITOT; LECOULS; FERNANDEZ, 2008). O material vegetal utilizado nesse trabalho foram folhas de variedades de cafeeiros resistentes e susceptíveis ao nematoide *Meloidogyne exigua*. Ao clonar CaWRKY1 no material vegetal desafiado pelo patógeno, os autores encontraram duas cópias distintas do gene, CaWRKY1a e CaWRKY1b. As espécies *C. canephora*, *C. eugenioides*, *C. congensis*, *C. liberica*, *C. racemosa* e *C. humilis* foram escolhidas para serem avaliadas como possíveis doadoras destes genes para o genoma híbrido de *Coffea arabica*. Essa escolha foi norteadada por análises filogenéticas conduzidas com diferentes marcadores moleculares

(LASHERMES et al., 1997; CROS et al., 1998). A similaridade na sequência de nucleotídeos de WRKY1 foi alta entre as espécies diploides, numa faixa de 94,4 a 99,3%. No entanto, uma similaridade ainda maior (99,9%) foi observada entre CaWRKY1a e o gene WRKY1 presente em *Coffea canephora* e entre CaWRKY1b e o gene WRKY1 de *Coffea eugenioides* (PETITOT; LECOULS; FERNANDEZ, 2008).

Análises de parcimônia realizadas com as sequências de WRKY1 produziram cladogramas construídos com alto índice de consistência (0,95) indicando baixo nível de homoplasia. A árvore filogenética obtida nos permite ver a relação estreita existente entre as sequências do gene WRKY1 presentes em *C. arabica*, *C. canephora* e *C. eugenioides* (Figura 1.2) (PETITOT; LECOULS; FERNANDEZ, 2008). Os resultados publicados nesse trabalho fortemente sugerem que os genes CaWRKY1a e CaWRKY1b são sequências homólogas resultantes do evento de alopoliploidização que deu origem à espécie *Coffea arabica*.

Como mencionado, além da análise filogenética, o trabalho destes autores forneceu também os primeiros dados moleculares sobre a expressão específica de cada subgenoma no café alopoliploide. Para verificar se CaWRKY1a e CaWRKY1b eram funcionais, a expressão gênica de plantas de café arábica inoculadas com *Meloidogyne exigua*, foi monitorada via qRT-PCR. Os ensaios em tempo real, mostraram que estes genes homólogos são concomitantemente ativados e exibem os mesmos padrões de expressão mediante a presença de patógenos. Estes resultados sugerem que CaWRKY1a e CaWRKY1b foram funcionalmente retidos no genoma do cafeeiro e que nenhuma divergência funcional ocorreu entre os genes duplicados no genoma de *Coffea arabica*.

Mais recentemente, em 2011, Marraccini e colaboradores analisaram a expressão do gene RBCS1 no alotetraploide *Coffea arabica*. Assim como o

encontrado por Petitot, Lecouls e Fernandez (2008), as análises *in silico* dos ESTs do cafeeiro apontavam para a existência de dois genes homólogos RBCS1 em *Coffea arabica*. Uma forma alélica, denominada pelos autores de CaCc, se alinhando perfeitamente a um unigene presente no genoma de *Coffea canephora* e outra exibindo maior similaridade com um unigene de *Coffea eugenioides* (alelo CaCe).

Os alelos caracterizados nessa pesquisa exibiram estrutura muito similar, compartilhando mais de 90% de identidade. A distinção entre os dois genes se dá por poucos polimorfismos de base única (SNPs) e alguns eventos de inserção e deleção (indels) ao longo da região 3' não traduzida (3' UTR) e das sequências intrônicas.

Primers específicos desenhados para cada um dos homólogos possibilitaram a avaliação da expressão das duas formas alélicas em plantas de café de diferentes espécies. Como esperado, os pesquisadores detectaram expressão de RBCS1-CaCe em *Coffea eugenioides* e *Coffea arabica* mas não em *Coffea canephora*. Por outro lado, grande quantidade de transcritos do alelo RBCS1-CaCc foi detectada em folhas das cultivares Tupi e IAPAR59, variedades de *Coffea arabica* recentemente introgrididas com *Coffea canephora* (MARRACCINI et al., 2011).

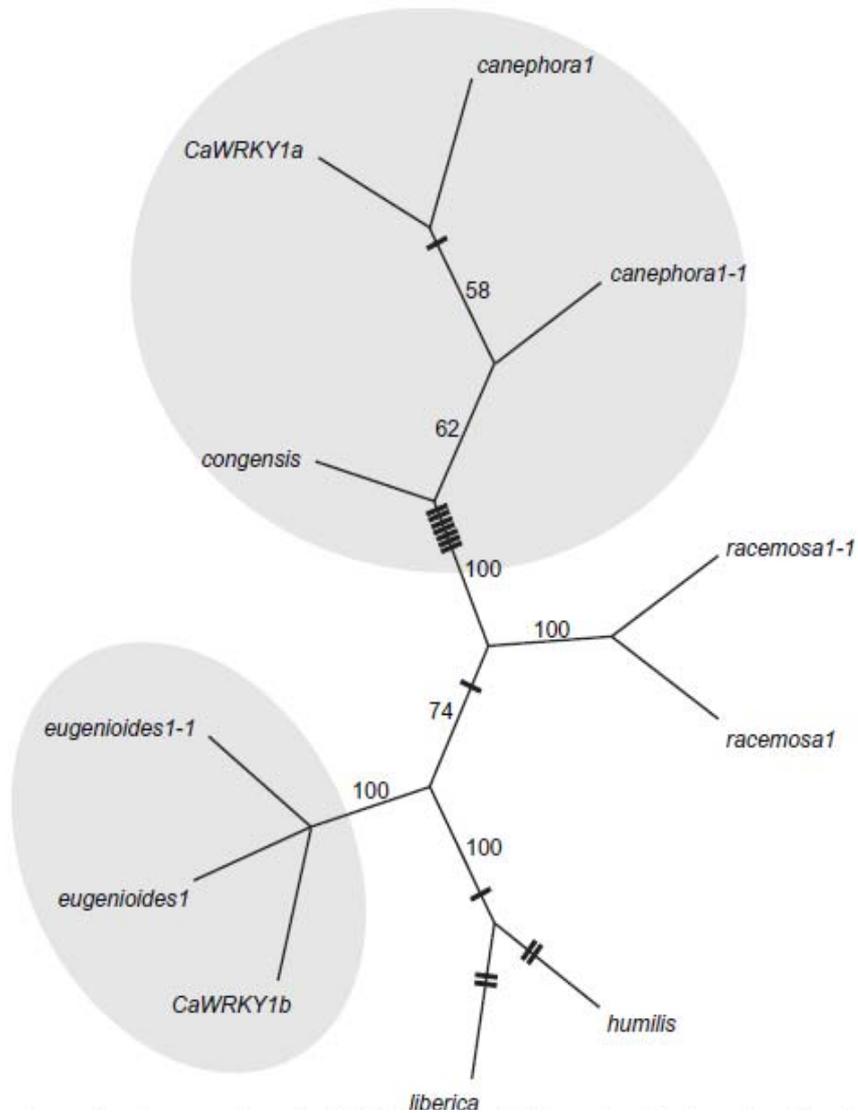


Figura 1.2 Árvore filogenética (máxima parcimônia) dos genes WRKY1 em espécies do gênero Coffea. Os números acima dos ramos são valores de bootstrap (100 repetições). Barras escuras (mutações inserção/deleção) foram atribuídas manualmente para cada ramo, mas não foram levadas em conta para a análise filogenética. Adaptado de Petitot, Lecouls e Fernandez (2008).

2.4- miRNAs em poliploides, padrões de expressão aditiva e não-aditiva.

A maioria das angiospermas sofre um ou mais eventos de poliploidização em sua linhagem, entre 2 a 4% de todos os eventos de especiação parece ocorrer através deste tipo de mecanismo (VAN DE PEER; MEYER, 2009). A poliploidia pode oferecer um nível adicional de oportunidades para as espécies recém formadas através de processos como o tamponamento de mutações, a sub e neo funcionalização dos genes duplicados e fixação da heterozigose. No entanto, apesar das possíveis vantagens, eventos de poliploidização podem induzir "choques genômicos" devido à redundância gênica, a expressão desequilibrada de genes antagônicos, problemas no emparelhamento dos cromossomos homólogos e a falta de sintonia no processo de replicação do DNA entre os múltiplos genomas (KENAN-EICHLER et al., 2011).

Estudos em diversas espécies têm demonstrado que uma ampla gama de respostas genéticas e epigenéticas ocorrem após os processos de hibridação e poliploidização. Tais respostas incluem alterações na estrutura da cromatina, rearranjos ou metilação de citosina, silenciamento gênico, ativação de transposons e até mesmo deleções de DNA (CHEN, 2007; DOYLE et al., 2008). Os pequenos RNAs têm sido associados com todos esses eventos atuando ativamente no controle de alterações epigenéticas que ocorrem em resposta à poliploidização (MATZKE; BIRCHLER, 2005; KENAN-EICHLER et al., 2011).

As alterações epigenéticas são potencialmente reversíveis e proporcionam um meio eficaz e flexível para os núcleos poliploides responderem aos choques genômicos (CHEN, 2007). Essas mudanças podem reprogramar a expressão gênica e os padrões de desenvolvimento nos novos alopoliploides e o silenciamento ou ativação reversível dos genes pode ser um

dos primeiros passos em direção a um estado fixo e geneticamente irreversível (WANG et al., 2006). As principais alterações epigenéticas são conduzidas por pequenos RNAs, que influenciam os principais processos celulares envolvidos com a funcionabilidade e estabilidade do genoma (KENAN-EICHLER et al., 2011).

Mudanças no padrão de expressão de alguns genes regulatórios, tais como fatores de transcrição e microRNAs podem influenciar muitos genes envolvidos em diferentes rotas metabólicas (CHEN; NI, 2006). Um único miRNA pode, por exemplo, regular centenas de genes envolvidos na transição de um estágio de desenvolvimento para outro (FARH et al., 2005; CHEN, 2007).

Embora os mecanismos de regulação via miRNAs sejam altamente conservados, variações na expressão dos genes que codificam estes RNAs e seus mensageiros alvos são relativamente comuns, inclusive entre espécies estreitamente relacionadas (HA et al., 2008). Estas variações podem refletir divergências nas redes regulatórias pós-transcricionais em poliploides. Em genomas híbridos, os loci de miRNAs herdados podem divergir em sequência e modo de expressão, mesmo quando se tratando de miRNAs de uma mesma família. As modificações na cromatina podem também influenciar os elementos cis e trans presentes nas sequências promotoras desses genes (AXTELL; BARTEL, 2005). Como os promotores são os responsáveis por determinar o local e o momento em que o processo de transcrição deve ocorrer, estes loci estão sujeitos a ganhar novos padrões de expressão em determinados órgãos/tecidos ou estágios de desenvolvimento. Alguns genes, inclusive, podem ser perdidos no processo de especiação (ASON et al., 2006).

Assim, a combinação de dois loci de miRNAs, tenham eles padrões de expressão diferenciais ou não, pode perturbar o balanço regulatório entre os RNAs catalíticos e seus genes alvo em células que contenham dois ou mais conjuntos de genomas, sejam elas híbridos interespecíficos ou aloploiploides

(HA et al., 2008). Como resultado, os níveis de acúmulo de miRNAs podem ser aditivos ou não-aditivos (CHEN; NI, 2006; WANG et al., 2006).

Em poliploides diz-se que a expressão de um gene é aditiva quando o nível de expressão desse gene é igual à soma dos dois loci parentais, (*MPV-mid-parent value*) (CHEN, 2007). Um exemplo simplificado desse padrão seria cada um dos parentais de um híbrido hipotético apresentando o valor de expressão para um dado gene. No híbrido formado, a combinação destes loci determinaria um valor de expressão igual a dois ($1 + 1 = 2$) (HA et al., 2009). De forma contrária, se a expressão de um gene é não-aditiva, o valor de expressão deste loci será diferente do MPV. No híbrido formado os loci combinados poderiam ser levados à ativação e apresentariam uma expressão ligeiramente maior que dois (> 2), ou então ser reprimidos e apresentar uma expressão menor que dois (< 2) (CHEN, 2007; HA et al., 2009). A Figura 1.3 traz um modelo para explicar regulação gênica aditiva e não aditiva.

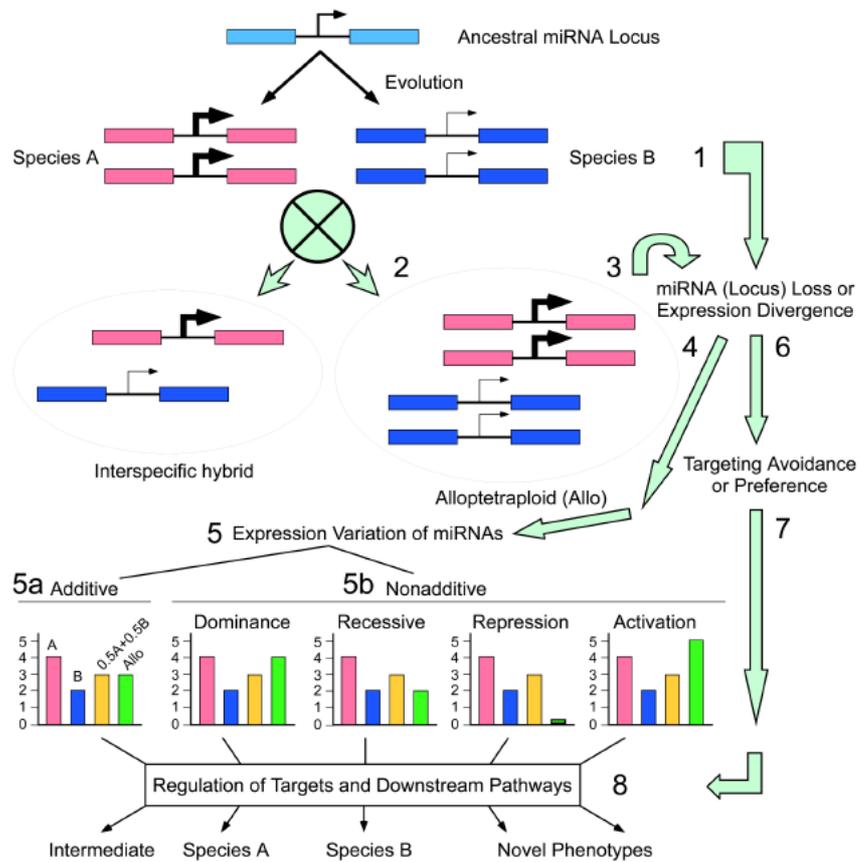


Figura 1.3 Modelo para representar variações interspecíficas da acumulação de miRNAs. Os loci de miRNAs são representados setas entre as sequências de codificantes (caixas rosa, azul e azul claro). O círculo com uma cruz representa cruzamento entre espécies, resultando em um híbrido interespecífico ou um alotetraploide. Setas verde claro junto aos números de 1 a 8 indicam possíveis alterações que podem ocorrer devido à divergência entre os loci de miRNAs presentes nas espécies A e B. Estas alterações podem ser mutações, perda do conjunto gênico de uma ou outra espécie e/ou alterações nos padrões de expressão. Setas grossas e finas indicam níveis altos e baixos de acumulação de miRNAs, respectivamente. Se ambos os loci estão presentes, eles podem se expressar de maneira diferente devido a mecanismos genéticos e epigenéticos. Os possíveis níveis de expressão são mostrados como, aditivo (5a) e não-aditivo (5b) (dominância, repressão, recessividade, ativação ou sobredominância). Os níveis de expressão dos miRNAs são mostrados em vermelho (para espécie A), azul (para espécie B), amarelo (para valores médios dos pais, MPV, que equivale a uma mistura igual de RNAs

dos parentais) e verde (para híbridos interespecífico ou alotetraploides). Adaptado de Ha et al (2008).

Em *Arabidopsis*, os alelos *FLOWERING LOCUS C* (FLC) mostram efeitos aditivos no período de floração (MICHAELS; AMASINO, 1999). Em híbridos alotetraploides que combinam os genomas de *A. thaliana* e *A. arenosa*, as expressões dos loci FLC somam-se e originam um fenótipo de floração tardia. Alguns trabalhos demonstram que até 90% do transcriptoma de alotetraploides resultantes da hibridação entre espécies relacionadas de *Arabidopsis* é expresso aditivamente (WANG et al., 2006). Embora para miRNAs o padrão de expressão mais comum seja o não-aditivo, em 2011, Kenan-Eichler e colaboradores encontraram que em trigo o percentual de pequenos RNAs correspondentes a miRNAs aumentou com o nível de ploidia. Os autores relatam que no alopoliploide, apesar de haver uma super-representação de miR168, a quantidade da maioria das famílias MIR foi semelhante aos valores médios dos pais.

Por outro lado, análises do transcriptoma de diferentes alopoliploides vegetais, incluindo também híbridos de *Arabidopsis*, têm indicado que muitos genes são expressos não-aditivamente (WANG et al., 2006; HEGARTY et al., 2006; SWANSON-WAGNE et al., 2006). Este valor médio de expressão diferente dos pais condiz com os padrões de dominância ou recessividade fenotípica geralmente observada entre os alopoliploides. Um dado interessante é que mais de 50% dos genes que dependem da regulação por miRNAs estão entre os loci que não são expressos de forma aditiva (revisado em HA et al., 2008). Isto nos mostra o quão importante é o controle da biogênese destes RNAs para que a expressão diferencial de alvos não aconteça de forma inadequada ou prejudicial trazendo, por exemplo, incompatibilidades, redução do *fitness* ou do vigor ao genoma poliploide.

Existem várias possibilidades para explicar o padrão de expressão não aditivo, dentre elas estão os efeitos de dominância, recessividade, repressão, ativação, sobredominância e a própria expressão não aditiva de genes relacionados ao processamento dos miRNAs (CHEN, 2007). Em *Arabidopsis*, por exemplo, genes como DCL1 e *AGO4* (este mais especificamente importante para biogênese e funcionamento correto de siRNAs) aparecem nas folhas e flores de híbridos alotetraploides de maneira não aditiva (HA et al., 2009).

Como dito anteriormente, muitos genes regulados por miRNAs codificam fatores de transcrição ou proteínas importantes para o desenvolvimento vegetal. Exemplos disso são miR164 que em *Arabidopsis* está associado com a morte celular programada (KIM et al., 2009); e miR172 que regula a mudança da fase do crescimento vegetativo para a fase de crescimento reprodutivo em plantas de algodão (ZHANG et al., 2007). A expressão não aditiva de miRNAs como estes pode levar a fenótipos mais semelhantes a uma ou outra das espécies parentais ou como é mais comum, produzir características totalmente inéditas na espécie híbrida (CHEN, 2007; NIWA; SLACK, 2007; HA et al., 2008).

Coletivamente estas informações sugerem que os aloploiploides podem obter vantagens ao controlar a expressão dos miRNAs e siRNAs herdados. Através de expressão diferencial, estes podem tornar-se verdadeiros amortecedores contra os choques genéticos resultantes da fusão de genomas (HA et al., 2009). Pelo silenciamento ou superativação de seus genes alvo os pequenos RNAs influem decisivamente no estabelecimento e desenvolvimento de novas combinações genômicas, sejam elas aloploiploides ou não.

3 OBJETIVO

Considerando a importância dos microRNAs na regulação da expressão gênica, é proposto o presente trabalho cujo objetivo geral é caracterizar, pela primeira vez, microRNAs atuantes nas vias de modelagem do transcriptoma das espécies *Coffea arabica* e *Coffea canephora*. Os objetivos específicos compreendem:

- Identificar no banco de ESTs “CAFEST”, *singlets* ou *contigs* que caracterizem prováveis candidatos a genes MIR.
- Apresentar as composições nucleotídicas e estruturas secundárias dos miRNAs do café, bem como seus possíveis genes regulados.
- Analisar, por PCR quantitativa, o padrão de expressão dos microRNAs na espécie tetraploide *Coffea arabica* e nas espécies diploides *Coffea canephora*, *Coffea congensis* e *Coffea eugenioides*.
- Usando a técnica *Rapid Amplification of cDNA Ends* (RACE) confirmar a regulação exercida por miR167 sobre os genes auxinas responsivos ARF8.

REFERÊNCIAS

- ABLETT, E. et al. Analysis of grape ESTs: global gene expression patterns in leaf and berry. **Plant Science**, Limerick, v. 159, n. 1, p. 87-95, Oct. 2000.
- ASON, B. et al. Differences in vertebrate microRNA expression. **Proceedings of the National Academy of Sciences United States of American**, New York, v. 103, n. 39, p. 14385–14389, Sept. 2006.
- AXTELL, M. J.; BARTEL, D. P. Antiquity of microRNAs and their targets in land plants. **The Plant Cell**, Rockville, v. 17, n. 6, p. 1658–1673, June 2005.
- BARBOSA, A. S.; LIN, C. J. Silenciamento de genes com RNA interferência: um novo instrumento para investigação da fisiologia e fisiopatologia do córtex adrenal. **Arquivos Brasileiros de Endocrinologia & Metabologia**, São Paulo, v. 48, n. 5, p. 612-619, Aug. 2004.
- 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**, Athens, v. 30, p. 29-35, 2011.
- BARTEL, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. **Cell**, Cambridge, v. 116, n. 2, p. 281–297, Jan. 2004.
- BENTWICH, I. Prediction and validation of microRNAs and their targets. **FEBS Letters**, Amsterdam, v. 579:5904–5910, Sept. 2005.
- BERNSTEIN, E. et al. Role for a bidentate ribonuclease in the initiation step of RNA interference. **Nature**, London, v. 409, n. 6818, p. 363-366, Jan. 2001.
- BM&FBOVESPA: a nova bolsa. Bolsa de Mercadorias & Futuros. São Paulo: BM&F, 2010. Disponível em: <<http://www.bmf.com.br>>. Acesso em: 06 dez. 2010.

BRITTEN, R. J.; DAVIDSON, E. H. Gene regulation for higher cells: a theory. **Science**, Washington, v. 165, n. 3891, p. 349-357, July 1969.

CHEN, Z. J. Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. **Annual Review of Plant Biology**, Palo Alto, v. 58, p. 377–406, June 2007.

CHEN, Z. J.; NI, Z. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. **BioEssays**, Cambridge, v. 28, n. 3, p. 240–252, Mar. 2006.

CORRIVEAU, J. L.; COLEMAN, A. W. Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. **American Journal of Botany**, Columbus, v. 75, n. 11, p. 1443–1458, Nov. 1988.

CROS, J. et al. Phylogenetic analysis of chloroplast DNA variation in *Coffea* L. **Molecular Phylogenetics and Evolution**, Orlando, v. 9, n. 1, p. 109–117, Feb. 1998.

DAVIS, A. et al. An annotated checklist of the genus *Coffea* L. (Rubiaceae). **Botanical Journal of the Linnean Society**, London, v. 152, n. 4, p. 465–512, Dec. 2006.

DENLI, A. M.; HANNON, G. J. RNAi: an ever-growing puzzle. **Trends Biochemistry Science**, Amsterdam, v. 28, n. 5, p. 196–201, May 2003.

DOYLE, J. J. et al. Evolutionary genetics of genome merger and doubling in plants. **Annual Review of Genetics**, Palo Alto, v. 42, p. 443-461, 2008.

DU, T.; ZAMORE, P. D. microPrimer: the biogenesis and function of microRNA. **Development**, Cambridge, v. 132, n. 21, p. 4645-4652, Nov. 2005.

DUGAS, V. D.; BARTEL, B. MicroRNA regulation of gene expression in plants. **Current Opinion in Plant Biology**, Amsterdam, v. 7, n. 5, p. 512–520, Oct. 2004.

FARH, K. K. et al. The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. **Science**, Washington, v. 310, n. 5755, p. 1817–1821, Dec. 2005.

FAZUOLI, L. C. et al. **Cultivares de café**. Campinas: Instituto Agrônomo, 1999.

FIGUEIREDO, R. C. et al. Dissecting the sugarcane expressed sequence tag (SUCEST) database: unraveling flower-specific genes. **Genetic and Molecular Biology**, Ribeirão Preto, v. 24, n. 1-4, p. 77-84, Jan./Dec. 2001.

FIRE, A. et al. Potent and specific genetic interference by double stranded RNA in *Caenorhabditis elegans*. **Nature**, London, v. 391, n. 6669, p. 806-811, Feb. 1998.

FOOD AND AGRICULTURE ORGANIZATION. Organização das Nações Unidas para Agricultura e Alimentação. Brasília: FAO, 2011. Disponível em: <<http://www.fao.org>>. Acesso em 12 dez. 2011.

HA, M. et al. Interspecies regulation of MicroRNAs and their targets. **Biochimica et Biophysica Acta**, Amsterdam, v. 1779, n. 11, p. 735–742, Nov. 2008.

HA, M. et al. Small RNAs serve as a genetic buffer against genomic shock in *Arabidopsis* interspecific hybrids and allopolyploids. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 106, n. 42, p. 17835-17830, Oct. 2009.

HAMMOND, S. M. et al. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. **Nature**, London, v. 404, n. 6776, p. 293–296, Mar. 2000.

HE, L.; HANNON, G. MicroRNAs: small RNAs with a big role in gene regulation. **Nature Reviews Genetics**, London, v. 5, n. 5, p. 522-531, May 2004.

HEGARTY, M. J. et al. Transcriptome shock after interspecific hybridization in senescio is ameliorated by genome duplication. **Current Biology**, London, v. 16, n. 16, p. 1652–1659, Aug. 2006.

KAWAMATA, T.; TOMARI Y. Making RISC. **Trends Biochemistry Science**, Amsterdam, v. 35, n. 7, p. 68-76, July 2010.

KENAN-EICHLER, M. et al. Wheat Hybridization and Polyploidization Results in Deregulation of small RNAs. **Genetics**, Austin, v. 188, n. 2, p. 263-272, June 2011.

KIM, J. H. et al. Trifurcate feed-forward regulation of age-dependent cell death involving miR164 in *Arabidopsis*. **Science**, Washington, v. 323, n. 5917, p. 1053–1057, Feb. 2009.

LASHERMES, P. et al. Molecular characterization and origin of the *Coffea arabica* L. genome. **Molecular and General Genetics**, New York, 261, n. 2, p. 259-266, 1999.

LASHERMES, P. et al. Phylogenetic relationships of coffee-tree species (*Coffea* L.) as inferred from ITS sequences of nuclear ribosomal DNA. **Theoretical and Applied Genetics**, Berlin, v. 94, p. 947–955, 1997.

LAU, N. C.; BARTEL, D. P. Censores internos. **Scientific American**, New York, v. 16, n. 2, p. 35-41, Sept. 2001.

LEE, R. C.; FEINBAUM, R. L.; AMBROS, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. **Cell**, Cambridge, v. 75, n. 5, p. 843–854, Dec. 1993.

LI, J. et al. Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in *Arabidopsis*. **Current Biology**, London, v. 15, n. 16, p. 1501–1507, Aug. 2005.

LIMA, A. A. et al. In silico characterization of putative members of the coffee (*Coffea arabica*) ethylene signaling pathway. **Genetic and Molecular Biology**, Ribeirão Preto, v. 10, n. 2, p. 1277-1289, June 2011.

MARRACCINI, P. et al. RBCS1 expression in coffee: *Coffea* orthologs, *Coffea Arabica* homeologs, and expression variability between genotypes and under drought stress. **BMC Plant Biology**, New York, v. 11, p. 85-108, May 2011.

MATZKE, M. A.; BIRCHLER, J. A. RNAi-mediated pathways in the nucleus. **Nature Reviews Genetics**, London, v. 6, n. 1, p. 24-35, Jan. 2005.

MENDOZA, H. et al. **Rubiaceae de Colômbia**: guia ilustrado de gêneros. Bogotá: Instituto de Investigación de Recursos Biológicos Alexandre Von Humboldt, 2004.

MICHAELS, S. D.; AMASINO, R. M. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. **The Plant Cell**, Rockville, v. 11, n. 5, p. 949–56, May 1999.

MIR, L. **Gênômica**. São Paulo: Atheneu, 2005.

NAPOLI, C.; LEMIEUX, C.; JORGENSEN, R. Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. **The Plant Cell**, Rockville, v. 2, n. 4, p. 279-289, Apr. 1990.

NEWMAN, T. et al. Genes Calore: a summary of methods for accessing results from large-scale partial sequencing of anonymous arabidopsis cDNA clone. **Plant Physiology**, Minneapolis, v. 106, n. 4, p. 1241-1255, Dec. 1994.

NIWA, R.; SLACK, F. J. The evolution of animal microRNA function. **Current Opinion in Genetics & Development**, London, v. 17, n. 2, p. 145–150, Apr. 2007.

OGIHARA, Y. et al. Correlated clustering and virtual display of gene expression patterns in the wheat life cycle by large-scale statistical analyses of expressed sequence tags. **Plant Journal**, Oxford, v. 33, n. 6, 1001–1011, Mar. 2003.

PETITOT, A-S.; LECOULS, A-C.; FERNANDEZ, D. Sub-genomic origin and regulation patterns of a duplicated WRKY gene in the allotetraploid species *Coffea arabica*. **Tree Genetics & Genomes**, New York, v. 4, p. 379–390, 2008.

RAINA, S.N.; MUKAI, Y.; YAMAMOTO, M. In situ hybridization identifies the diploid progenitor species of *Coffea arabica* (Rubiaceae). **Theoretical and Applied Genetics**, Berlin, v. 97, p. 1204-1209, 1998.

RIBAS, A. F.; PEREIRA, L. F. P.; VIEIRA, L. G. E. Genetic transformation of coffee. **Brazilian Journal Plant of Physiology**, Londrina, v. 8, n. 1, p. 83-94, Jan./Mar. 2006.

RODRIGUEZ, A. et al. Identification of mammalian microRNA host genes and transcription units. **Genome Research**, Cold Spring Harbor, v. 14, n. 10a, p. 1902-1910, Oct. 2004.

RONNING, C. et al. Comparative analyses of potato expressed sequence tag libraries. **Plant Physiology**, Minneapolis, v. 131, v. 2, p. 419–442, Feb. 2003.

SCHWARZ, D. S. et al. Asymmetry in the assembly of the RNAi enzyme complex. **Cell**, Cambridge, v. 115, n. 2, p. 199–208, Oct. 2003.

SERA, T. Coffee genetic breeding at IAPAR. **Crop Breeding and Applied Biotechnology**, Londrina, v. 1, n. 2, p. 179-199, 2001.

SILVA, M. C. et al. Coffee resistance to the main diseases: leaf rust and coffee berry disease. **Brazilian Journal of Plant Physiology**, Piracicaba, v. 18, n. 1, p. 119-147, Jan./Mar. 2006.

SWANSON-WAGNER, R. A. All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents. **Proceedings of the National Academy of Sciences, New York**, v. 103, n. 18, p. 6805–6810, May 2006.

TOMARI, Y.; MATRANGA, C. A protein sensor for siRNA asymmetry. **Science**, Washington, v. 306, n. 5700, p. 1377-1380, Nov. 2004.

VAN DE PEER, Y., S. M.; MEYER, A. The evolutionary significance of ancient genome duplications. **Nature Reviews Genetics**, London, v. 10, p. 725-732, Oct. 2009.

VAN DER HOEVEN, R. et al. Deductions about the number, organization, and evolution of genes in the tomato genome based on analysis of a large expressed sequence tag collection and selective genomic sequencing. **The Plant Cell**, Rockville, v. 14, n. 7, p. 1441–1456, July 2002.

VAN DER KROL, A. R. et al. Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. **The Plant Cell**, Rockville, v. 2, n. 4, p. 291-299, Apr. 1990.

VAUCHERET, H. Post-transcriptional small RNA pathways in plants: mechanisms and regulations. **Genes & Development**, Cold Spring Harbor, v. 20, n. 7, p. 759–771, Apr. 2006.

VETTORE, A. L. et al. The libraries that made Sucest. **Genetic and Molecular Biology**, Ribeirão Preto, v. 24, n. 1, p. 1-7, Jan./Dec. 2001.

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

VOINNET, O. Origin, biogenesis, and activity of plant MicroRNAs. **Cell**, Cambridge, v. 136, n. 4, p. 669–687, Feb. 2009.

WANG, J. et al. Genomewide nonadditive gene regulation in Arabidopsis allotetraploids. **Genetics**, Austin, v. 172, n. 1, p. 507–517, Sept. 2006.

WU, X-M. et al. Identification and characterization of microRNAs from citrus expressed sequence tags. **Tree Genetics & Genomes**, New York, v. 7, p. 117–133, 2011.

YAMAMOTO, K.; SASAKI, T. Large-scale EST sequencing in rice. **Plant Molecular Biology**, Dordrecht, v. 35, n. 1-2, p. 135–144, 1997.

YANG, T.; XUE, L.; AN, L. Functional diversity of miRNA in plants. **Plant Science**, Limerick, v. 172, n. 3, p. 423-432, Mar. 2007.

ZAMORE, P. D. Ancient pathways programmed by small RNAs. **Science**, Washington, v. 296, n. 5571, p. 1265-1269, May 2002.

ZHANG, B. et al. Identification of cotton microRNAs and their targets. **Gene**, Amsterdam, v. 397, n. 1-2, p. 26-37, Aug. 2007.

CAPÍTULO 2 Caracterização *in silico* dos possíveis microRNAs contidos no banco de ESTs do café.

1 INTRODUÇÃO

Após a descoberta do primeiro miRNA (lin-4) em *Caenorhabditis elegans* (LEE; FEINBAUM; AMBROS, 1993), milhares de miRNAs foram identificados em genomas eucarióticos (GRIFFITHS-JONES et al., 2008). Estima-se que entre 1 a 5% dos genes transcritos em plantas e animais contenham miRNAs, estes números tornam os miRNAs uma das classes mais abundantes e dinâmicas de reguladores genéticos (LIM et al., 2005).

Muitos miRNAs de plantas, assim como seus genes alvo foram identificados por técnicas de clonagem (EMERY et al., 2003; TANG et al., 2003). Entretanto, a evolução das técnicas de sequenciamento que tem oferecido uma riqueza de dados cada vez maior, somada à disponibilidade de eficientes ferramentas de bioinformática têm possibilitado uma amostragem muito mais abrangente dos pequenos RNAs. As abordagens *in silico* ampliam as chances de identificação de miRNAs não expressos ou daqueles que são especificamente expressos em determinados tecidos ou estágios de desenvolvimento (BONNET et al., 2004). Os recentes métodos computacionais fornecem uma maneira rápida, precisa, de baixo custo e, conseqüentemente, muito conveniente para recuperar seqüências precursoras de miRNA em bases públicas de dados (ZANCA et al., 2010).

O uso destas estratégias é bem sucedido tanto para caracterização de miRNAs inéditos quanto para a caracterização por similaridade daqueles conservados entre diferentes espécies. Especialmente em plantas, onde a complementaridade quase perfeita entre os miRNAs e seus mRNAs facilita a predição computacional, a bioinformática tem importância muito significativa para a identificação dos genes codificadores de RNA de silenciamento e dos

genes dependentes da regulação destes RNAs catalíticos (JONES-RHOADES; BARTEL, 2004). Centenas de miRNAs vegetais foram identificados pela semelhança com homólogos descritos nas espécies *Arabidopsis thaliana* e *Oriza sativa* (XIE; FRAZIER.; ZHANG, 2011; WU et al., 2011).

A plataforma de dados e registro de miRNAs, o miRBase, tem depositados em sua versão mais recente (*release* 18) mais de 15.000 loci de microRNAs e 17.000 sequências maduras distintas em cerca de 140 espécies (GRIFFITHS-JONES et al., 2008). Há sequências de miRNAs disponíveis para 29 plantas dicotiledôneas, 10 plantas monocotiledôneas, dentre outras como para a alga *Chlamydomonas reinhardtii*. Nesta versão, o miRBase apresenta um total de 581 miRNAs caracterizados para *Oriza sativa*, 291 em *Arabidopsis thaliana*, 234 em *Populus trichocarpa* e 172 em *Zea mays*, dentre outros. No entanto, não há nenhum depósito referente a espécies do gênero *Coffea*. Até o momento, nenhum trabalho científico abordando os miRNAs do café foi publicado. Neste trabalho descrevemos, pela primeira vez, a composição nucleotídica e a estrutura secundária dos possíveis miRNAs atuantes nas vias de regulação pós-transcricional do cafeeiro. Além disso, apresentamos também possíveis genes alvos controlados por estes pequenos RNAs.

2 MATERIAIS E MÉTODOS

A metodologia adotada para discriminar os possíveis miRNAs do cafeeiro, foi descrita por Zhang et al. (2006) e vem sendo amplamente empregada na caracterização de pequenos RNAs homólogos em espécies vegetais (ZHANG et al., 2006; XIE et al., 2007; QIU et al., 2007). A Figura 2.1 representa um esboço da metodologia. Um total de 1.549 miRNAs maduros, pertencentes a 20 espécies vegetais, sendo elas *Arabidopsis thaliana*, *Brassica*

napus, *Brassica oleracea*, *Brassica rapa*, *Carica papaya*, *Glycine Max*, *Gossypium herbecium*, *Gossypium hirsutum*, *Gossypium rammindii*, *Lotus japonicus*, *Medicago truncatula*, *O. sativa*, *Physcomitrella patens*, *Populus trichocarpa*, *Saccharum officinarum*, *Solanum lycopersicum*, *Sorghum bicolor*, *Triticum aestivum*, *Vigna unguiculata*, *Vitis vinifera* e *Zea mays*, foram utilizados na caracterização dos potenciais miRNAs contidos no banco de ESTs do café.

As sequências destes RNAs foram retiradas do miRBase (*Release 13.0*, Março de 2009)(GRIFFITHS-JONES et al., 2006). Para evitar a sobreposição entre os dados obtidos, sequências redundantes foram eliminadas. As 809 sequências que permaneceram foram definidas como base de dados referência (sequências *query*). Por meio da interface *on-line* Gene Project (www.lge.ibi.unicamp.br/cafe) buscou-se *reads* de *Coffea arabica* e *Coffea canephora* contendo *singlets* ou *contigs* que caracterizassem prováveis candidatos a genes MIR.

Para selecionar sequências similares entre os ESTs de café e os miRNAs do miRBase, submetemos as sequências *query* a uma busca BLAST (*Basic Local Alignment Search Tool*) (ALTSCHUL et al., 1997). O algoritmo deste programa procura alinhar apenas as regiões mais conservadas entre as sequências *query* analisadas e um banco de dados. Este método de comparação é utilizado para determinar o grau de similaridade entre duas ou mais sequências, ou a similaridade entre fragmentos destas sequências.

A qualidade de um alinhamento é determinada pela soma dos pontos obtidos por cada unidade pareada (*match*) menos as penalidades pela introdução de gaps e posições não pareadas (*mismatch*). Sua unidade de medida é o *e-value*, que representa a probabilidade do alinhamento ter ocorrido por acaso, não sendo as sequências similares. A variação do *software* utilizado para realizar a busca no banco ESTs foi o BLASTn 2.2.19 com parâmetros padrão.

Apenas sequências de café apresentando menos de 4 *mismatches*, incluindo inserções e/ou deleções de nucleotídeos, com relação aos miRNAs *query*, seguiram como candidatos. As 769 sequências apuradas, foram confrontadas com sequências de mRNAs armazenadas na plataforma de dados do NBCI, buscando-se eliminar com este passo as sequências que codificavam proteínas. Após este filtro o número de candidatos foi significadamente reduzido. Apenas 141 sequências atenderam as condições de alinhamento e se mostraram distintas de RNAs codantes.

Por fim, as estruturas secundárias dos ESTs contendo sequências MIR características, foram preditas através do *software online mfold 3,1* (ZUKER, 2003), o algoritmo está publicamente disponível em <http://www.bioinfo.rpi.edu/applications/mfold/rna/form1.cgi>. (THE RNA INSTITUTE, 2011). Todos os parâmetros do *software* foram utilizados no formato padrão. Foram analisados aspectos de energia livre (Δ Gkcal / mol), a constituição de nucleotídeos (A, G, C e U) e a localização da região correspondente ao miRNA maduro. Sabe-se que as sequências precursoras de miRNAs geralmente têm *minimal folding free energy index* (MFEI) significativamente maior do que outros RNAs. Uma sequência de RNA tem maior propensão a ser uma molécula de silenciamento quando apresentar valores de MFEI maiores do que 0,85 (ZHANG et al., 2006). Assim para evitar a designação errônea de outros RNAs como candidatos a microRNAs, MFEI foi um critério cuidadosamente analisado na predição das estruturas secundárias. O valor de MFEI pode ser calculado pela equação:

$$\text{MFEI} = \text{AMFE}/(\text{G+C})\%$$

Onde temos:

(AMFE) - energia mínima livre ajustável – representa a energia mínima disponível de 100 nucleotídeos.

(G + C)% - proporção de bases nitrogenadas do tipo guaninas e citosinas existentes na sequência analisada.

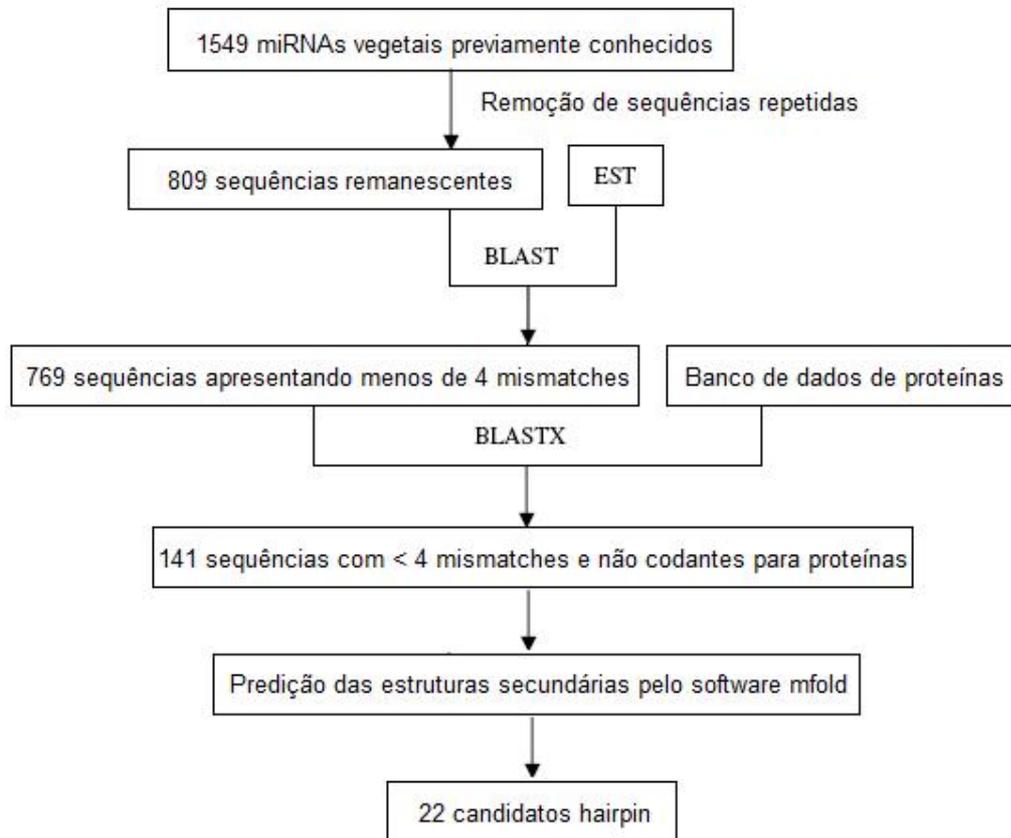


Figura 2.1 Fluxograma apresentando a metodologia utilizada na discriminação dos possíveis miRNAs do café.

Assim, contudo, apenas foram consideradas candidatas as sequências de RNA que se enquadraram nos seguintes critérios: (1) a estrutura do miRNA maduro a ser predita, não deve conter mais que quatro substituições nucleotídicas (*mismatches*), em comparação com a estrutura do miRNA maduro previamente conhecido; (2) a sequência de RNA deve se dobrar em uma

estrutura secundária apropriada (estrutura *hairpin* adequada); (3) conter a sequência do miRNA maduro em um dos braços do grampo e a sequência do miRNA* no outro; (4) não deve haver, além disso, mais de seis *mismatches* entre as sequências do miRNA maduro e sua sequência oposta na estrutura secundária; (5) laços ou rupturas nas sequências de miRNAs também não foram permitidas; (6) e por fim, as estruturas previstas devem apresentar MFEI superior a 0,85 (ZHANG et al., 2006). Depois de analisar esses seis critérios, o número total de candidatos a miRNAs foi drasticamente reduzido para vinte e dois.

3 RESULTADOS E DISCUSSÃO

A partir de 1.549 sequências iniciais, caracterizamos um total de 22 possíveis miRNAs para o cafeeiro. Estes RNAs pertencem a 8 diferentes famílias, e se distribuem de maneira desuniforme entre elas. Por exemplo, foram identificados 7 membros para a família miR167 enquanto apenas 3 foram descritos para a família miR172. Em estudos feitos nas espécies *Oriza sativa*, (ZHANG et al., 2006) *Brassica napus* (XIE et al., 2007) e *Vitis vinifera* (YAN- DU et al., 2008) temos que, diferente dos resultados obtidos neste trabalho, a família miR167 é composta apenas por 2 membros, enquanto miR172 contém um número bem mais expressivo de representantes. A variação no número de membros que constitui cada família, pode ser indicativo de sua função, sugerindo que eles podem oferecer funções únicas para a regulação da biogênese dos miRNAs e da expressão gênica em cada uma das espécies vegetais em que estão presentes (ZHANG et al., 2006) (Figura 2.2).

Os miRNAs apurados foram diversos também em outros aspectos, tais como na localização dos miRNAs maduros no contexto da estrutura *hairpin*, no

comprimento das sequências dos precursores e em sua constituição de bases nitrogenadas. Os miRNAs maduros podem ser encontrados em cada um dos dois braços da estrutura secundária (BONNET et al., 2004). Corroborando com resultados obtidos por outros trabalhos (ZHANG et al., 2006; QIU et al., 2007), a maioria dos miRNAs descritos estavam no sentido 5' das sequências do pré-miRNAs. Entre os 22 miRNAs identificados, 7 foram localizados na extremidade 3', enquanto 15 estavam localizados no braço que compreende a extremidade 5'.

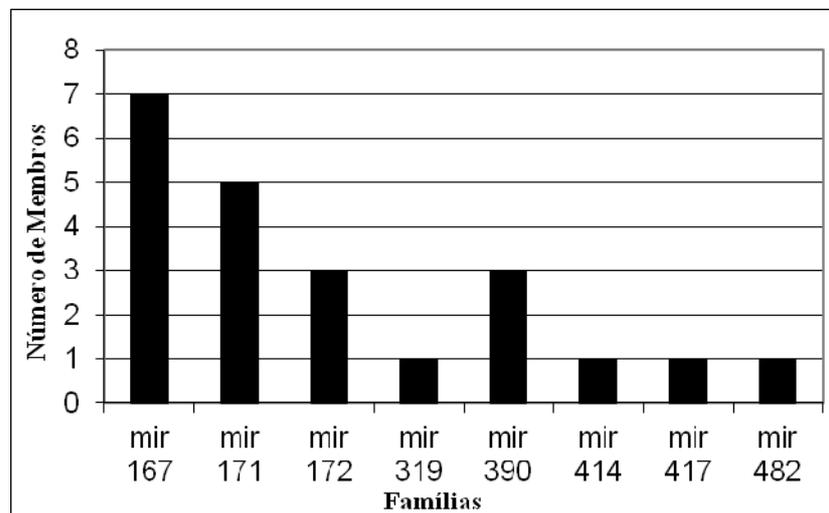


Figura 2.2 Número de membros por família de miRNA identificadas em *Coffea arabica* e *Coffea canephora*.

No café, o comprimento dos pré-miRNA variou de 37 a 125 nucleotídeos. Em relação ao tamanho dos precursores, os pré-miRNAs de plantas são mais diversificados em estrutura e tamanho, do que os descritos em animais. Enquanto pré-miRNA de mamíferos têm tipicamente de 70 a 80 nucleotídeos, precursores de miRNAs vegetais podem apresentar de 60 a 509 nucleotídeos, sendo que a maioria deles (73,5%) são constituídos por 81 a 160 nucleotídeos. A estrutura dos precursores é uma característica importante e

exerce grande influência nas funções das moléculas de silenciamento e, assim consequentemente, no controle da expressão gênica (ZHANG et al., 2006).

Além disso, as sequências dos candidatos a pré-miRNAs foram avaliadas quanto ao conteúdo de pares de bases nitrogenadas adenina (A) + uracila (U). Nossos resultados estão em acordo com os de trabalhos anteriores, (ZHANG et al., 2005; ZHANG et al., 2006; ZHANG et al., 2007), tendo este parâmetro variado de 48,00% a 75,00% (Tabela 1). Aspectos termodinâmicos são também critérios geralmente muito úteis para distinguir miRNAs de outros tipos de RNAs codantes ou não. Os pré-miRNAs caracterizados neste estudo apresentaram uma média alta de MFEI, aproximadamente 0.9334, valor significativamente mais elevado do aquele descrito para tRNAs (0,64), rRNAs (0,59), e mRNAs (0.62-0.66) (ZHANG et al., 2006), (Tabela 1).

Estudos recentes sugerem que cerca de 1% de genes preditos por meio de abordagens computacionais são genes MIR (LAI et al., 2003). No entanto, o número de miRNAs identificados em café, assim como o encontrado em tomate (YIN et al., 2008), por exemplo, foi relativamente pequeno. A aparente discrepância entre estas informações pode ser devida a duas razões não mutuamente-exclusivas. Primeiro, os precursores de miRNA são normalmente muito curtos (até 200 nucleotídeos) (ZHANG et al., 2006), e os *ESTs* geralmente contêm mais de 300 nucleótidos.

Assim, a maioria dos pré-miRNA não podem ser clonados para *ESTs*. Em segundo lugar, a maioria das moléculas primárias dos miRNAs são rapidamente processadas no núcleo, assim os precursores de miRNA têm uma menor probabilidade de serem clonados como *ESTs* (ZHANG et al., 2005).

Em contrapartida, como *ESTs* são provenientes da expressão real do gene em questão, esse tipo de análise fornece maiores evidências e confiança na descoberta de putativos miRNAs e os genes por eles regulados. Essas análises oferecem uma via rápida e de boa relação custo-benefício para a descoberta e

Tabela 1

isolamento de sequências de miRNAs conservados (ZHANG et al., 2006). Além disso, devido ao grande número de projetos de sequenciamento de genomas vegetais, mais clones de cDNA irão se tornar disponíveis em bancos de dados públicos, assim como novos RNAs de silenciamento poderão ser descobertos.

Neste contexto, as probabilidades de sucesso na mineração de dados entre os clones de cDNA aumentam, de forma que, muito provavelmente, mais miRNA serão caracterizados para as plantas em geral, inclusive para as espécies alvo deste estudo.

As famílias de miRNA descritas *in silico* para o cafeeiro, estão relacionadas a diferentes processos biológicos em espécies como *Arabidopsis thaliana*, *Oryza sativa*, *Brassica napus*, *Gossypium hirsutum*, dentre outras. Processos estes que, conforme descrito abaixo, incluem padrões de crescimento e o desenvolvimento vegetal, processos metabólicos, respostas hormonais, degradação de proteínas, formação do mRNA maduro, respostas a estresses e rotas de sinalização celular (Tabela 2).

Tabela 2

➤ **miR167**

Tanto para *Arabidopsis thaliana* quanto para *Oriza sativa*, miR167 apresenta-se associado aos mRNAs que codificam fatores de resposta a auxinas (*ARF transcription factors*) (BONNET et al., 2004).

As proteínas *ARF* podem se vincular a elementos *cis* das regiões promotoras de alguns genes auxina-responsivos (ULMASOV et al., 1997), podendo então ativar ou reprimir a expressão destes genes. A regulação dos níveis de transcritos é importante para o desenvolvimento de várias respostas do vegetal aos sinais de auxina que, por exemplo, incluem o alongamento celular, a divisão e diferenciação das raízes laterais e o desenvolvimento de estames e do gineceu nas flores (WU; TIAN; REED, 2006).

Plantas transgênicas de *Arabidopsis* (35S: MIR167b), apresentaram filamentos anormalmente curtos, anteras deformadas e grãos de pólen com dificuldade em germinar, sugerindo uma importante ligação entre a via regulatória da expressão gênica mediada por miRNAs e a rede de sinalização da auxina, para a promoção do desenvolvimento reprodutivo das plantas (RU et al., 2006).

Além disso, Coruzzi et al. (2009) observaram que superexpressão de miR167 em *Arabidopsis thaliana* induziu alterações fenotípicas em plantas jovens, aumentando as raízes laterais. Plantas com este fenótipo ampliaram suas possibilidades de exploração da superfície e assim demonstraram maior eficiência metabólica, maior absorção de nutrientes e maiores taxas de crescimento e produtividade. Por fim, alguns trabalhos descrevem a expressão diferencial da família em situações de estresse hídrico (LIU et al., 2008).

➤ **miR171**

Os alvos preditos para essa família são três proteínas SCARECROW-like, essa família de fatores de transcrição são influentes na padronização radial de raízes, sinalização pelo fito-hormônio giberelina e por estímulos luminosos (DI LAURENZIO et al., 1996). Os mRNA que codificam proteínas *cap methyltransferase* também são regulados pela família miR171 (ZHANG et al., 2007). Além disso, trabalhos realizados em *A. thaliana* descrevem a família diferencialmente se expressando em condições de estresse hídrico (LIU et al., 2008).

➤ **miR172**

Em *Arabidopsis thaliana* 20 dos 21 nucleotídeos dos transcritos de MIR172 são complementares a um trecho da sequência de RNA do gene homeótico floral APETALA2 (AP2). Essa complementaridade sugere regulação dos níveis de expressão deste gene através dos miRNAs membros desta família (PARK et al., 2002). A regulação pode resultar em defeitos nos órgãos florais e tempos alternativos de floração. Em algodão essa família de miRNAs também se encontra associada ao controle de proteínas *apetal-like*, que regulam a formação floral nas plantas de algodão e a mudança da fase do crescimento vegetativo para a fase de crescimento reprodutivo (ZHANG et al., 2007). Além disso, outras possíveis proteínas alvo de miR172 estão envolvidas com processos de desenvolvimento de fibras vegetais. Sendo elas, a *fosfoenolpiruvato carboxilase*, o receptor glutamato envolvido com as proteínas *dehydrogenase/reductase (SDR)* de cadeia curta, e com a biosíntese do fitorregulador ácido abscísico (ABA) (LEE; WOODWARD; CHEN 2007).

➤ **miR319**

Também conhecido como miR-JAW. A família está envolvida com a regulação de um grupo de fatores de transcrição planta-específicos chamado TCPs (PALATNIK et al., 2003). Os TCPs estão ligados a importantes eventos da fisiologia e do desenvolvimento foliar (CUBA'S et al., 1999), incluindo várias moléculas de sinalização relacionadas com a senescência das folhas e hormônios vegetais como ácido salicílico, etileno, citocininas e ácido jasmônico (VAN DER GRAAFF et al., 2006). Além disso, esse miRNA controla a expressão do gene *Lanceolate*, que está envolvido no controle da morfologia de folhas de tomate (ORI et al., 2007). Já para *Brassica napus*, o miRNA ath-miR319a, aparece regulando uma proteína homóloga às basic/helix-loop-helix (bHLH), sugerindo seu envolvimento com o controle da proliferação celular e/ou do estabelecimento de células meristemáticas (MASSARI; MURRE, 2000).

➤ **miR390**

Em *Arabidopsis* esta família MIR atua reconhecendo e clivando transcritos de loci denominados TAS (RNAs longos não codificantes que tem importante função ao longo do desenvolvimento vegetal). A ação regulatória de miR390 sobre o gene não codante TAS3, promove a correta formação de raízes laterais nesta espécie (YOON et al., 2009). Nagasaki et al. (2007) relatam que em conjunto com outros miRNAs esta família está envolvida com a manutenção da identidade e polaridade dos órgãos em arroz.

Já em plantas de algodão as proteínas possivelmente visadas por essa família incluem os fatores de transcrição *MYB* e *zinc finger*, as famílias de proteínas *glicosil transferease* e *action / hydrogen exchanger* e proteínas do

tipo miosina de cadeia pesada (ABDURAKHMONOV et al., 2008). Estas proteínas atuam nas rotas de embriogênese de óvulos e de desenvolvimento de fibras vegetais. Além disso, algumas destas têm se mostrado importantes para as sínteses da celulose, dos componentes do citoesqueleto e em transportes de prótons e cátions (LEE; WOODWARD; CHEN, 2007).

➤ **miR414**

Em *Gossypium hirsutum* encontra-se associado, principalmente, a famílias gênicas que regulam a transdução de sinais, respostas hormonais, processos metabólicos e de desenvolvimento (QIU et al., 2007). Devido à sua elevada afinidade de ligação às proteínas cromossômicas não-histônicas mais abundantes no núcleo de eucariotos superiores, as *high-mobility-group* (HMG) sugerem que a regulação exercida por este miRNA possa estar ligada à recombinação genética e à transcrição nos núcleos (QIU et al. 2007). Além disso, as proteínas WRKY, são também alvos da família miR414 (EULGEM, 2000). Tais proteínas parecem regular vários processos fisiológicos incluindo defesa a patógenos, a senescência e o desenvolvimento de tricomas. Outro alvo são as proteínas *basic leucine zipper proteins* (bZIP) responsáveis pela regulação de diversos processos fisiológicos e de respostas a estímulos bióticos e abióticos (SCHLOGL et al., 2004). Por fim, para *Brassica napus*, genes que controlam processos intracelulares incluindo sinais de transdução de cálcio, a rota de degradação de proteínas via ubiquitina e o transporte de proteínas também parecem ser regulados por essa família de pequenos RNAs (MOON; PARRY; ESTELLE, 2004), miR414 pode ligar-se perfeitamente ao mRNA codante para a ubiquitina sugerindo então sua capacidade de regular a expressão do nível desta proteína.

➤ **miR417**

Em *Arabidopsis* esta família apresenta expressão constante durante as fases de crescimento de todos os órgãos, incluindo caules, raízes, folhas e flores. Além disso, estudos recentes mostram que miR417 é transitoriamente mais expresso em resposta ao estresse osmótico (plantas submetidas a tratamentos com manitol), e que plantas transgênicas super expressando essas moléculas apresentam quadros de hipersensibilidade ao ABA (JUNG; KANG, 2007). Nesse mesmo trabalho, verificou-se ainda que miR417 tem um impacto negativo sobre a germinação de sementes e a taxa de sobrevivência de plantas de *Arabidopsis* em condição de estresse salino.

➤ **miR482**

Em *Brassica napus* esta família regula genes que codificam fatores transcrição, proteínas sinalizadoras, relacionados à resistência a doenças, a choques térmicos e à oxidação. Um dos alvos importantes de miR482 são os genes codificadores das heme-peroxidases, envolvidas com deposição de lignina e suberina (QUIROGA et al., 2000) na biosíntese de metabólitos secundários e no enrijecimento da parede celular através da imobilização de extensinas e das ligações entre os polissacarídeos da matriz celular (HIYAMA; LAM; STON, 1994).

4 CONCLUSÕES

Abordagens de bioinformática possibilitaram a caracterização de 22 putativos miRNAs desuniformemente distribuídos em oito famílias MIR. A identificação destes RNAs a partir da similaridade entre suas sequências e as sequências de seus homólogos em outras espécies, mostra o quão conservado é o processo de regulação pós-transcricional via microRNAs. Mesmo entre grupos filogeneticamente distantes.

Além dos pequenos RNAs foram descritos também alguns de seus potenciais genes alvo em *Coffea arabica* e *Coffea canephora*. A grande variedade de genes descritos indica a plasticidade funcional dessas moléculas e reforça a importância do entendimento das vias de silenciamento gênico em geral. Esta abordagem computacional poderá ser utilizada como uma orientação para o delineamento experimental. Possibilitando desenho de *primers*, norteando a escolha de organismos, processos, tecidos e até mesmo de rotas específicas para análises mais aprofundadas.

Os próximos passos deste estudo serão, portanto, analisar experimentalmente as sequências sugeridas por nossa análise *in silico*, determinando as funções moleculares e elucidando qualquer relação significativa entre os miRNAs e seus genes alvo.

REFERÊNCIAS

ABDURAKHMONOV, I. Y. et al. Small RNA regulation of ovule development in the cotton plant, *G. hirsutum* L. **BMC Plant Biology**, New York, v. 8, p. 93-101, Sept. 2008.

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

BARTEL, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. **Cell**, Cambridge, v. 116, n. 2, p. 281–297, Jan. 2004.

BONNET, E. et al. Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes. **Proceedings of the National Academy of Sciences of United States of American**, New York, v. 101, n. 31, p. 11511–11516, Aug. 2004.

CORUZZI, G. et al. **Patent application title:** methods of affecting plant growth with microRNA. USPC class 800278, nº 20090094711, 04 set. 2009.

CUBA'S, P. et al. The TCP domain: a motif found in proteins regulating plant growth and development. **Plant Journal**, Oxford, v. 18, p. 215–222, 1999.

DI LAURENZIO, L. et al. The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the *Arabidopsis* root. **Cell**, Cambridge, v. 86, n. 9, p. 423–433, Aug. 1996.

EMERY, J. F. et al. Radial patterning of *Arabidopsis* shoots by class III HD-ZIP and KANADI genes. **Current Biology**, London, v. 13, n. 20, p. 1768-1774, Oct. 2003.

EULGEM, T. The WRKY superfamily of plant transcription factors. **Trends Plant Science**, Oxford, v. 5, n. 5, p. 199–206, May 2000.

GRIFFITHS-JONES, S. et al. miRBase: tools for microRNA genomics. **Nucleic Acids Research**, Oxford, v. 36, p. 154-158, Jan. 2008.

GRIFFITHS-JONES, S. miRBase: integrating microRNA annotation and deep-sequencing data. **Nucleic Acids Research**, Oxford, v. 36, p. 154-158, Jan. 2008.

IYAMA, K.; LAM, T.; STON, B. A. Gene expression changes and early events in cotton fibre development. **Plant Physiology**, Minneapolis, v. 104, n. 2, p. 315-320, Feb. 1994.

JONES-RHOADES, M. W.; BARTEL, D. P. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. **Molecular Cell**, Cambridge, v. 14, n. 6, p. 787-799, June 2004.

JUNG, H. J.; KANG, H. Expression and functional analyses of microRNA417 in *Arabidopsis thaliana* under stress conditions. **Plant Physiology and Biochemistry**, Paris, v. 45, n. 10-11, p. 805–811, Oct./Nov. 2007.

LAI, E. et al. Computational identification of Drosophila microRNA genes. **Genome Biology**, Oxford, v. 4, n. 7, p. 1–20, Apr. 2003.

LEE, J. J.; WOODWARD, A. W.; CHEN, Z. J. Gene Expression Changes and Early Events in Cotton Fibre Development. **Annals of Botany**, London, v. 100, n. 7, p. 1391–1401, Dec. 2007.

LEE, R. C.; FEINBAUM, R. L.; AMBROS, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. **Cell**, Cambridge, v. 75, n. 5, p. 843–854, Dec. 1993.

LIM, L. P. et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. **Nature**, London, v. 433, n. 7027, p. 769-773, Feb. 2005.

LIU, H. H. et al. Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. **RNA**, New York, v. 14, n. 5, p. 836-843, May 2008.

MASSARI, M. E.; MURRE, C. Helix-loop-helix proteins: regulators of transcription in eukaryotic organisms. **Molecular Cell Biology**, London, v. 20, n. 2, p. 429-440, Jan. 2000.

MIRBASE. **miRBase**: the microRNA database. Manchester: [s.n.], 2009. Disponível em: < www.mirbase.org >. Acesso em: 15 mar. 2011.

MOON, J.; PARRY, G.; ESTELLE, M. The Ubiquitin-Proteasome Pathway and Plant Development. **The Plant Cell**, Rockville, v. 16, n. 12, p. 3181-3195, Dec. 2004.

NAGASAKI, H. et al. The small interfering RNA production pathway is required for shoot meristem initiation in rice. **Proceedings of the National Academy of Sciences of United States of American**, v. 104, n. 37, p. 14867-14871, Sept. 2007.

ORI, N. et al. Regulation of LANCEOLATE by miR319 is required for compound leaf development in tomato. **Nature Reviews Genetics**, London, v. 39, n. 10, p. 787-791, Oct. 2007.

PALATNIK, J. F. et al. Control of leaf morphogenesis by microRNAs. **Nature**, London, v. 425, n. 6955, p. 257-263, Sept. 2003.

PARK, W. J. et al. CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. **Current Biology**, London, v. 12, p. 1484-1495, 2002.

QIU, C. X. et al. Computational identification of microRNAs and their targets in *Gossypium hirsutum* expressed sequence tags. **Gene**, Amsterdam, v. 395, n. 1-2, p. 49–61, June 2007.

QUIROGA, M. et al. Gene expression changes and early events in cotton fibre development. **Plant Physiology**, Minneapolis, v. 122, n. 4, p. 1119–1127, Apr. 2000.

RU, P. et al. Plant fertility defects induced by the enhanced expression of microRNA167. **Cell Research**, New York, v. 16, p. 457-465, May 2006.

SCHLOGL, P. S. et al. **Plant Science**, Limerick, v. 167, n. 3, p. 583–595, Sept. 2004.

TANG, G. et al. A biochemical framework for RNA silencing in plants. **Genes and Development**, Cold Spring Harbor, v. 17, p. 49-63, Mar. 2003.

THE RNA INSTITUTE. College of Arts and Sciences. **Rna folding form**. New York: The RNA Institute, 2011. Disponível em: < <http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form> >. Acesso em: 14 jun. 2011.

ULMASOV, T. et al. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. **The Plant Cell**, Rockville, v. 9, p. 1963–1971, Nov. 1997.

VAN DER GRAAFF, E. et al. Transcription Analysis of Arabidopsis Membrane Transporters and Hormone Pathways during Developmental and Induced Leaf Senescence. **Plant Physiology**, Minneapolis, v. 141, n. 2, p. 776–792, Feb. 2006.

WU, M. F.; TIAN, Q.; REED, J. W. Arabidopsis microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. **Development**, Cambridge, v. 133, n. 21, p. 4211-4218, Nov. 2006.

WU, X-M. et al. Identification and characterization of microRNAs from citrus expressed sequence tags. **Tree Genetics & Genomes**, New York, v. 7, p. 117–133, 2011.

XIE, F. L. et al. Computational identification of novel microRNAs and targets in *Brassica napus*. **FEBS Letters**, Amsterdam, v. 581, n. 7, p. 1464–1474, 2007.

XIE, F.; FRAZIER, T. P.; ZHANG, B. Identification, characterization and expression analysis of MicroRNAs and their targets in the potato (*Solanum tuberosum*). **Gene**, Amsterdam, v. 473, n. 1, p. 8–22, Oct. 2011.

YAN-DU, L. et al. Identification and Characterization of MicroRNAs and their targets in grapevine (*Vitis vinifera*). **Agricultural Sciences in China**, Amsterdam, v. 7, v. 8, p. 929–943, Aug. 2008.

YIN, Z. et al. Identification of conserved microRNAs and their target genes in tomato (*Lycopersicon esculentum*). **Gene**, Amsterdam, v. 414, n. 7, p. 60–66, Sept. 2008.

YOON, E. K. et al. Auxin regulation of the microRNA390-dependent transacting small interfering RNA pathway in Arabidopsis lateral root development. **Nucleic Acids Research**, Oxford, v. 38, p. 1382–1391, Dec. 2009.

ZANCA, A. S. et al. Identification and expression analysis of microRNAs and targets in the biofuel crop sugarcane. **BMC Plant Biology**, New York, v. 10, p. 260–269, Nov. 2010.

ZHANG, B. et al. Conservation and divergence of plant microRNA genes. **The Plant Journal**, Texas, v. 46, p. 243–259, 2006.

ZHANG, B. et al. Identification of cotton microRNAs and their targets. **Gene**, Amsterdam, v. 397, n. 1–2, p. 26–37, Aug. 2007.

ZHANG, B. H. et al. Evidence that miRNAs are different from other. **Cellular and Molecular Life Sciences**, Basel, v. 63, n. 2, p. 246–254, 2006a.

ZHANG, B. H. et al. Identification and characterization of new plant microRNAs using EST analysis. **Cell Research**, New York, v. 15, n. 5, p. 336-360, 2005.

ZUKER, M. Mfold web server for nucleic acid folding and hybridization prediction. **Nucleic Acids Research**, Oxford, v. 31, n. 13, p. 3406–3415, July 2003.

CAPÍTULO 3 Análise da expressão de miRNAs, via qPCR, em *Coffea arabica*, *Coffea canephora*, *Coffea congensis* e *Coffea eugenioides*.

1 INTRODUÇÃO

Embora muitas espécies do gênero *Coffea* carreguem características importantes para programas de melhoramento, é *Coffea arabica* a espécie de maior relevância comercial. Os cafés de melhor qualidade são produzidos a partir de variedades dessa espécie, que representa cerca de 70% da produção mundial e é cultivada em muitos países em desenvolvimento representando uma fonte importante de renda e emprego.

Algumas dúvidas ainda permanecem sobre os cruzamentos que originaram essa espécie que é a única alotetraploide descrita para o gênero. Entre as espécies diploides, *Coffea canephora*, *Coffea congensis* e *Coffea eugenioides* são apontadas como as mais estreitamente relacionadas à *Coffea arabica* (RAINHA; MUKAI; YAMAMOTO, 1998; LASHERMES et al., 1999; PETITOT; LECOULS; FERNANDEZ, 2008).

A poliploidia é uma característica genômica marcante nos eucariontes, especialmente em plantas. Estima-se que a ocorrência de hibridação seja frequente em cerca de 10% das espécies animais e de 25% entre as plantas (MALLET, 2007). A ocorrência frequente de poliploidia sugere que haja uma vantagem evolutiva de se ter vários conjuntos de material genético para a adaptação.

Apesar disso, no entanto, a geração de novas espécies por meio de hibridações interespecífica, intergenéricas ou por poliploidização envolve muitas alterações genéticas e epigenéticas. Quando dois genomas diferentes são combinados em uma única célula, eles precisam adaptar-se às consequências

desta duplicação. O aumento no número de genes pode induzir desenvolvimento anormal, susceptibilidade a doenças, redução do *fitness* e do vigor híbrido (CHEN, 2007; DOYLE et al. 2008; KENAN-EICHLER et al., 2011).

Estudos em diversas espécies têm demonstrado que os pequenos RNAs, entre eles os miRNAs, são os principais moduladores das respostas epigenéticas à hibridação de dois ou mais genomas. Esses trabalhos têm demonstrado que os RNAs de silenciamento podem funcionar como “tampões genéticos” influenciando a funcionabilidade e estabilidade de genomas recém combinados.

Neste Capítulo a técnica *stem-loop* RT-PCR foi aplicada para monitorar a expressão de miRNAs maduros no alotetraploide e em três espécies diploides relacionadas, *Coffea canephora*, *Coffea congensis* e *Coffea eugenoides*. Essa análise nos permite conhecer o padrão de expressão apresentado por estes RNAs, verificar a presença e a relação destes em espécies com diferentes níveis de ploidia e inferir sobre a ocorrência de possíveis alterações em sequência durante o processo de especiação. Além disso, é apresentado também nesta secção a técnica de *Rapid Amplification of cDNA Ends* (RACE) que nos permitiu confirmar a regulação de miR167 sobre seus genes alvo predito *in silico*.

2 MATERIAIS E MÉTODOS

2.1 Coleta do material vegetal

O material vegetal utilizado nas análises *in vitro* foi obtido na Fazenda Santa Elisa, Centro de Análise e Pesquisa Tecnológica do Agronegócio do Café "Alcides Carvalho", um dos centros de pesquisa do Instituto Agrônomo de Campinas (IAC), Campinas-SP, Brasil. Atualmente, a coleção germoplasma do

IAC é uma das coleções *ex situ* mais representativas do gênero *Coffea*. A coleção inclui acessos de 14 espécies do gênero *Coffea* L. e três espécies do gênero *Psilanthus* Hook F. As espécies mais bem representadas na coleção são *Coffea arabica* L. e *Coffea canephora* Pierre ex A. Froehner (SILVESTRINI et al., 2008).

Foram coletadas folhas das espécies *Coffea arabica*, *Coffea canephora*, *Coffea congensis* e *Coffea eugenioides*. As folhas de cada espécie foram destacadas de três indivíduos diferentes, obtendo-se sempre folhas dos ramos medianos das plantas. As folhas foram divididas em três amostras (amostras 1, 2 e 3) perfazendo no total amostras de folhas de 12 indivíduos. O tecido vegetal foi colocado em tubos Falcon, e imediatamente mergulhado em N₂ líquido. Posteriormente o material coletado foi armazenado em freezer -80°C até a extração do RNA.

2.2- Extração de RNA

O RNA total do tecido foliar das referidas amostras foi extraído utilizando-se o método do reagente CONCERT® (Invitrogen). Dois gramas de tecido de cada amostra foi macerado em N₂ líquido até a obtenção de um pó fino. A cada microtubo contendo o tecido pulverizado foram adicionados 0,5 mL de tampão de extração (reagente CONCERT), homogeneizando a solução por agitação. A mistura foi deixada à temperatura ambiente por dois minutos. Procedeu-se a centrifugação à máxima velocidade da centrífuga (10.000 g) a 4°C por dois minutos. A fase líquida foi coletada e transferida para um novo tubo, adicionado 100 µL de NaCl 5M e 300 µL de clorofórmio, homogeneizando novamente a solução por agitação. O material foi submetido à centrifugação por 10 minutos à máxima velocidade a 4°C. A fase superior foi coletada e transferida para um novo tubo adicionando o mesmo volume de isopropanol. O

material ficou por 30 minutos à temperatura ambiente para promover a precipitação dos ácidos nucléicos. Posteriormente foi realizada centrifugação a 4°C em máxima velocidade por 25 minutos. O líquido do tubo foi descartado tomando cuidado com o precipitado que foi formado no fundo. Foram adicionados 600 µL de etanol 75% e centrifugado por 5 minutos a 4°C na máxima velocidade da centrífuga. O etanol foi descartado tomando cuidado com o precipitado que foi deixado em banho seco a 37°C até secar. O material foi eluído em 20 µL água miliQ autoclavada. A integridade do RNA extraído foi visualizada em gel de agarose 1,2%, corado com brometo de etídio, sob luz ultravioleta. As amostras foram analisadas em espectrofotômetro (Nanodrop® Spectrophotometer ND-1000) para detecção de concentração e qualidade.

2.3 Tratamento com DNase

As amostras de RNA total foram tratadas com o *kit DNA Free* (Applied Biosystems) segundo especificações do fabricante, para eliminação de eventual DNA presente. Uma reação contendo 10 µg de RNA total, 1 µL de enzima rDNase, 5 µL de tampão de reação 10X e água destilada, para completar um volume final de 50 µL. Essa reação foi incubada a 37°C por 30 minutos e adicionou-se 5 µL da solução para inativar DNase (DNase inactivation). As amostras foram incubadas por 2 minutos à temperatura ambiente, centrifugadas a 10.000 g por 1,5 minutos e o sobrenadante foi transferido para novos tubos. O RNA livre de DNA foi então armazenado a -20° C.

2.4 Síntese cDNA

A técnica utilizada para quantificar a expressão de microRNAs maduros difere um pouco da técnica tradicional, uma vez que amplifica fragmentos muito pequenos (19 a 24 bases). Para amplificar estes RNAs são necessários três *primers*. O *primer stem-loop* tem aproximadamente 50 nucleotídeos e forma um grampo estável apenas a baixas temperaturas. Sua complementaridade com o miRNA alvo é de apenas 6 bases da região 3'. O *primer forward* é praticamente igual ao miRNA alvo, exceto pelas bases que se anelam ao *stem-loop* e uma extensão 5' que tem função termodinâmica. O reverso é usado em qualquer amplificação de miRNAs, este *primer* é complementar a uma sequência universal presente no *primer stem-loop*. A Figura 3.1 ilustra o processo resumidamente. Estes *primers* foram desenhados de acordo com a metodologia descrita por Chen et al. (2005). Com esta metodologia, a especificidade do ensaio tem início já na transcrição reversa dos pequenos RNAs, um cDNA deve ser construído para cada família MIR que se queira analisar.

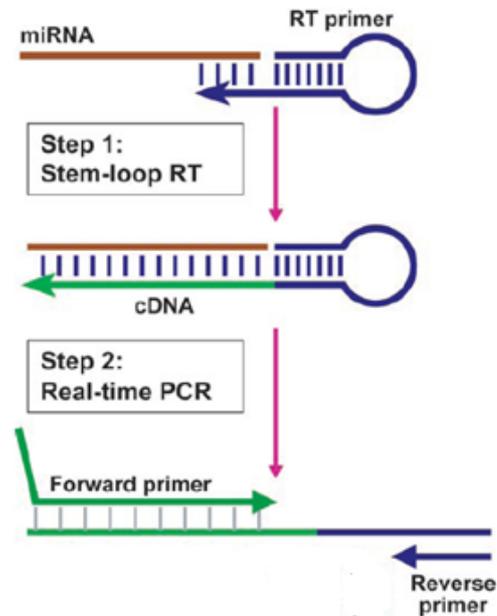


Figura 3.1 Resumo do processo de amplificação de miRNAs em *stem-loop*-RT-PCR. O *primer stem-loop* é utilizado para estender o miRNA maduro durante a síntese de cDNA. Pelo tamanho reduzido do *primer forward*, bases adicionais na extremidade 5' podem ser necessárias para garantir a estabilidade ideal. O *primer* reverso é universal, sempre se anela ao trecho que forma o grampo do primeiro oligonucleotídeo. Adaptado de Chen et al. (2005).

Neste trabalho a síntese de cDNA foi realizada a partir do RNA total tratado, utilizando-se o *Kit Impron-II Reverse Transcriptase* (Promega). Os RNAs utilizados foram escolhidos baseados na qualidade avaliada por espectrofotômetro através das relações 260/280 e 260/230 e pela visualização das bandas 18 e 28S em gel de agarose 1%. Utilizamos 1 mg (aprox. 5.0 μL) de RNA, 1.0 μL de *Primer oligodT*, 1.0 μL de *Primer RT* específico, 1.0 μL de dNTP mix, 4.0 μL do tampão, 2.4 de Mg^{+2} , 0.5 μL de RNase out, 1.0 μL da enzima Impron-II e completamos o volume da reação para 20 μL com água livre de DNase/RNase. As condições do termociclador foram: 16°C por 30 min, 30°C

por 30 seg, 42°C por 30 seg, 50°C por 1 seg (as temperaturas 1, 2 e 3 foram repetidas 59 vezes), depois desse ciclo a reação é exposta a uma temperatura de 85°C por 5 minutos. Ao final as amostras foram armazenadas em freezer a -20°C.

2.5 Expressão dos genes MIR por *stem-loop* RT-PCR

Para a análise de expressão dos genes MIR 167/171/172/319/390/414/417/482, foi utilizado o aparelho ABI PRISM 7500 *Real-Time* PCR (Applied Biosystems), com método de detecção *SYBR Green*, utilizando como template as amostras de cDNA obtidas das folhas de diferentes espécies de café.

Os *primers forward* e *stem-loop* foram desenhados de acordo com as sequências apuradas pela análise *in silico* apresentada no Capítulo 2. Para as famílias MIR compostas por mais de um candidato, a sequência com menos *mismatches* foi utilizada no desenho dos iniciadores.

Foram construídas curvas padrão com as diluições cDNA: água de 1/5, 1/25, 1/125 e 1/625, para verificar a eficiência dos *primers* desenhados. As condições térmicas da reação foram 2 minutos a 50°C, 10 minutos a 95°C, seguidos por 40 ciclos de 15 segundos a 95°C e 1 minuto a 60°C, e finalizando com 15 segundos a 95°C. Os dados foram coletados e armazenados no programa *7500 Fast Software* (Versão 2.1). Para cada reação, foram utilizados 1 µL de cDNA na diluição de 1/25, 1,5 µM de cada *primer* e 5,0 µL de *Master Mix SYBR green* (Applied) para um volume final 10,0 µL/amostra. As amostras foram processadas em triplicatas. Os resultados foram normalizados usando C_{T_s} (Ciclo *Threshold*) obtidos pela expressão dos genes endógenos β -actina e gliceraldeído-3-fosfato desidrogenase (GAPDH) presentes na mesma reação.

Os valores de emissão de fluorescência, captados pelo equipamento durante a amplificação do molde na reação de PCR, são diretamente proporcionais à quantidade de cDNA da amostra. Esses valores são traduzidos pelo *software* e plotados em um gráfico que mostra a intensidade de fluorescência versus o número de ciclos. Quanto maior a quantidade de cDNA molde presente no início da reação, menor é o número de ciclos necessários para detecção de uma intensidade de fluorescência estatisticamente significativa. Esse ponto é definido como limiar ou Ct (*cycle threshold*) e ocorre durante a fase exponencial de amplificação. Assim, cada amostra apresenta um valor de Ct específico. O C_T foi determinado pelo número de ciclos no qual a fluorescência gerada dentro de uma reação cruza a linha *Threshold*.

Para o cálculo da expressão, utilizamos o método do C_T comparativo. Cada amostra é normalizada com os controles endógenos, utilizando a equação $\Delta C_T = C_T$ (gene alvo) - C_T (controle endógeno). O fator de normalização foi a média geométrica dos genes β -actina e gliceraldeído-3-fosfato desidrogenase (GAPDH), a quantificação relativa foi obtida pela fórmula $2^{-\Delta\Delta C_T}$ (LIVAK; SCHMITTGEN, 2001). A Tabela 3.1 contém as sequências dos *primers* utilizados no experimento de *stem-loop* RT-PCR.

Tabela 3.1 *Primers* utilizados no experimento qRT-PCR.

miR167	sequência miR <i>RT primer</i> <i>Forward primer</i>	UGAAGCUGCCAGCAUGAUCU GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC TCGCGTGAAGCTGCCAGCAT
miR171	sequência miR <i>RT primer</i> <i>Forward primer</i>	GUGAGCCGAACCAUAUCACU GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC TTCCTTGATTGAGCCGCGCC
miR172	sequência miR <i>RT primer</i> <i>Forward primer</i>	UGAAUCUUGAUGAUGCACAC GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACATGTAG CCTGAGTGAATCTTGATGATG
miR319	sequência miR <i>RT primer</i> <i>Forward primer</i>	AUUGGACUGAAGGGAGCUCC GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGGGCGA CCTGAGATTGACTGAAGGGA
miR390	sequência miR <i>RT primer</i> <i>Forward primer</i>	AAGCUCAGGAGGGAUAGCGCC GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGGCGCT CCTGAGAAGCTCAGGAGGGAT
miR414	sequência miR <i>RT primer</i> <i>Forward primer</i>	UCAUCCUCAUCAUCGUCC GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGGACGA CCTGAGTCATCCTCATCATCA
miR417	sequência miR <i>RT primer</i> <i>Forward primer</i>	GAAUGUAGUGAAUUUGUCCA GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTGGAAC CCTGAGGAATGTAGTGAATTT
miR482	sequência miR <i>RT primer</i> <i>Forward primer</i>	UCUUCCCUACUCCUCCAUUCC GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGGAATG CCTGAGTCTTCCCTACTCCTCC
universal	<i>Reverse primer</i>	GTGCAGGGTCCGAGGT
	<i>Forward primer</i>	TGTGCTCTTTAGCTTCCAAACG
14-3-3	<i>Reverse primer</i>	CTTCACGAGACATATTGTCTTACTCAA
	<i>Forward primer</i>	TTGAAGGGCGGTGCAA
Gapdh	<i>Reverse primer</i>	AGCAAGGATGCACCCATGTT
	<i>Forward primer</i>	AAGACAGCTTCAACAGAGTACAGCAT
Ubiquitina	<i>Reverse primer</i>	GGCAGGACCTTGGCTGACTATA
	<i>Forward primer</i>	AATTGTCCGTGACATCAAGGAA
β-actina	<i>Reverse primer</i>	TGAGCTGCTTTGGCTGTTC

2.6 Seleção dos genes alvo

Para selecionar os genes alvo a serem analisados via RACE, a metodologia revisada em Zhang et al., (2006) foi adotada. A sequência madura de miR167 serviu como *query* para a busca BLASTn feita no banco de ESTs do café. Os parâmetros do algoritmo *Expected threshold* e *Word length* foram colocados em 1.000 e 7, respectivamente. Foram selecionados os ESTs com identidade nucleotídica maior ou igual a 95% e, como os miRNAs são complementares aos genes que regulam, as sequências que estivessem no sentido antisenso foram selecionadas. Com relação a *mismatches* os seguintes parâmetros foram considerados, das bases 1 a 9 do miRNA: máximo de 2 *mismatches* foram aceitos; das bases de 10-15: no máximo de 1 *mismatch*; das bases de número 16 a 21: no máximo 3 *mismatches*. As sequências apuradas foram utilizadas para desenhar os *primers* utilizados no experimento de confirmação da clivagem direcionada por miR167.

2.7 Análise de 5' RACE

Cinco microgramas de RNA total de folhas de *Coffea arabica* e *Coffea canephora* foram ligados a 0.25 µg do oligo de RNA em uma reação contendo 1 µL de T4 RNA Ligase (5 U/µL), 1 µL de RNase out 40U/µL, 1 µL de ATP 10 mM, 3 µL de tampão da ligase 10X completando o volume com água para 30 µL. Posteriormente a reação foi incubada a 37°C por 1 hora, e depois colocada no gelo. O RNA total foi purificado com fenol:clorofórmio e precipitado com etanol, acetato de sódio e glicogênio. O *pellet* resultante foi então ressuspenso em 10 µl de água DEPC autoclavada.

Este RNA ligado ao oligo de RNA GeneRacer foi utilizado como molde para reação de transcrição reversa utilizando a enzima Superscript III (Invitrogen) seguindo as recomendações do fabricante. O cDNA foi utilizado

para amplificação dos fragmentos clivados do gene ARF8 utilizando o iniciador direto 5' GENE RACER do *kit* e o iniciador reverso ARF8 R2, específico para a sequência do mRNA-alvo (ver Tabela 3.2).

Em seguida o produto da primeira reação foi usado como molde para a PCR NESTED fazendo uso dos iniciadores 5' GENE RACER NESTED e reverso ARF8 R1 (Tabela 3.2). Depois da amplificação, o produto fresco 5' RACE foi clonado no vetor TOPO PCR 2.1. Após triagem de colônias, cinco clones de cada espécie foram sequenciados.

A Tabela 3.2 contém as sequências dos *primers* utilizados no experimento de 5' Race.

Tabela 3.2 *Primers* utilizados no experimento 5' RACE e de qRT-PCR.

Gene alvo ARF8	Iniciador	Sequência
	<i>Gene Racer 5'</i>	CGACTGGAGCACGAGGACACTGA
	<i>Gene Racer 5'Nested primer forward</i>	GGACACTGACATGGACTGAAGGAGTA ATCGAAGGGTTGCTTGAAGA
	<i>primer reverse 1ARF8</i>	GCACTTCCTCTGGAGACAGG
	<i>primer reverse 2ARF8</i>	ATGCGACTCAAGGATTCACC

3 RESULTADOS E DISCUSSÃO

3.1 *Stem-loop* RT-PCR

O presente trabalho foi proposto com o objetivo de investigar a expressão de miRNAs maduros em folhas de diferentes espécies de café. De oito famílias MIR caracterizadas no Capítulo 2, apenas a metade apresentou padrões satisfatórios de amplificação. A eficiência do ensaio tem um grande impacto sobre a precisão da quantificação da expressão gênica. Nos testes de eficiência com curvas padrão, os *primers* desenhados para os miRNAs miR167, miR171, miR172 e miR390 apresentaram eficiência de mais de 90%, com R2 de 0.998,

0.992, 0.998 e 0.996, respectivamente. Os outros quatro conjuntos de *primers* apresentaram curvas de *melting* com mais de um pico, indicando ampliações inespecíficas, e por isso foram excluídos da análise de expressão. A revelação do produto de PCR em gel de agarose 3% mostra que os *primers* desenhados para miR319, miR414, miR417 e miR482 são capazes de se anelar tanto a genes MIR (amplicon menores que 100pb) quanto a genes codantes de RNAs mensageiros (amplicons maiores que 100pb) (Figura 3.2).

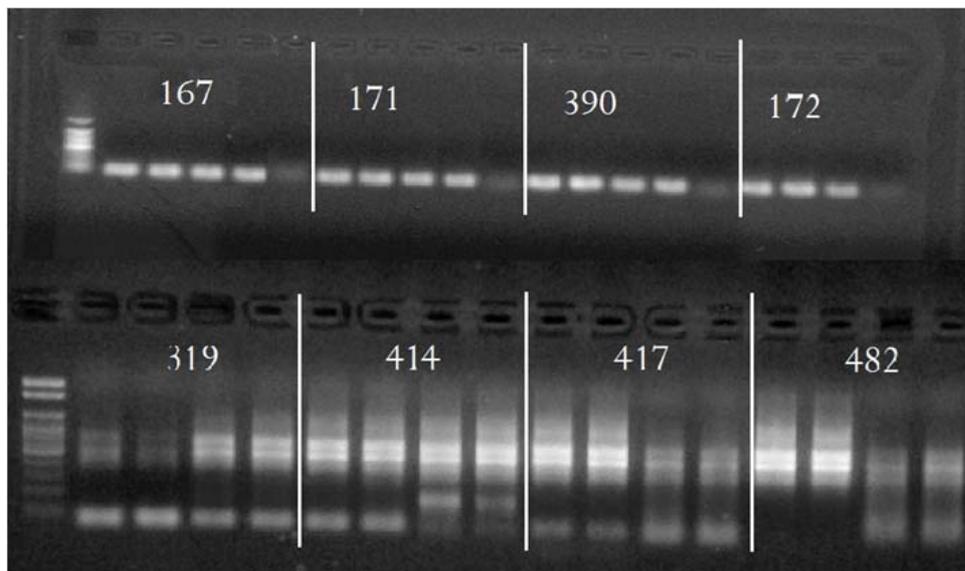


Figura 3.2 Eletroforese em gel de agarose 3% do produto de PCR de diferentes famílias MIR. A porção superior da figura revela amplificação correta dos miRNAs pertencentes às famílias miR167/171/172/390. Na porção inferior, diferentes bandas são vistas em cada canaleta, indicando amplificação inespecífica dos *primers* desenhados para as famílias miR319/414/417/482. A amplificação destes oligonucleotídeos foi verificada em quatro espécies de café, a ordem em que aparecem no gel é *Coffea arabica*, *Coffea canephora*, *Coffea congensis* e *Coffea eugenioides*. Na parte superior da figura os controles negativos (NTCs) correspondem à última canaleta de cada grupo. Para o segundo conjunto de *primers* os NTCs não foram representados na figura.

Transcritos de todas as famílias MIR que permaneceram na análise *in vitro* foram detectados nas quatro espécies avaliadas neste estudo. Como discutido mais adiante, o nível de expressão dessas famílias mostrou-se proporcional ao nível de ploidia do genoma.

Quatro genes constitutivos, entre eles, β -actina, GAPDH, ubiquitina e 14.3.3, foram testados como padrões internos para normalizar os valores de expressão do microtranscriptoma. No entanto, nenhum destes controles endógenos apresentou estabilidade adequada, apresentando grandes desvios entre os valores de Ct encontrados para as amostras de *Coffea eugenioides* e às demais espécies estudadas. Em todos os testes a expressão destes genes foi tardia em *Coffea eugenioides*, apresentando diferença de cerca de 3 ciclos *threshold*. A falta de estabilidade dos genes endógenos prejudicou os resultados de expressão relativa. Como o nível de expressão dos miRNAs foi muito parecido entre as quatro espécies avaliadas, o desvio nos valores de expressão dos genes referência em *Coffea eugenioides* superestima o resultado encontrado para esta.

Ao plotar os resultados em gráfico, é representada uma super-expressão de todas as famílias MIR em *Coffea eugenioides*. Por exemplo, a expressão do microRNA miR172 em *Coffea canephora* é de 0,43 em relação à *Coffea arabica*. Analisando os ciclos *threshold* das triplicatas técnicas desta espécie, temos os seguintes valores 31,979/ 31,632/ 31,371. Para *Coffea eugenioides* os valores dos ciclos *threshold* são praticamente os mesmos, 31,951/ 31,340/ 31,725 (ver Figura 3.3). No entanto, mesmo com níveis de acumulação de miR172 praticamente iguais, o gráfico referente a este miRNA mostra uma expressão de 3,6 vezes maior de *Coffea eugenioides* com relação a *Coffea arabica* (Figura 3.4).

Sample N...	Target Na...	Task	Dyes	Ct	Ct Mean	Ct SD
canephora	mir 172	UNKNOWN	SYBR-None	31,979	31,66	0,305
canephora	mir 172	UNKNOWN	SYBR-None	31,632	31,66	0,305
canephora	mir 172	UNKNOWN	SYBR-None	31,371	31,66	0,305
canephora	gapdh	UNKNOWN	SYBR-None	23,489	23,511	0,038
canephora	gapdh	UNKNOWN	SYBR-None	23,537	23,511	0,038
canephora	gapdh	UNKNOWN	SYBR-None	23,566	23,511	0,038
eugenioides	mir 172	UNKNOWN	SYBR-None	31,951	31,672	0,309
eugenioides	mir 172	UNKNOWN	SYBR-None	31,340	31,672	0,309
eugenioides	mir 172	UNKNOWN	SYBR-None	31,725	31,672	0,309
eugenioides	gapdh	UNKNOWN	SYBR-None	26,640	26,665	0,094
eugenioides	gapdh	UNKNOWN	SYBR-None	26,852	26,665	0,094
eugenioides	gapdh	UNKNOWN	SYBR-None	26,609	26,665	0,094

Figura 3.3 Representação dos ciclos *threshold* dos genes miR172 e GAPDH nas espécies *Coffea canephora* e *Coffea eugenioides*.

Eliminando os dados de expressão de *Coffea eugenioides*, os gráficos construídos mostram uma relação de proporcionalidade entre os valores de expressão de *Coffea arabica* e *Coffea canephora*. Para todas as famílias MIR o experimento de qPCR mostrou que a quantidade de miRNAs acumulados em *Coffea canephora* corresponde a um valor muito próximo da metade do valor encontrado para as amostras de *Coffea arabica*. O mesmo não foi visto para *Coffea congensis* que apresentou valores de expressão totalmente variáveis. (Figuras 3.5 a 3.8).

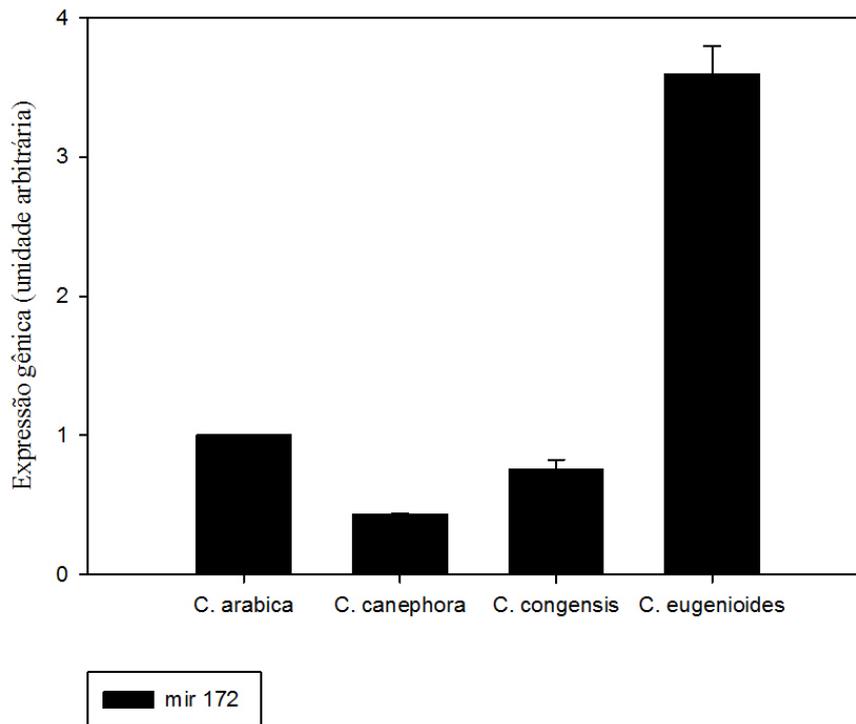


Figura 3.4 Perfil de expressão relativa de miR172 em folhas jovens de *Coffea arabica*, *Coffea canephora*, *Coffea congensis* e *Coffea eugenioides*. O nível do transcrito é representado como uma razão (expressão relativa) do valor absoluto da expressão do gene MIR pelo valor absoluto da expressão dos genes normalizadores β -actina e GAPDH. No gráfico os valores de expressão de *Coffea canephora* e *Coffea congensis* foram plotados tendo o valor de expressão de *Coffea arabica* como referência. As barras verticais representam o desvio padrão de duas amostras biológicas.

Como visto anteriormente, mesmo para miRNAs conservados, os níveis e/ou padrões de expressão não se mantêm necessariamente iguais em espécies diferentes ou em diferentes estágios de desenvolvimento de uma mesma espécie (WANG et al., 2006; HEGARTY et al., 2006).

Ha et al. (2009) analisaram a expressão de siRNAs e miRNAs em duas espécies estreitamente relacionadas, *Arabidopsis thaliana* e *Arabidopsis arenosa* e em linhagens (de F1 a F7) de alotetraploides derivados dessas espécies. Nesse trabalho os autores encontraram que os padrões de expressão de miRNAs foram altamente variáveis entre os alotetraploides e seus progenitores. A maioria dos miRNAs testados foram expressos não-aditivamente e a regulação dos genes alvos destoou do que é encontrado tanto para *Arabidopsis thaliana* quanto para *Arabidopsis arenosa*. Em contraste, com relação aos siRNAs, nenhuma ligação foi vista entre estes e o padrão de controle da expressão gênica adotado pelo genomas híbridos. Os autores discutem que a herança estável de siRNAs pode estar relacionada à manutenção da cromatina e estabilidade do genoma, enquanto que a variação dos miRNAs pode levar a mudanças na expressão gênica global, mudanças fenotípicas, vigor híbrido e adaptação. Essas alterações podem ser causadas por um ou vários fatores em combinação, podendo ser vistas principalmente nas gerações imediatamente seguintes à hibridação. Essas são características de um padrão de expressão não-aditivo (HEGARTY et al., 2006; HA et al., 2009).

Comai (2000) e Wang et al. (2006), dentre outros autores, discutem que os padrões de regulação gênica de um genoma híbrido depende muito da divergência de expressão entre os genomas parentais. Sendo que, quanto mais diversos forem os genes envolvidos no processo de hibridação, maiores são as chances de aparecimento de alterações *cis* e *trans* nas sequências promotoras dos loci envolvidos na biogênese dos pequenos RNAs. Por outro lado, se os loci de miRNAs são originários de espécies intimamente relacionadas a probabilidade de variação nos padrões de expressão é menor. Essa proximidade entre as espécies progenitoras minimiza possíveis divergências nas redes pós-transcricionais no poliploide formado, diminuindo a possibilidade de variação

em sequência, estrutura e preferência por alvos nas vias metabólicas reguladas (WANG et al., 2006; CHEN, 2007).

O experimento de qPCR mostrou que os miRNAs de café apresentam uma resposta à poliploidização diferente daquela vista para espécies como *Senecio cambrensis* (HEGARTY et al., 2006), *Gossypium hirsutum* L. (FLAGEL et al., 2008) e *Arabidopsis thaliana* (HA et al., 2009). Para o café, o nível de expressão dos genes MIR aumentou proporcionalmente com nível de ploidia do genoma.

Nossos resultados corroboram com os dados publicados por Kenan-Eichler et al. (2011). A fim de caracterizar os eventos epigenéticos que ocorrem durante a especiação via aloploidização em plantas de trigo, os pesquisadores sequenciaram o microtranscriptoma de parentais, de híbridos intergenéricos e de aloploidos. Os autores relatam que, enquanto o percentual de siRNAs diminuiu, o número de pequenos RNAs correspondentes a miRNAs aumentou de acordo com o nível de ploidia. A menor porcentagem de loci MIR com relação à quantidade total de pequenos RNAs foi relatada para o genoma diploide *Aegilops tauschii* (21%); uma porcentagem de 33% foi relatada para o híbrido triploide; 38% para o tetraploide *Triticum turgidum* e 44% do microtranscriptoma sequenciado era de miRNAs no hexaploide sintético.

No caso de aloploidia, a expressão aditiva pode fornecer a base molecular necessária para o equilíbrio e compensação da dose extra de genes redundantes reunidos em um único núcleo. Esse padrão de expressão pode conferir estabilidade ao desenvolvimento da espécie híbrida (VEITIA; BOTTANI; BIRCHLER, 2008). Condizente com o modelo aditivo de expressão, em *Coffea arabica* que apresenta número de cromossomos $2n = 4x = 44$, a acumulação das famílias MIR corresponde a aproximadamente o dobro dos valores encontrados para espécie *Coffea canephora* ($2n = 2x = 22$ cromossomos) (Figura 3.5 a 3.8).

Como visto anteriormente, os trabalhos de Lashemers et al. (1999) e Petitot, Lecouls e Fernandez (2008) discutem e apontam evidências sobre a ancestralidade de *Coffea arabica*. Esses autores defendem que o processo de hibridação que originou o genoma alotetraploide deu-se pela interação entre as espécies *Coffea canephora* e *Coffea eugenioides*. Embora não seja visual nos gráficos de expressão, por motivos já discutidos, em nosso experimento os níveis de acúmulo dos miRNAs foi praticamente o mesmo para estas espécies diploides (Figura 3.3).

É importante notar que, embora as modificações nas redes de regulação pós-transcricional sofram rápidas mudanças entre as gerações e por isso podem ser melhor analisadas em linhagens imediatamente seguintes ao processo de hibridação, o presente estudo não analisou plantas que estabelecessem uma relação direta de parentesco. Apesar disso, assim como a hibridização *in situ* conduzida por Lashermes et al., (1999) e a análise funcional do gene WRKY1 publicada por Petitot, Lecouls e Fernandez (2008), os nossos resultados apontam para uma relação mais evidente entre *Coffea arabica*, *Coffea canephora* e *Coffea eugenioides*.

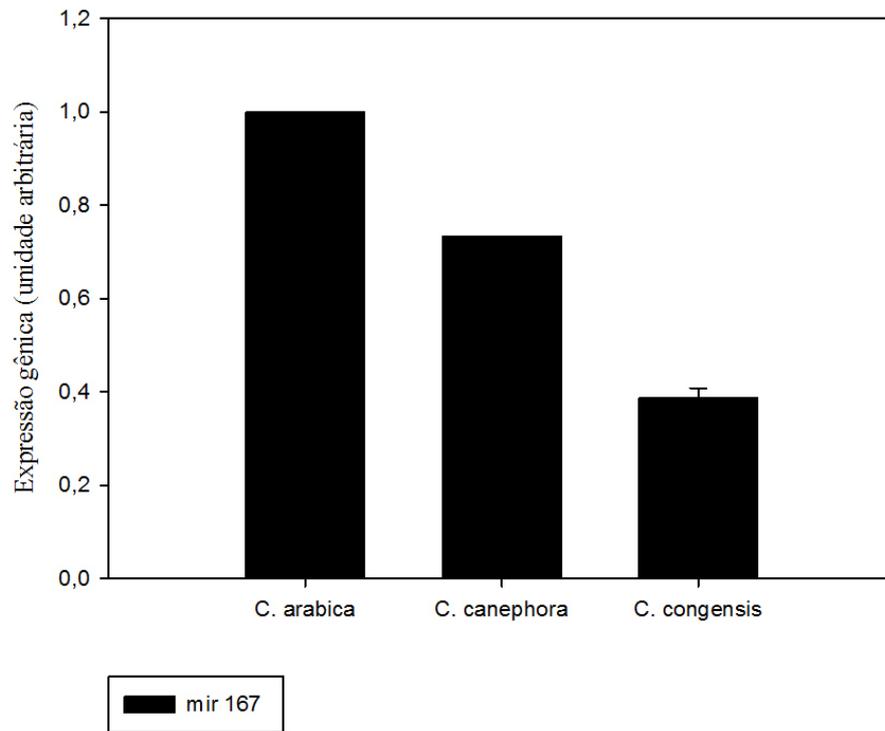


Figura 3. 5 Perfil de expressão relativa de miR167 em folhas jovens de *Coffea arabica*, *Coffea canephora* e *Coffea congensis*. O nível do transcrito é representado como uma razão (expressão relativa) do valor absoluto da expressão do gene MIR pelo valor absoluto da expressão dos genes normalizadores β -actina e GAPDH. No gráfico os valores de expressão de *Coffea canephora* e *Coffea congensis* foram plotados tendo o valor de expressão de *Coffea arabica* como referência. As barras verticais representam o desvio padrão de duas amostras biológicas.

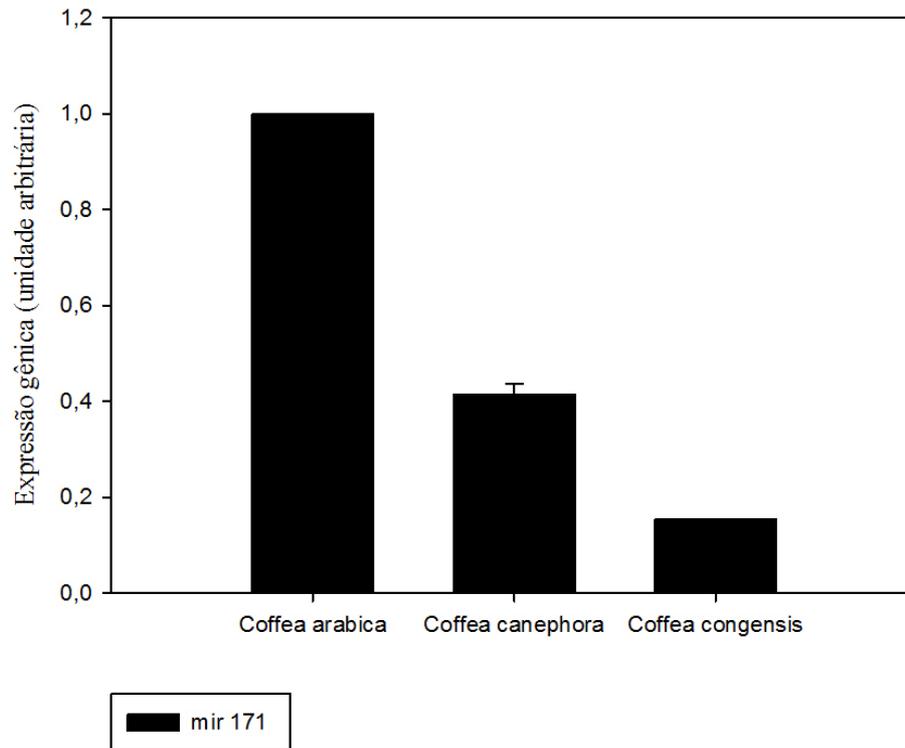


Figura 3.6 Perfil de expressão relativa de miR171 em folhas jovens de *Coffea arabica*, *Coffea canephora* e *Coffea congensis*. O nível do transcrito é representado como uma razão (expressão relativa) do valor absoluto da expressão do gene MIR pelo valor absoluto da expressão dos genes normalizadores β -actina e GAPDH. No gráfico os valores de expressão de *Coffea canephora* e *Coffea congensis* foram plotados tendo o valor de expressão de *Coffea arabica* como referência. As barras verticais representam o desvio padrão de duas amostras biológicas.

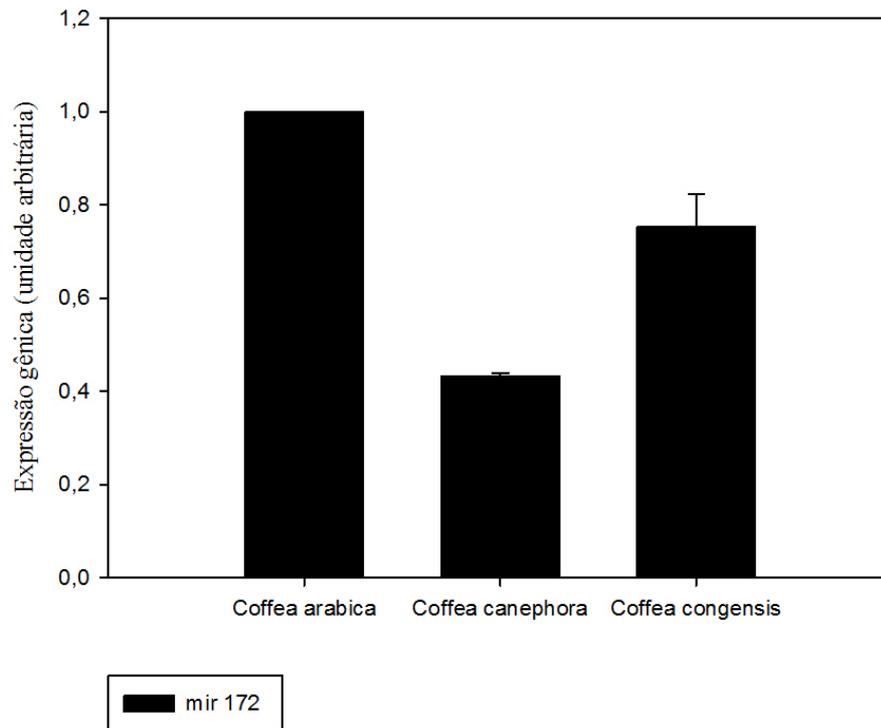


Figura 3.7 Perfil de expressão relativa de miR172 em folhas jovens de *Coffea arabica*, *Coffea canephora* e *Coffea congensis*. O nível do transcrito é representado como uma razão (expressão relativa) do valor absoluto da expressão do gene MIR pelo valor absoluto da expressão dos genes normalizadores β -actina e GAPDH. No gráfico os valores de expressão de *Coffea canephora* e *Coffea congensis* foram plotados tendo o valor de expressão de *Coffea arabica* como referência. As barras verticais representam o desvio padrão de duas amostras biológicas.

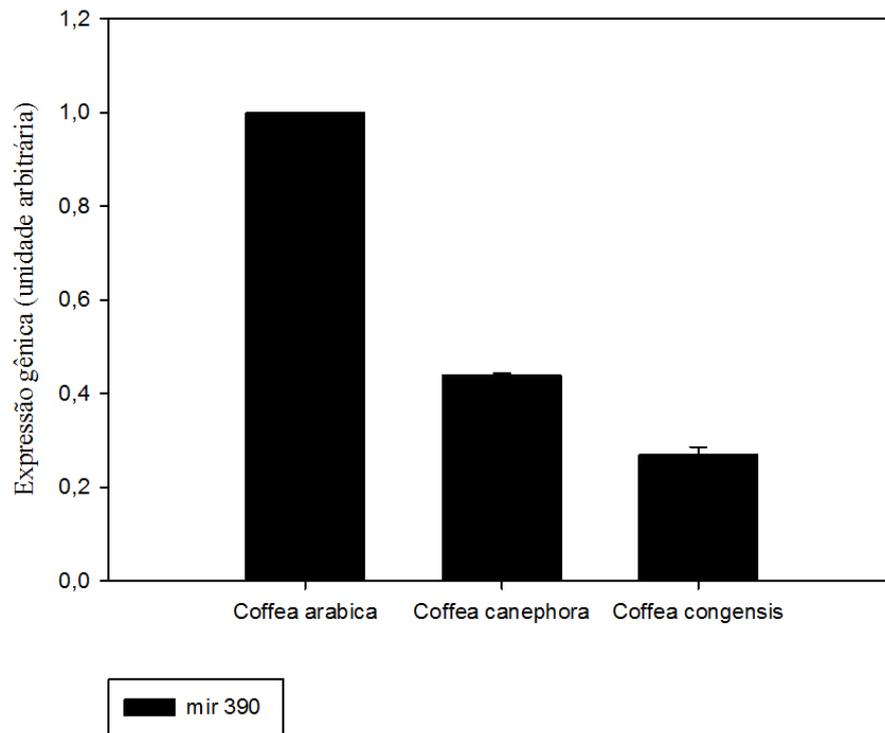


Figura 3.8 Perfil de expressão relativa de miR390 em folhas jovens de *Coffea arabica*, *Coffea canephora* e *Coffea congestis*. O nível do transcrito é representado como uma razão (expressão relativa) do valor absoluto da expressão do gene MIR pelo valor absoluto da expressão dos genes normalizadores β -actina e GAPDH. No gráfico os valores de expressão de *Coffea canephora* e *Coffea congestis* foram plotados tendo o valor de expressão de *Coffea arabica* como referência. As barras verticais representam o desvio padrão de duas amostras biológicas.

A relação de proporção encontrada entre os valores de expressão destas três espécies não é visto em *Coffea congestis* e nos leva a inferir que cada um dos dois subgenomas diploides presentes em *Coffea arabica* contribuiu de forma equitativa com as informações genéticas correspondente a miRNAs.

Com relação ao desvio de expressão visto para miR167 que em *Coffea canephora* apresentou expressão de 0,73 em relação ao genoma tetraploide (Figura 3.5), alterações na estrutura da cromatina e a herdabilidade diferencial de alguns genes podem ser a explicação. Se por um lado, os miRNAs e seus alvos tendem a coevoluir e serem corregulados alguns componentes da cromatina, as proteínas *Argonaute*, *Dicer* e seus cofatores podem evoluir de forma independente entre as espécies (CHAPMAN; CARRINGTON, 2007). Possíveis desbalanços nestes que são fatores centrais do mecanismo de silenciamento, podem modular variações no processo de edição, maturação e no controle do silenciamento dos genes alvo (CHAPMAN; CARRINGTON, 2007; HA et al., 2008).

Além disso, alterações na estrutura da cromatina podem mecanicamente explicar peculiaridades no perfil de expressão de alguns pequenos RNAs. Mudanças nas sequências promotoras dos genes MIR ou dos loci diretamente envolvidos com a biogênese dos miRNAs, podem alterar o local e o momento em que o processo de transcrição ocorrerá, podendo fortemente influenciar o modo de expressão desses genes (MALLORY; VAUCHERET, 2009). Tirosh et al. (2009) demonstram que em híbridos interespecíficos, novas interações entre elementos ativadores *cis* e *trans* e as sequências promotoras podem levar à superexpressão ou, alternativamente, a supressão de 10% dos genes.

3.2 5' RACE

Dentre os genes alvos apurados por nossas análises *in silico*, os alvos de miR167, fatores de transcrição auxina responsivos ARF8 foram escolhidos para serem analisados *in vitro*. A família dos Fatores Auxina Responsivos (ARF) são proteínas planta específicas que se ligam a elementos promotores de transcrição das auxinas e controlam as respostas vegetais a esse fito-hormônio. As auxinas desempenham papéis críticos durante o crescimento vegetal (MALLORY;

BARTEL; BARTEL, 2005), elas exercem influência sobre processos vitais como a regulação da embriogênese, o desenvolvimento das raízes e a formação dos órgãos florais (WANG et al., 2005).

Nagpal et al. (2005) relatam que flores de *Arabidopsis* duplo *knockout* ARF6/ARF8, apresentam retardo no tempo de abertura dos botões florais, estames e pétalas reduzidos, anteras incapazes de liberar o pólen, gineceu imaturo, má formação das papilas estigmáticas e fêmea estéril. Os autores relatam também que nos mutantes simples, *knockout* ARF8, embora os filamentos mais curtos dos estames e o atraso na deiscência da antera tenham induzido redução de fecundidade, a mutação não provocou a esterilidade completa. Ambos os fatores de transcrição são negativamente regulados por miR167.

Wu, Tian e Reed (2006) descrevem que em *Arabidopsis* mutações no sítio de clivagem de miR167 causaram expressão ectópica dos genes ARF8 e ARF6. A falta da atuação deste microRNA resultou em atrofia do tegumento dos óvulos e em anteras reduzidas e incapazes de liberar corretamente os grãos de pólen. Resultados como este mostram que miR167 é essencial para a padronização correta da expressão dos genes relacionados à fertilidade vegetal. E que a magnitude das respostas aos miRNAs em plantas difere da que é normalmente vista para os animais, onde estes RNAs geralmente reforçam ou mantêm padrões de expressão transcricionalmente estabelecidos (WU; TIAN; REED, 2006).

A sequência de mi167 caracterizada neste trabalho é quase perfeitamente complementar à sequência de ARF8 em *Arabidopsis* e ao possível gene de Café. Apenas um *missmatche* ocorre no décimo quarto resíduo da sequência do miRNA (Figura 3.9).

miR167	3'TGAAGCTGCCAGCA AT GATCT5'
ARF8 em <i>Arabidopsis</i>	5'*** TTCGACGGTCG G ACTAGA3'
ARF8' em <i>Coffea arabica</i>	5'*** TTCGACGGTCG G ACTAGA3'

Figura 3.9 Sítio do transcrito do gene ARF8 complementar ao microRNA miR 167. *Missmatche* entre as sequências está destacado em cinza.

A análise de expressão de ARF8 e miR167 está representada na Figura 3.10. Os iniciadores utilizados para amplificar o gene alvo estão representados na Tabela 3.2. Como mostra a ilustração, não foi possível identificar qualquer correlação negativa entre a expressão de miR167 e o gene supostamente regulado. Este resultado é, provavelmente, devido ao fato de o material vegetal utilizado não ter sido coletado sob nenhuma condição específica, onde a ativação do mecanismo de silenciamento pós-transcricional pudesse ser evidenciada.

Sabe-se que muitos miRNAs são restritos a específicos tecidos ou a estágios e condições de desenvolvimento, o que sugere uma regulação espacial e temporal da acumulação dos miRNAs e com isso da expressão do seu gene alvo (BARTEL, 2004; MALLORY; BARTEL; BARTEL, 2005).

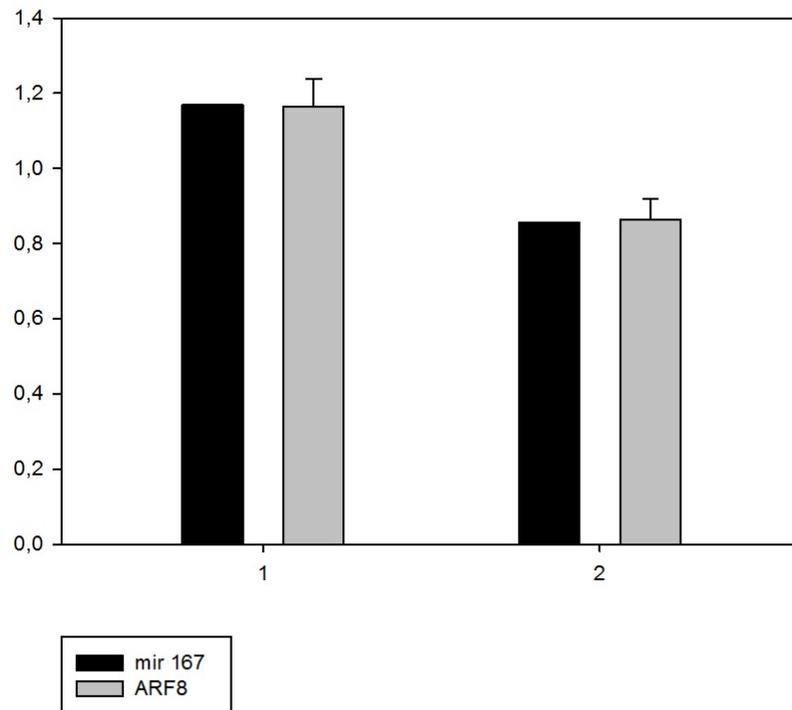


Figura 3.10 Perfil de expressão relativa de miR167 e ARF8 em folhas jovens de *Coffea arabica* (barras com número 1) e *Coffea canephora* (barras com número 2). O nível dos transcritos é representado como uma razão (expressão relativa) do valor absoluto da expressão dos genes de interesse pelo valor absoluto da expressão dos genes normalizadores β -actina e GAPDH. No gráfico os valores de expressão de *Coffea arabica* e *Coffea canephora* foram plotados tendo o valor médio de expressão das duas espécies como referência. As barras verticais representam o desvio padrão de duas amostras biológicas.

Assim, para tal análise seria ideal a utilização de material submetido a estresses bióticos, ambientais, ou a condições específicas de desenvolvimento, e que houvesse também coletas em diferentes intervalos de tempo. Dessa forma a resposta ao silenciamento direcionado por miR167, ou qualquer outro miRNA, poderia ser sistematicamente quantificada.

Como demonstrado na Figura 3.11 os iniciadores utilizados para detectar os transcritos do ARF 8 foram desenhados flanqueando o sítio de reconhecimento desses genes por miR167, isso indica que somente transcritos intactos de ARF8 foram detectados por RT-qPCR. Para confirmar a possível regulação negativa exercida por miR167 sobre transcritos de ARF8, realizou-se um ensaio de 5'RACE.

A técnica de 5'RACE é uma variante da RT-PCR em que um adaptador universal é capaz, por meio da ação da enzima t4 RNA ligase, de se ligar a fragmentos de RNA que apresentem extremidade OH livre. Esse RNA ligado ao adaptador é utilizado como molde para reação de transcrição reversa e a especificidade do ensaio acontece nas etapas de amplificação. Os iniciadores específicos utilizados para amplificar os fragmentos clivados do gene ARF8 estão representados na Tabela 3.2, (Figura 3.11).

Com a clonagem e sequenciamento dos produtos de PCR foi possível, além de confirmar a clivagem de transcritos ARF8, determinar a posição nucleotídica em que o corte guiado pelo RISC ocorreu.

> Coffea arabica CS1 Coffea arabica cDNA clone CA00-XX-CS1-055-H03-EQ.

```
GCTGATTCAAACACTAAATTTTCAGCATCAATAGCTCCTTCAAGCATGCAAAATGTGCTAGGCTCACTCTGTTC
AGAGGGAAGTACTAGTCTTTTGAACCTTCTCCAGAACGACTCAGACCATGCTAAATGAGCAGCCACCTCAGCAAT
CCTGGGTATCTAAATATTCTCATGGACAAGTCAATACTTGTTCGAGTCTGCGTCACTTCCCCATATCTGGGA
AAAGATGGTCCAGTGGGCAGGAAGCTTGTGGCTTGGATGCACAGAATCAAGCTCTTTTGGTGCCAAACATGGA
CTCTGCTGCTCTTTTACTTCCCCTACTGTCTCCAGTGTCCGGTACTTCTTCAATTGATGCTGGTGTGCTTCTA
TGCCATTAGGGGATTCTGGCTTTCAGAATCCTTTATACGGTTTTGTGCATGACTCTTCCGAGTTGTTGCATGGT
ACTGGTCAAGATTACCTACACGAACCCGCACATTTGTCAAGGTCCACAAGTTTGGGTGTGTTGGGAGGTCATT
GGACATCAGCCGGTTCAACAGCTATCATGAGCTCGACAGGAACCTCGGTGAGATGATGGGATCGAAGGGTTGC
TTGAAGACCCCTCAAAGATCAGGCTGGCAGCTTGTATTTTGTGCGACAGGGAGAACGATGTGCTTCTCCTTGGAGAC
GACCCATGGGAGGCATTTGTTAATAATGTCTGGTATATCAAGATCCTGTCTCCAGAGGAAGTGCAGAAGCTGGG
AAAGCAAGAGGGTGAATCCTTGAGTCCGATTTGCTGCTGAGAGGATGAACAGCAGCAGTGGTGTATGGGCGAGACA
TGGTCGCTAGGTTTCCTTTTAGGATCACTAGAATATTGAGAGGGAGTAATCACTAATCACT
```

Figura 3.11 Possível sequência do Fator Responsivo a Auxina 8 de café (sequência parcialmente totalmente representada). Destacado em cinza escuro está a porção do gene ARF8 complementar a mi167, em cinza claro estão os sítios de pareamento para os iniciadores reversos utilizados no 5'RACE, sublinhado estão os sítios de pareamento para os iniciadores direto e reverso utilizados no qRT-PCR.

Para todos os clones de *Coffea arabica* e *Coffea canephora* o sequenciamento mostrou que o sítio de clivagem de miR167 é o mesmo para as duas espécies. Corroborando com o resultado encontrado em *Arabidopsis* por diversos autores (MALLORY; BARTEL; BARTEL, 2005; ALLEN et al., 2005), o corte do transcrito aconteceu no centro da complementaridade entre o miRNA e o mRNA, mais precisamente no décimo primeiro nucleotídeo da porção de ARF8 que se parecia com o miR167 (Figura 3.12).

(A)

```

ARA23 TGAA-GGAGTAGAAAGGCAGCTTGTATTTGTCGACAGGGAGAAC-GATGTGCTTCTC-TT 171
ARA28 TGAA-GGAGTAGAAAGGCAGCTTGTATTTGTCGACAGGGAGAAC-GATGTGCTTCTCCTT 172
ARA24 TGAA-GGAGTAGAAAGGCAGCTTGTATTTGTCGACAGGGAGAAC-GATGTGCTTCTCCTT 160
ARA27 TGAA-GGAGTAG---GGCAGCTTGTATTTGTCGACAGGGAGAAC-GATGTGCTTCTCCTT 168
ARA37 TGAAACGGAGTAGAAAGGCAGCTTGTATTTGTCGACAGGGAGAACAGACGTGCTTCTCCTT 162
      *****
      *****

ARA23 GGAGACGACCCATGGGAGGCATTT-GTTAATAATGTCT-GGTATATCAAGATCCTGTCTC 229
ARA28 GGAGACGACCCATGGGAGGCATTT-GTTAATAATGTCT-GGTATATCAAGATCCTGTCTC 230
ARA24 GGAGACGACCCATGGGAGGCATTTAGTTAATAATGTCT-GGTATATCAAGATCCTGTCTC 219
ARA27 GGAGACGACCCATGGGAGGCATTT-GTTAATAATGTCT-GGTATATCAAGATCCTGTCTC 226
ARA37 GGAGACGACCCATGGGAGGCATCTGTTAATAATGTCTTGGTATATCAAGATCCTGTCTC 222
      ***** * *****

ARA23 CAGAGGAAGTGC AAGGGCGAATTCAGCACACTGGCGCCGTTACTAGTGGATCCGAGCT 289
ARA28 CAGAGGAAGTGC AAGGGCGAATTCAGCACACTGGCGCCGTTACTAGTGGATCCGAGCT 290
ARA24 CAGAGGAAGTGC AAGGGCGAATTCAGCACACTGGCGCCGTTACTAGTGGATCCGAGCT 279
ARA27 CAGAGGAAGTGC AAGGGCGAATTCAGCACACTGGCGCCGTTACTAGTGGATCCGAGCT 286
ARA37 CAGAGGAAGTGC AAGGGCGAATTCAGCACACTGGCGCCGTTACTAGTGGATCCGAGCT 282
      *****

```

(B)

```

CAN3 CGAGGACACTGACATGGACTGA-----AGGGCAGCTTGTATTTGTCGACAGGG 160
CAN19CGANGACACTGACATGGACTGACAGTGAGTAGAAATGGCAGCTTGTATTTGTCGACAGGG 176
CAN8 CGAGGACACTGACATGGACTGAAGG--AGTAGA-AAGGCAGCTTGTATTTGTCGACAGGG 162
CAN4 CGAGGACACTGACATGGACTGAAGG--AGTAGA-AAGGCAGCTTGTATTTGTCGACAGGG 162
CAN6 CGAGGACACTGACATGGACTGAAGG--AGTAGA-AAGGCAGCTTGTATTTGTCGACAGGG 160
      *** ***** * *****

CAN3 AGAACGATGTGCTTCTCCTTGGAGACGACCCATGGGAGGCATTTGTTAATAATGTCT-GG 219
CAN19AGAACGATGTGCTTCTCCTTGGAGACGACCCATGGGAGGCATTTGTTACTAATGTCTTGG 236
CAN8 AGAACGATGTGCTTCTCCTTGGAGACGACCCATGGGAGGCATTTGTTAATAATGTCT-GG 221
CAN4 AGAACGATGTGCTTCTCCTTGGAGACGACCCATGGGAGGCATTTGTTAATAATGTCT-GG 221
CAN6 AGAACGATGTGCTTCTCCTTGGAGACGACCCATGGGAGGCATTTGTTAATAATGTCT-GG 219
      ***** **

CAN3 TATATCAAGATCCTGTCTCCAGAGGAAGTGC AAGG-GCGAATTCAGCACACTGGCGGCC 278
CAN19TATATCAAGATCCTGTCTCCAGAGGAAGTGC AAGGTGCGAATTCAGCACACTGGCGGCC- 295
CAN8 TATATCAAGATCCTGTCTCCAGAGGAAGTGC AAGG-GCGAATTCAGCACACTGGCGGCC- 279

```

```

CAN4 TATATCAAGATCCTGTCTCCAGAGGAAGTCCAAGG-GCGAATTCCAGCACACTGGCGGCC 280
CAN6 TATATCAAGATCCTGTCTCCAGAGGAAGTCCAAGG-GCGAATTCCAGCACACTGGCGGCC 278
***** ** *****

```

Figura 3.12 Alinhamento entre trecho das seqüências ARF8 de *Coffea arabica* (A) e *Coffea canephora* (B) inseridas em cinco clones bacterianos. A seqüência não foi completamente representada. Destacado em cinza claro está o adaptador GeneRACER, em cinza escuro está a porção do gene ARF8 complementar a mi167 e em azul está o primer *Reverse 1* ARF8. Os asteriscos representam homologia de bases. Alinhamento feito com ferramenta disponível em <http://www.ebi.ac.uk/clustalw/>.

4 CONCLUSÕES

De oito famílias MIR caracterizadas *in silico*, apenas quatro apresentaram padrões satisfatórios de amplificação. Os iniciadores desenhados para as famílias miR319, miR414, miR417 e miR482 não se mostraram específicos e por isso não foi possível avaliar o perfil de expressão destes miRNAs no café.

Para as quatro famílias MIR avaliadas o experimento de *stem-loop* RT-PCR indicou que, condizente com o modelo aditivo de expressão, a quantidade de miRNAs acumulados em *Coffea canephora* ($2n = 2x = 22$) corresponde a um valor muito próximo da metade do valor encontrado para as amostras de *Coffea arabica* ($2n = 4x = 44$).

Com a utilização da técnica de 5'RACE foi possível confirmar a presença de transcritos ARF8 clivados no transcriptoma do cafeeiro. O que confirma a ação de miR167 sobre o gene alvo predito *in silico*.

Para todos os clones de *Coffea arabica* e *Coffea canephora* sequenciados o sítio de clivagem de miR167 foi o mesmo. O corte do transcrito aconteceu no 11º resíduo do trecho de complementaridade entre miR167 e ARF8.

5 PERSPECTIVAS

Nossas análises *in silico* possibilitaram a predição de diversos genes alvo no banco de ESTs do café. A próxima etapa desta pesquisa será selecionar os alvos ou vias metabólicas mais interessantes e com sequenciamento do “degradoma” de RNAs confirmar a atuação dos miRNAs sobre tais genes. Vale ressaltar que este foi um trabalho preliminar, de caracterização, e que os resultados obtidos servirão de base para pesquisas futuras sobre o papel funcional dos genes MIR e seus alvos no cafeeiro.

Nestas futuras pesquisas, análises mais acuradas que avaliem, por exemplo, a capacidade de degradação dos transcritos, o tempo e a intensidade da resposta ao silenciamento e os possíveis padrões de expressão órgão/tecido específicos, serão propostas. Para tanto, materiais vegetais mais específicos e métodos complementares como análises metabolômicas, proteômicas e de RNA-seq deverão ser adotados para investigar os efeitos globais das respostas dos miRNAs aos diferentes estímulos ou condições ambientais. É importante que tenhamos de forma clara a medida quali-quantitativa da resposta geral da planta a mudanças no microtranscriptoma.

Com relação à ploidia, nossos resultados indicam que os miRNAs são expressos em *Coffea arabica* de forma aditiva, mostrando um valor de acumulação próximo do dobro daquele encontrado para cada uma das prováveis espécies genitoras. Usar indivíduos aloploidos imediatamente seguintes a um processo de hibridação (indivíduos da geração F1) tornaria os resultados mais informativos. Nessa condição seria possível monitorar como as vias moduladas por miRNAs interagem numa mesma célula que comporta dois ou mais genomas diferentes e como o controle dessas vias modelam a evolução fisiológica e morfológica da nova espécie.

Espera-se que a avaliação da atuação dos microRNAs aliada à caracterização dos principais fatores de transcrição, forneçam subsídio para

futuros trabalhos que visem a manipulação e o desenvolvimento da cultura do café e que colabore com conhecimento significativo e aplicável para a obtenção de genótipos melhorados, mais tolerantes, dessa *commodity* que tem lugar de grande destaque no agronegócio nacional.

REFERÊNCIAS

- ALLEN, E. et al. microRNA-directed phasing during trans-acting siRNA biogenesis in plants, **Cell**, Cambridge, v. 121, n. 2, p. 207–221, Apr. 2005.
- BARTEL, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. **Cell**, Cambridge, v. 116, n. 2, p. 281–297, Jan. 2004.
- CHAPMAN, E. J.; CARRINGTON, J. C. Specialization and evolution of endogenous small RNA pathways. **Nature Reviews Genetics**, London, v. 8, n. 11, p. 884–896, Nov. 2007.
- CHEN, C. et al. Real-time quantification of microRNAs by stem-loop RT-PCR. **Nucleic Acids Research**, Oxford, v. 179, p. 1-9, Nov. 2005.
- CHEN, Z. J. Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. **Annual Review of Plant Biology**, Palo Alto, v. 58, p. 377–406, June 2007.
- COMAI, L. Genetic and epigenetic interactions in allopolyploid plants. **Plant Molecular Biology**, Dordrecht, v. 43, n. 2-3, p. 387–399, 2000.
- DOYLE, J. J. et al. Evolutionary genetics of genome merger and doubling in plants. **Annual Review of Genetics**, Palo Alto, v. 42, p. 443-461, 2008.
- FLAGEL, L. et al. Duplicate gene expression in allopolyploid *Gossypium* reveals two temporally distinct phases of expression evolution. **BMC Biology**, London, v. 6, p. 6-16, Jan. 2008.
- HA, M. et al. Interspecies regulation of MicroRNAs and their targets. **Biochimica et Biophysica Acta**, Amsterdam, v. 1779, n. 11, p. 735–742, Nov. 2008.

HA, M. et al. Small RNAs serve as a genetic buffer against genomic shock in *Arabidopsis* interspecific hybrids and allopolyploids. **Proceedings of the National Academy of Sciences**, New York, v. 106, n. 42, p. 17835–17840, Oct. 2009.

HEGARTY, M. J. et al. Transcriptome shock after interspecific hybridization in senescence is ameliorated by genome duplication. **Current Biology**, London, v. 16, n. 16, p. 1652–1659, Aug. 2006.

KENAN-EICHLER, M. et al. Wheat Hybridization and Polyploidization Results in Derepression of small RNAs. **Genetics**, Austin, v. 188, n. 2, p. 263-272, June 2011.

LASHERMES, P. et al. Molecular characterization and origin of the *Coffea arabica* L. genome. **Molecular and General Genetics**, New York, 261, n. 2, p. 259-266, 1999.

LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta CT) Method. **Methods**, San Diego, v. 25, n. 4, p. 402–408, Dec. 2001.

MALLET, J. Hybrid speciation. **Nature**, London, v. 446, n. 7133, p. 279-283, Mar. 2007.

MALLORY, A. C.; BARTEL, D. P.; BARTEL, B. MicroRNA-directed regulation of *Arabidopsis* auxin response factor17 is essential for proper development and modulates expression of early auxin response genes. **The Plant Cell**, Rockville, v. 17, n. 5, p. 1360-1375, May 2005.

MALLORY, A. C.; VAUCHERET, H. ARGONAUTE 1 homeostasis invokes the coordinate action of the microRNA and siRNA pathways. **EMBO Reports**, Oxford, v. 10, n. 5, p. 521-526, May 2009.

NAGPAL, P. et al. Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. **Development**, Cambridge, v. 132, n. 18, p. 4107-4118, Sept. 2005.

PETITOT, A-S.; LECOULS, A-C.; FERNANDEZ, D. Sub-genomic origin and regulation patterns of a duplicated WRKY gene in the allotetraploid species *Coffea arabica*. **Tree Genetics & Genomes**, New York, v. 4, p. 379–390, 2008.

RAINA, S. N.; MUKAI, Y.; YAMAMOTO, M. In situ hybridization identifies the diploid progenitor species of *Coffea arabica* (Rubiaceae). **Theoretical and Applied Genetics**, Berlin, v. 97, p. 1204-1209, 1998.

SILVESTRINI, M. et al. Genetic diversity of a *Coffea* Germplasm Collection assessed by RAPD markers. **Genetic Resources and Crop Evolution**, Dordrecht, v. 55, n. 6, p. 901–910, 2008.

TIROSH, I. et al. A yeast hybrid provides insight into the evolution of gene expression regulation. **Science**, Washington, v. 324, n. 5938, p. 659-662, July 2009.

ULMASOV, T. et al. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. **The Plant Cell**, Rockville, v. 9, p. 1963-1971, Nov. 1997.

VEITIA, R. A.; BOTTANI, S.; BIRCHLER, J. A. Cellular reactions to gene dosage imbalance: genomic, transcriptomic and proteomic effects. **Trends Genetics**, v. 24, n. 8, p. 390–397, Aug. 2008.

WANG, J. et al. Genomewide nonadditive gene regulation in *Arabidopsis* allotetraploids. **Genetics**, Austin, v. 172, n. 1, p. 507–517, Sept. 2006.

WANG, J. W. et al. Control of root cap formation by MicroRNA-targeted auxin response factors in *Arabidopsis*. **The Plant Cell**, Rockville, v. 17, n. 8, p. 2204-2216, Aug. 2005.

WU, M. F.; TIAN, Q.; REED, J. W. Arabidopsis microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. **Development**, Cambridge, v. 133, n. 21, p. 4211-4218, Nov. 2006.

ZHANG, B. et al. Computational identification of microRNAs and their targets. **Computational Biology and Chemistry**, Oxford, v. 30, p. 395-407, 2006.

ANEXO A