KELLEN KAUANNE PIMENTA DE OLIVEIRA

***Coffea* sp. siRNAs:** FROM MICROSPOROGENESIS REGULATION TO WARMING TOLERANCE

LAVRAS-MG
2022

KELLEN KAUANNE PIMENTA DE OLIVEIRA

***Coffea* sp. siRNAs:** FROM MICROSPOROGENESIS REGULATION TO WARMING TOLERANCE

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

Prof. Antonio Chalfun-Junior, PhD.
Orientador

Dr. Raphael Ricon de Oliveira
Coorientador

Dra. Christiane Noronha Fernandes-Brum
Coorientadora

LAVRAS-MG
2022

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

de Oliveira, Kellen Kauanne Pimenta.

Coffea sp. siRNAs: from microsporogenesis regulation to warming tolerance / Kellen Kauanne Pimenta de Oliveira. - 2022.

68 p. : il.

Orientador(a): Kellen Kauanne Pimenta de Oliveira.

Coorientador(a): Raphael Ricon de Oliveira, Christiane Noronha Fernandes-Brum.

Tese (doutorado) - Universidade Federal de Lavras, 2022.

Bibliografia.

1. MicroRNAs. 2. Reproductive Development. 3. Regulation of Gene Expression. I. de Oliveira, Kellen Kauanne Pimenta. II. de Oliveira, Raphael Ricon. III. Fernandes-Brum, Christiane Noronha. IV. Título.

KELLEN KAUANNE PIMENTA DE OLIVEIRA

***Coffea* sp. siRNAs:** FROM MICROSPOROGENESIS REGULATION TO WARMING TOLERANCE

***Coffea* sp. siRNAs:** DA REGULAÇÃO DA MICROSPOROGÊNESE À TOLERÂNCIA AO AUMENTO DE TEMPERATURA

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

APROVADA em 23 de setembro de 2022.
Dra. Elisa Monteze Bicalho - UFLA
Dra. Thaís Barros Rodrigues - GreenLight Biosciences
Dr. José Domingos Cochicho Ramalho - Universidade de Lisboa
Dr. Fabio Tebaldi Silveira Nogueira - USP

Prof. Antonio Chalfun-Junior, PhD.
Orientador

Dr. Raphael Ricon de Oliveira
Coorientador

Dra. Christiane Noronha Fernandes-Brum
Coorientadora

LAVRAS – MG
2022

À minha mãe Sandra, por todo amor, carinho, apoio, exemplo de fé, garra e justiça que sempre me deu. Dedico

AGRADECIMENTOS

Agradeço à Deus por estar à frente de tudo, principalmente nos momentos mais difíceis dessa incrível caminhada.

À toda minha família, especialmente aos meus pais Waldemir e Sandra, e meus irmãos Khaius e Kayky. Obrigada por terem me dado uma base tão forte, por me fazerem acreditar em mim, por sempre me apoiarem e se orgulharem de mim.

À minha família mineira, Regina, Sérgio, Pâmela e Marcela por terem me acolhido com carinho como membro de sua família em sua casa em Lavras. Serei eternamente grata!

Ao meu companheiro de vida José Diogo, por compartilhar comigo os melhores e piores momentos dessa jornada, com muita paciência, amor e carinho.

Ao meu orientador, professor Dr. Antonio Chalfun Júnior, pela oportunidade e confiança depositada em mim.

Aos meus coorientadores Raphael Ricon de Oliveira e Christiane Noronha Fernandes-Brum pela coorientação e apoio.

Aos meus amigos Iasminy, Gabriel, Rafael, Jacqueline, Bruno, Lillian, Júlia, Robert e Dawyson pela amizade, apoio e companheirismo dentro e fora do laboratório.

Aos membros e ex-membros do Laboratório de Fisiologia Molecular de Plantas (LFMP) com os quais convivi durante todos esses anos. Agradeço pela ajuda durante as diversas análises deste trabalho, todos vocês me ajudaram a concluí-lo de alguma forma.

Ao Joel, por cuidar de nossas plantas com todo cuidado e atenção necessários.

À Universidade Federal de Lavras, especialmente ao Programa de Biotecnologia Vegetal e ao Setor de Fisiologia Vegetal pela oportunidade e estruturas fornecidas.

À Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e ao Instituto Nacional de Ciência e Tecnologia do Café (INCT Café) pelo suporte financeiro.

À banca examinadora pela gentileza em aceitar contribuir com este trabalho.

Muito obrigada!

“Nada na biologia faz sentido, exceto à luz da evolução” (Theodosius Dobzhansky).

RESUMO

Os siRNAs são pequenos RNAs que regulam a expressão gênica pós-transcricionalmente em diversos processos vegetais, desde a formação dos gametas até o desenvolvimento, atuando também em resposta a estresses bióticos e abióticos. Diante das mudanças climáticas esperadas para os próximos anos, investigamos neste trabalho, o papel de siRNAs na microsporogênese e tolerância ao aumento de temperatura utilizando espécies do gênero *Coffea* como modelo de plantas poliploides de importância socioeconômica. Essa tese foi estruturada da seguinte maneira: uma introdução e referencial teórico (Primeira parte) contendo as informações relevantes para compreensão dos dois artigos gerados (Segunda parte), finalizando com uma conclusão geral e perspectivas futuras. No primeiro artigo desta tese, relatamos siRNAs relacionados à microsporogênese de *C. arabica* e vias moleculares relacionadas à sua biossíntese. Identificamos duas famílias das proteínas mais importantes envolvidas neste processo, CaAGOs e CaDCLs, além de miRNAs canônicos e novos "triggers" para phasiRNAs reprodutivos de 21 e 24 nucleotídeos. Também identificamos a dinâmica temporal dos phasiRNAs nas fases de desenvolvimento dos botões florais na microsporogênese de *C. arabica*. No segundo artigo desta tese, identificamos um novo miRNA que regula um transportador de GABA, uma molécula relatada por atuar na regulação fisiológica de plantas frente ao estresse térmico e que tem seu acúmulo reduzido em cafeeiros sensíveis ao calor sob elevação de temperatura. Além disso, também relacionamos um miRNA conservado com o acúmulo de antocianina em cafeeiros sensíveis ao calor e à maior tolerância de plantas mutantes de Arabidopsis ao estresse térmico. Por fim, com estes resultados inéditos, aumentamos o conhecimento científico sobre siRNAs no cafeeiro que estão envolvidos em dois importantes processos biológicos, a reprodução e regulação da temperatura. Dessa forma, esperamos contribuir com o desenvolvimento de novas ferramentas que ajudem a regular os efeitos do aquecimento global em crops, como no desenvolvimento de produtos à base de miRNAs sintéticos que regulam moléculas da fisiologia do estresse, ou produtos que regulam a expressão destes miRNAs.

Palavras-chave: Antocianina. *Coffea arabica*. Desenvolvimento Reprodutivo. GABA. MicroRNAs. Microsporogênese. PhasiRNAs. Regulação da Expressão Gênica.

ABSTRACT

SiRNAs are small RNAs that regulate post-transcriptional gene expression in several plant processes, from gamete formation to development, also acting in response to biotic and abiotic stresses. Given the climate changes expected for the coming years, we investigated in this work the role of siRNAs in microsporogenesis and tolerance to temperature increase using species of the genus *Coffea* as a model of polyploid plants of socioeconomic importance. This thesis was structured as follows: an introduction and theoretical framework (First part) containing the relevant information for understanding the two papers generated (Second part), ending with a general conclusion and future perspectives. In the first paper of this thesis, we reported siRNAs related to the microsporogenesis of *C. arabica* and molecular pathways related to their biosynthesis. We identified two families of the most important proteins involved in this process, CaAGOs and CaDCLs, in addition to canonical miRNAs and new triggers for 21 and 24 nucleotides reproductive phasiRNAs. We also identified the temporal dynamics of phasiRNAs in the stages of flower bud development in *C. arabica* microsporogenesis. In the second paper of this thesis, we identified a new miRNA that regulates a GABA transporter, a molecule reported to act in the physiological regulation of plants in the face of heat stress and that has its accumulation reduced in heat-sensitive coffee plants under temperature elevation. In addition, we also related a conserved miRNA with the accumulation of anthocyanin in heat-sensitive coffee plants and the higher tolerance of Arabidopsis mutant plants to heat stress. Finally, with these unprecedented results, we have increased scientific knowledge about siRNAs in coffee plants that are involved in two important biological processes, reproduction and temperature regulation. In this way, we hope to contribute to the development of new tools that help regulate the effects of global warming on crops, such as the development of products based on synthetic miRNAs that regulate molecules of stress physiology, or products that regulate the expression of these miRNAs.

Keywords: Anthocyanin. *Coffea arabica.* GABA. MicroRNAs. Microsporogenesis. PhasiRNAs. Reproductive Development. Regulation of Gene Expression.

SUMÁRIO

PRIMEIRA PARTE **1**

1. INTRODUÇÃO **1**

2. REFERENCIAL TEÓRICO **3**

2.1. Cafeeiro: Origem e fenologia 3

2.2. Impactos das mudanças climáticas na cafeicultura 6

2.3. Aspectos fisiológicos do cafeeiro frente às mudanças climáticas 7

2.4. Como os componentes genéticos e moleculares contribuem para as respostas do cafeeiro às mudanças climáticas 8

2.5. siRNAs: Biogênese e mecanismo de ação 9

2.6. miRNAs na resposta ao estresse térmico 14

REFERÊNCIAS **16**

SEGUNDA PARTE - ARTIGOS **21**

ARTIGO 1: REPRODUCTIVE PHASIRNAS PATHWAY DURING THE *COFFEA ARABICA* MICROSPOROGENESIS **21**

ARTIGO 2: NOVEL AND CONSERVED MICRORNA ACT IN RESPONSE TO TEMPERATURE ELEVATION IN *COFFEA* REGULATING GABA AND ANTHOCYANIN. **45**

CONCLUSÕES E PERSPECTIVAS FUTURAS **68**

PRIMEIRA PARTE

1. INTRODUÇÃO

O estresse térmico afeta negativamente a distribuição e a produtividade de plantas importantes comercialmente em todo o mundo. Quando a temperatura aumenta 5°C ou mais acima do ideal, é experimentada como estresse térmico por todos os organismos vivos (GUAN et al., 2013). Com o cafeeiro não é diferente, no início da florada, temperaturas elevadas associadas a um intenso déficit hídrico provocam a morte dos tubos polínicos por desidratação, causando o abortamento das flores (CAMARGO; CAMARGO, 2001) ou abertura precoce dos botões florais (DAMATTA et al., 2019), o que acaba gerando perdas na produção. Além disso, a exposição contínua de *Coffea arabica* a temperaturas próximas a 30°C pode resultar não só no crescimento reprimido, mas também em anomalias como amarelecimento de folhas (DAMATTA; RAMALHO, 2006) e crescimento de tumores na base do caule (FRANCO, 1958).

Devido à importância econômica do cafeeiro, tem se desenvolvido cada vez mais estudos de manejo e desenvolvimento de variedades com características agronômicas importantes, tais como tolerância a estresses bióticos e abióticos e maior produtividade (TONIETTO et al., 2012). Dos vários objetos de estudo visando o melhoramento, destacam-se os mecanismos moleculares que atuam durante todo o desenvolvimento da planta. MicroRNAs (miRNAs) por exemplo, são uma extensa classe de pequenas moléculas de RNA endógenas que atuam como reguladores pós-transcricionais da expressão gênica, atuando em inúmeros processos de desenvolvimento (ZHANG, 2015). Estas moléculas foram descobertas inicialmente em *Caenorhabditis elegans* (LAU et al., 2001), mas já foram identificadas em diferentes espécies de animais e plantas (ARTEAGA-VÁSQUEZ et al., 2006).

Uma das desvantagens do melhoramento genético, no entanto, é que uma característica de uma nova variedade de culturas pode ser equilibrada pelo comprometimento de uma ou mais outras características benéficas. Uma das formas de se superar este problema é a utilização de moléculas reguladoras pós transcricionais específicas para um determinado processo, como é o caso dos siRNAs. Estes pequenos RNAs atuam na degradação ou bloqueio do mRNA alvo, regulando diferentes processos, como desenvolvimento, diferenciação, florescimento, metabolismo e defesa das plantas (TANG, J.; CHU, C., 2017; DAMATTA; RAMALHO, 2006; VARSHNEY et al., 2020; VOINNET, 2009).

Alguns estudos já foram realizados, nos quais microRNAs (miRNAs) responsivos ao estresse térmico por altas temperaturas foram caracterizados, como em *Populus tomentosa* (CHEN

et al., 2012), *Arabidopsis thaliana* (GUAN et al., 2013), *Oryza sativa* (SAILAJA et al., 2014), *Saccharina japônica* (LIU et al., 2015); *Solanum lycopersicum* (ZHOU et al., 2016), *Gossypium hirsutum* (WANG et al., 2016) e *Musa* spp. AAA (ZHU et al., 2019). Contudo, apesar de alguns miRNAs em comum, a lista de miRNAs com expressão alterada varia com o grau de estresse e entre as espécies. Entretanto, não existem trabalhos com cafeeiro relacionando o papel dos siRNAs no controle da temperatura, embora sejam descritas as vias relacionadas e em resposta ao aumento da temperatura de maneira separada (DE OLIVEIRA et al., 2020; NORONHA FERNANDES-BRUM et al., 2017).

Muitos fatores afetam a produção de café, dentre eles, destacam-se a influência da bienalidade negativa e a incidência de altas temperaturas, atrelada à escassez de chuvas em período importante do ciclo (CONAB, 2022; ICO, 2022). A longo prazo, mudanças climáticas combinadas a fenômenos como secas mais longas e imprevisíveis ou, em contrapartida, chuvas excessivas, ameaçam a sustentabilidade da produção agrícola em escala global (DAMATTA et al., 2010, BUNN et al., 2015). As projeções indicam que, no final deste século, o CO_2 atmosférico, um dos principais gases do efeito estufa, poderá aumentar entre 421 e 936 μL L^{-1} de ar, paralelamente à elevação da temperatura global de 0,3 a 1,7°C, no melhor cenário e de 2,6 a 4,8°C no cenário mais pessimista, em relação ao período de 1986 a 2005 (IPCC, 2013; IPCC, 2014).

Portanto, considerando esse cenário de mudanças climáticas e o atual estado da arte sobre o papel dos siRNAs no cafeeiro, faz-se necessário um estudo aprofundado sobre estas moléculas nesse gênero. Com isso, espera-se contribuir para o entendimento dos complexos processos biológicos que ocorrem nessas plantas, desde a sua reprodução, até a tolerância a altas temperaturas, para que se possa facilitar a manipulação de características agronômicas de interesse comercial de forma mais eficiente e precisa.

2. REFERENCIAL TEÓRICO

2.1. Cafeeiro: Origem e fenologia

O cafeeiro pertence à família Rubiaceae, gênero *Coffea* L., com 103 espécies descritas, sendo a grande maioria diploide e alogâmica, porém, apenas *C. arabica* (café Arábica), *C. canephora* (café Robusta/Conilon), *C. liberica* (café Libérica), e *C. dewevrei* (café Excelsa) são cultivados (DAVIS et al., 2006). Destes, apenas duas possuem importância econômica: *C. arabica* e *C. canephora*. No Brasil, 80% da produção total é de *C. arabica* (CONAB, 2022).

C. canephora é uma espécie diplóide, com 2n=22 cromossomos (CONAGIN, 1961; LASHERMES et al., 1999), alógama, originária de uma ampla região quente, úmida e de baixa altitude, do continente africano, que se estende da Guiné à República Democrática do Congo. *C. canephora* é amplamente cultivada nos continentes africano, americano e asiático, em locais de baixa altitude e temperaturas mais elevadas, com média anual entre 22°C e 26°C (MENDES et al., 2002).

Em regiões de temperaturas mais elevadas e com abundante umidade, plantas da espécie *C. canephora* podem atingir até 5 metros de altura. Geralmente são multicaules, mesmo em cultivos comerciais com desbrotas frequentes. É uma espécie altamente polimorfa, com extensa distribuição geográfica, exibindo grande capacidade adaptativa a variadas condições de ambiente (GUERREIRO-FILHO et al., 2008). Esta espécie foi introduzida no Brasil por volta de 1920, no Espírito Santo e suas variedades possuem grande variabilidade em relação às diferentes características agronômicas e morfológicas devido a própria origem da espécie (AGUIAR et al., 2005).

Já a espécie *C. arabica* teve sua origem nas regiões do sudoeste da Etiópia, sudeste do Sudão e norte do Quênia (DAVIS et al., 2006; GUERREIRO-FILHO et al., 2008), a partir da hibridação natural de gametas não reduzidos das espécies diplóides, *C. eugenioides* e *C. canephora* (LASHERMES et al., 1999; SAKIYAMA et al., 1999). Portanto, é uma espécie alotetraplóide, 2n=4x=44 e considerada autógama, com cerca de 10% de alogamia (CARVALHO; MÔNACO,1964).

Atualmente, a espécie *C. arabica* tem ampla dispersão, sendo cultivada em regiões de altitudes mais elevadas e temperaturas mais amenas, entre 18°C e 21°C, nos continentes americano e asiático, além de algumas regiões da África (GUERREIRO-FILHO et al., 2008). *C. arabica* chegou ao Brasil em 1727, com introdução de somente três plantas, sendo assim, sua base genética é bastante estreita e todas as cultivares disponíveis da espécie são provenientes das duas formas botânicas, Typica e Bourbon. Apesar disso, *C. arabica* apresenta cultivares com grande variabilidade devido a fatores como cruzamentos, mutações, práticas agronômicas e o ambiente (GUERREIRO-FILHO et al., 2008; ANTHONY et al., 2001).

Ao contrário da maioria das angiospermas, que emitem inflorescências e frutificam no mesmo ano fenológico, o café precisa de dois anos para completar seu ciclo (CAMARGO; CAMARGO, 2001). O primeiro ano se inicia com crescimento vegetativo caracterizado pelo desenvolvimento dos ramos plagiotrópicos a partir dos ramos ortotrópicos. A primeira fase, que compreende os meses de setembro a março, tem início a partir da formação das gemas axilares nos nós dos ramos plagiotrópicos primários (CAMARGO; CAMARGO, 2001; LIVRAMENTO, 2010; RENA E MAESTRI, 1986).

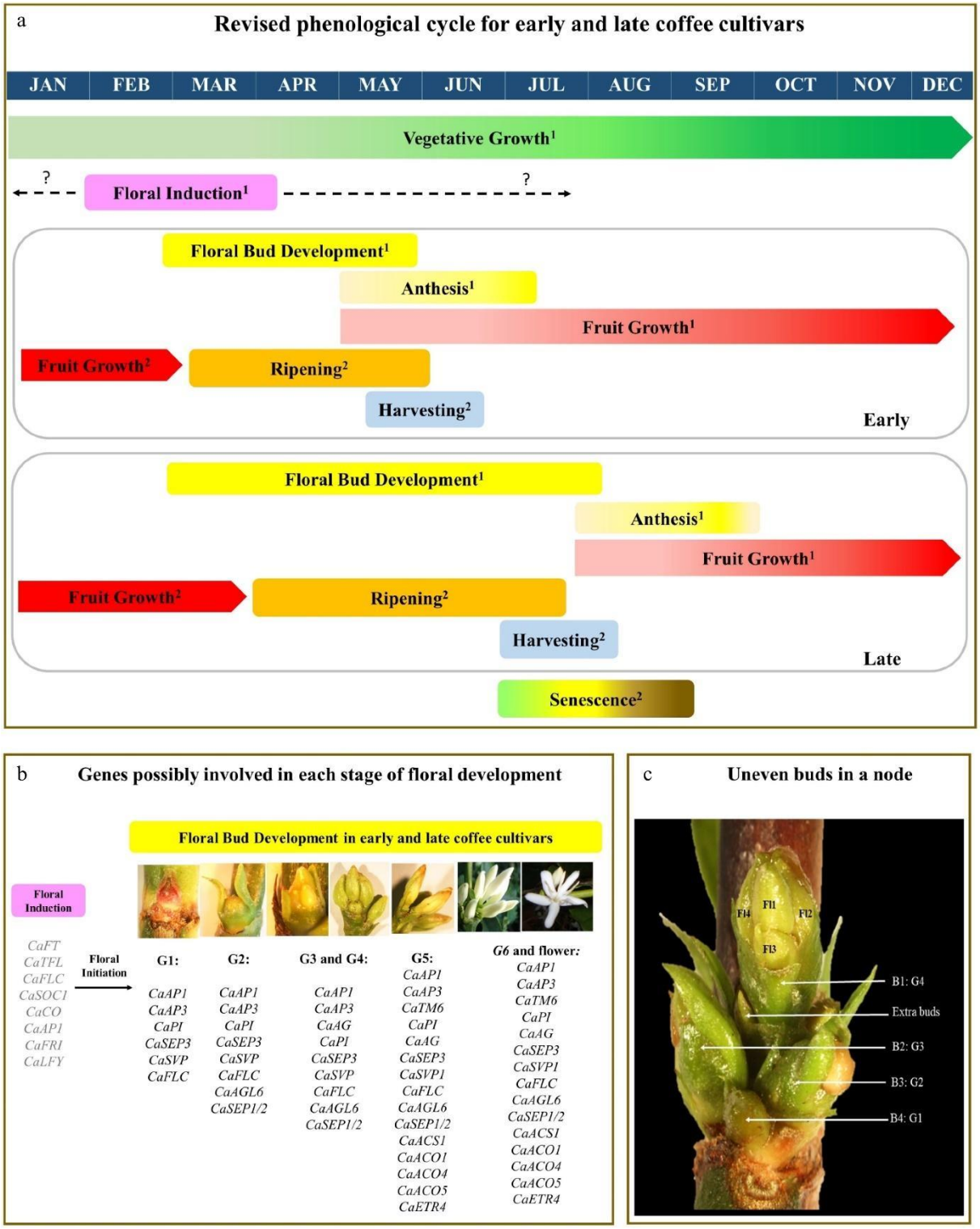
De janeiro a julho, começa a diferenciação das gemas florais entre cultivares precoces e tardias de *C. arabica* (LÓPEZ et al., 2021). Sabe-se que a qualidade e intensidade da luz, fotoperíodo ou duração do dia afetam o crescimento da planta, influenciando diretamente o desenvolvimento da flor e muitas outras características. No entanto, o estímulo indutivo do ciclo reprodutivo do cafeeiro ainda não está claro e pode ser afetado pela interação de diferentes fatores ambientais, nos quais os aspectos moleculares são pouco explorados (LÓPEZ et al., 2021). Recentemente, López et al., (2021) revisaram os fatores endógenos e ambientais relacionados ao processo de floração de *C. arabica*, e propuseram um modelo para a indução e desenvolvimento floral de *C. arabica* em condições ambientais brasileiras (Fig. 1).

Em cada axila da folha, é possível ocorre de quatro a cinco botões, que estão em diferentes estádios de desenvolvimento desde a sua origem (MAJEROWICZ; SONDAHL, 2005; DE OLIVEIRA et al., 2014. Os botões florais são classificados de acordo com seu tamanho, variando de G1 a G6 de acordo com Morais et al., (2008). Cada botão floral pode dar origem a quatro flores que terminam seu desenvolvimento com a antese (DE OLIVEIRA et al., 2014).

Quando as gemas florais atingem a maturação, momento em que todos os verticilos florais já estão formados (DE OLIVEIRA et al., 2014), entram em estádio de dormência, que corresponde aos meses de inverno, quando o potencial hídrico diminui (CAMARGO; CAMARGO, 2001). As gemas florais permanecem nesse estádio de latência até que ocorra o choque hídrico (CAMARGO; CAMARGO, 2001). Esse choque hídrico, causado por chuva ou irrigação é o principal fator para desencadear a florada, mas o aumento da umidade relativa do ar pode também iniciar este processo (CAMARGO; FRANCO, 1985).

O segundo ano fenológico inicia-se com a florada seguida pela formação dos chumbinhos, que precede a expansão dos grãos. Em seguida ocorre a granação dos frutos e a fase de maturação. Por fim, ocorre a senescência, morte dos ramos plagiotrópicos terminais (auto-poda). Na primavera do próximo ano, brotam novos ramos vegetativos, que se transformam em reprodutivos, permitindo nova produção (CAMARGO; CAMARGO, 2001).

Figura 1 – Indução e desenvolvimento floral de *C. arabica* em condições ambientais brasileiras.



Legenda: a) Modelo proposto para a indução e desenvolvimento floral de *C. arabica* em condições ambientais brasileiras. Modelo revisado para o desenvolvimento vegetativo e reprodutivo de cultivares de *C. arabica* quanto ao tempo de floração para cultivares precoces e tardias. b) Genes que estão possivelmente envolvidos em cada estádio do desenvolvimento floral. c) Foto descrevendo o desenvolvimento desigual de botões florais em um nó localizado no galho de um cafeeiro. Quatro botões florais (B1, B2, B3, B4) são apresentados em ordem crescente de emergência com a respectiva identificação dos estádios de desenvolvimento (G4, G3, G2, G1). No botão floral B1 estão indicadas as quatro flores que terminam seu desenvolvimento com a antese. Assim, cada nó pode potencialmente produzir 16 flores (quatro flores de cada botão), mas B3 e B4 geralmente permanecem latentes e, em alguns casos, botões florais extras podem ser desenvolvidos.

Fonte: López et al. (2021).

2.2. Impactos das mudanças climáticas na cafeicultura

O CO_2 é o maior contribuinte para o efeito estufa e consequentemente, para o aquecimento global. Sua concentração aumentou aproximadamente 43% desde os níveis pré-industriais e a temperatura média global aumentou 0,85°C no mesmo período (IPCC, 2013; IPCC, 2014). Essas mudanças à longo prazo, combinadas com a variabilidade climática, tais como secas mais longas e imprevisíveis, e chuvas excessivas, ameaçam a sustentabilidade da produção agrícola em escala global (DAMATTA et al., 2010; DAMATTA et al., 2019).

A seca e as temperaturas desfavoráveis sempre foram descritas como as principais limitações climáticas para a produção de café (CHESEREK; GICHIMU, 2012). Os principais efeitos das mudanças climáticas são o deslocamento de zonas de crescimento ótimas, mudanças na precipitação (quantidade e distribuição), mudanças na dinâmica de doenças/pragas das culturas e perda de terras agrícolas devido ao aumento do nível do mar e/ou desertificação (KIMEMIA, 2010). À medida que as mudanças no clima global são reconhecidas e o cultivo de café se espalha para diversas regiões onde antes não era cultivado, aumentam-se as preocupações com estes fatores e as possíveis consequências para a produção desta cultura (DA MATTA; RAMALHO 2006; KIMEMIA, 2010; CHESEREK; GICHIMU, 2012).

De fato, a temperatura pode limitar a exploração econômica do café, em parte porque o crescimento do café é particularmente afetado por temperaturas altas e baixas (BARROS et al., 1997; SILVA et al., 2004). A faixa média de temperatura ótima para *C. arabica* é relatada como sendo entre 18 e 21°C. Alguns estudos já relataram que temperaturas acima de 23°C podem acelerar o desenvolvimento e amadurecimento dos frutos, muitas vezes levando à perda de qualidade (DAMATTA; RAMALHO, 2006; CAMARGO; CAMARGO 2001) e a exposição contínua das plantas a temperaturas próximas a 30°C pode resultar em crescimento reprimido e anormalidades, como o amarelecimento de folhas (DAMATTA; RAMALHO, 2006) e crescimento de tumores na base do caule (FRANCO, 1958).

Entretanto, deve-se observar que cultivares selecionadas sob condições de manejo intensivo, permitiram que as plantações de *C. arabica* fossem espalhadas para regiões com temperaturas médias anuais de até 24 - 25°C, com rendimentos satisfatórios, como no Nordeste e no Norte do Brasil (FAZUOLI et al., 2007; BERGO et al., 2008). Por outro lado, em regiões com uma temperatura média anual abaixo de 18°C, o crescimento é amplamente reprimido (CAMARGO, 2009). De qualquer modo, está claro que grandes variações na temperatura causam defeitos nos grãos, modificam sua composição bioquímica e alteram a qualidade final da bebida (LOBELL; SCHLENKER; COSTA-ROBERTS, 2011; CARR, 2001; SILVA et al., 2005).

2.3. Aspectos fisiológicos do cafeeiro frente às mudanças climáticas

A sensibilidade do café a altas temperaturas, mais especificamente *C. arabica*, foi descrita em diversos estudos que avaliaram diversos aspectos desta cultura (DAMATTA et al., 2019), incluindo reduções extensivas no zoneamento agroclimático e perdas de áreas adequadas (ZULLO et al., 2011; BUNN et al., 2015; MAGRACH; GHAZOUL, 2015; OVALLE-RIVERA et al., 2015; MOAT et al., 2017), diminuição do rendimento das culturas (GAY et al., 2006; CRAPARO et al., 2015), qualidade da bebida (LÄDERACH et al. 2017), impactos negativos nas populações selvagens de *C. arabica* (DAVIS et al. 2012; MOAT et al., 2017), maior incidência de pragas (AVELINO et al., 2015; MAGRACH; GHAZOUL, 2015) e aumento das vulnerabilidades agrícolas, sociais e econômicas (BACA et al., 2014).

No entanto, os dados pessimistas obtidos até então sobre como as mudanças climáticas afetam a cultura do café, não haviam considerado os potenciais efeitos positivos da elevada concentração de CO_2 atmosférico na fotossíntese do cafeeiro ou o papel do CO_2 na tolerância ao calor e a resiliência reconhecida dos genótipos de café elite para se aclimatar a condições estressantes (DAMATTA et al., 2019).

No aspecto fisiológico, de modo geral, as plantas sentem e respondem diretamente ao aumento da concentração de CO_2 atmosférico por meio do aumento na taxa de fotossíntese líquida (A) e redução na condutância estomática (g_s), esta é a base para o efeito de fertilização por CO_2 com aumento correspondente nos rendimentos (LONG et al. 2006; AINSWORTH; ROGERS, 2007). Aumentos em A e na concentração de CO_2 no cloroplasto de plantas C3, estão associados a estimulação da taxa de carboxilação da ribulose-1,5-bisfosfato carboxilase/oxigenase (RuBisCO) (AINSWORTH; ROGERS, 2007).

Além disso, tem sido demonstrado sistematicamente, mas não universalmente, que g_s diminui com níveis elevados de concentração de CO_2 (ENGINEER et al., 2016). Em geral, diminuições em g_s frequentemente levam a taxas de transpiração mais baixas e maior eficiência no uso da água. No entanto, isso também resulta em perda de calor latente, aumentando as temperaturas das folhas, o que pode prejudicar o desempenho das culturas em um cenário de aquecimento global (DAMATTA et al., 2019).

Plantas de *Coffea arabica* cultivadas sob concentração de CO_2 elevada, foram acompanhadas por 3 anos e apresentaram maior rendimento de produção de grãos, do que plantas cultivadas sob concentração de CO_2 ambiente (GHINI et al., 2015). No entanto, os efeitos da combinação de elevadas temperaturas e concentrações de CO_2 sobre o rendimento da produção e a qualidade dos grãos devem ser melhor explorados.

Em um estudo com plantas de *Coffea arabica* L. cv. Icatu e IPR108; e *Coffea canephora* cv. Conilon CL153, cultivadas por 10 meses a 25/20°C (dia/noite) sob 380 e 700 µl de CO_2 L^{-1} e depois submetidas ao aumento de temperatura (0,5°C por dia) até 42/34°C (dia/noite), maiores concentrações de CO_2 contribuíram fortemente no impacto da temperatura em ambas as espécies, promovendo maior eficiência do uso da água (31/25°C), não provocando regulação fotossintética negativa, além de aumentar as moléculas de proteção, bem como a atividade de algumas enzimas antioxidantes. Elevadas concentrações de CO_2, portando, têm um papel na sustentabilidade da cultura sob cenários futuros de mudanças climáticas (RODRIGUES et al., 2016; MARTINS et al., 2016).

Todavia, várias questões cruciais ainda precisam ser totalmente elucidadas, como por exemplo, como a concentração elevada de CO_2 juntamente com o aumento da temperatura pode afetar o equilíbrio planta-peste/doenças, a formação anormal de flores e o desenvolvimento e qualidade dos frutos. Além disso, como o aumento de déficit de pressão de vapor (DPV) pode afetar o crescimento e a produtividade do café, e como isso pode ser mitigado pela concentração elevada de CO_2 ainda não foram esclarecidos (DAMATTA et al., 2019).

2.4. Como os componentes genéticos e moleculares contribuem para as respostas do cafeeiro às mudanças climáticas

No aspecto genético e molecular, as capacidades adaptativas e evolutivas de *C. arabica* e seus genitores (*C. eugenioides* e *C. canephora*) foram avaliadas sob quatro regimes térmicos (RTs; 18–14°C, 23–19°C, 28–24°C e 33–29°C). Os dados indicam que *C. arabica* aciona o genoma herdado de cada genitor para diferentes condições de temperatura. Por exemplo, a taxa de crescimento foi semelhante à de *C. canephora* sob o RT mais quente e à de *C. eugenioides* sob o mais frio RT. Para teores de metabólitos, *C. arabica* mostrou comportamento semelhante ao de *C. canephora* no RT mais quente. No nível de expressão gênica, poucas diferenças entre o alopoliploide e seus genitores foram observadas para genes estudados ligados à fotossíntese, respiração e ao relógio circadiano, enquanto genes ligados à atividade redox mostraram uma maior capacidade do alopoliplóide para a homeostase. Desta forma, a resposta transcricional global aos RTs de *C. arabica* era mais homeostática em comparação com seus pais e isso proporcionou maior homeostase fenotípica quando confrontada com ambientes inadequados para as espécies parentais diploides (BERTRAND et al., 2016).

Já um estudo recente, comparando genótipos de *C. arabica* (Acauã e Catuaí), explorando a fisiologia foliar, transcriptoma e composição de carboidratos/proteínas sob dois regimes térmicos (23/19°C-Opt e 30/26°C-Wat), demonstrou que a cultivar Acauã apresentou menores níveis de temperatura foliar nas duas condições em relação à cv. Catuaí, enquanto pouca ou nenhuma

diferença foi observada para os demais parâmetros fisiológicos foliares. No entanto, os genótipos apresentaram uma restrição transcricional na temperatura mais quente. Genes diferencialmente expressos responsivos a WaT revelaram genes compartilhados e específicos de cada genótipo, principalmente relacionados ao metabolismo de carboidratos. Esses achados revelaram diferenças intraespecíficas nas vias transcricionais e metabólicas interconectadas que respondem a temperaturas mais quentes, o que está potencialmente ligado à termotolerância no cafeeiro (DE OLIVEIRA et al., 2020).

2.5. siRNAs: Biogênese e mecanismo de ação

Grande parte das sequências transcritas de um genoma não codificam proteínas e são chamadas de RNAs não-codantes (ncRNAs). Dentre os RNAs não codantes, há um grupo de pequenos RNAs (sRNAs) que atuam no silenciamento mediado por RNA, controlando a expressão de muitos genes. Apesar dos sRNAs serem bioquimicamente muito próximos, eles diferem entre si do ponto de vista funcional e em sua biogênese, por exemplo, os miRNAs são transcritos a partir de genes MIR, enquanto siRNAs são produtos da clivagem de outros RNAs (Fig. 2) (VOINNET, 2009; ZHAO et al., 2016; WANG; CHEKANOVA, 2016).

Os miRNAs são uma extensa classe de pequenos RNAs endógenos, não codantes, que variam entre 20 e 24 nucleotídeos de comprimento e estão envolvidos na mediação e silenciamento da expressão gênica pós-transcricional. Apesar da inibição da tradução por clivagem ser um processo classicamente mediado por pequenos RNAs de interferência (siRNAs), miNAs também exercem esse papel (VOINNET, 2009; BUDAK; AKPINAR, 2015).

Os genes *MIR* geralmente codificam uma unidade de transcrição e são transcritos pela RNA polimerase II (POL II), enquanto os sRNAs são geralmente transcritos a partir da ação das polimerases IV e V exclusivas de plantas (LAW; JACOBSEN, 2010). Os transcritos miRNA primários (pri-miRNA) são 5'-capsulados e poliadenilados e são capazes de formar estruturas de dobra imperfeitas denominadas "*hairpins*", que possuem um braço 5p e um 3p, com tamanho semelhante aos transcritos que codificam proteínas (TANG, 2010; BUDAK; AKPINAR, 2015).

Embora os genes *MIR* frequentemente codificarem um transcrito de RNA único, existem alguns miRNAs policistrônicos documentados em plantas, um fenômeno que é muito mais comum em animais. A transcrição de alguns *loci* de miRNAs podem gerar transcritos com estruturas e sequências de "*hairpin*" diferentes, mas que ainda podem ser processadas pela maquinaria de biogênese dos miRNAs, resultando em sequências "maduras" diferentes (VOINNET, 2009; AXTELL et al., 2011, BUDAK; AKPINAR, 2015).

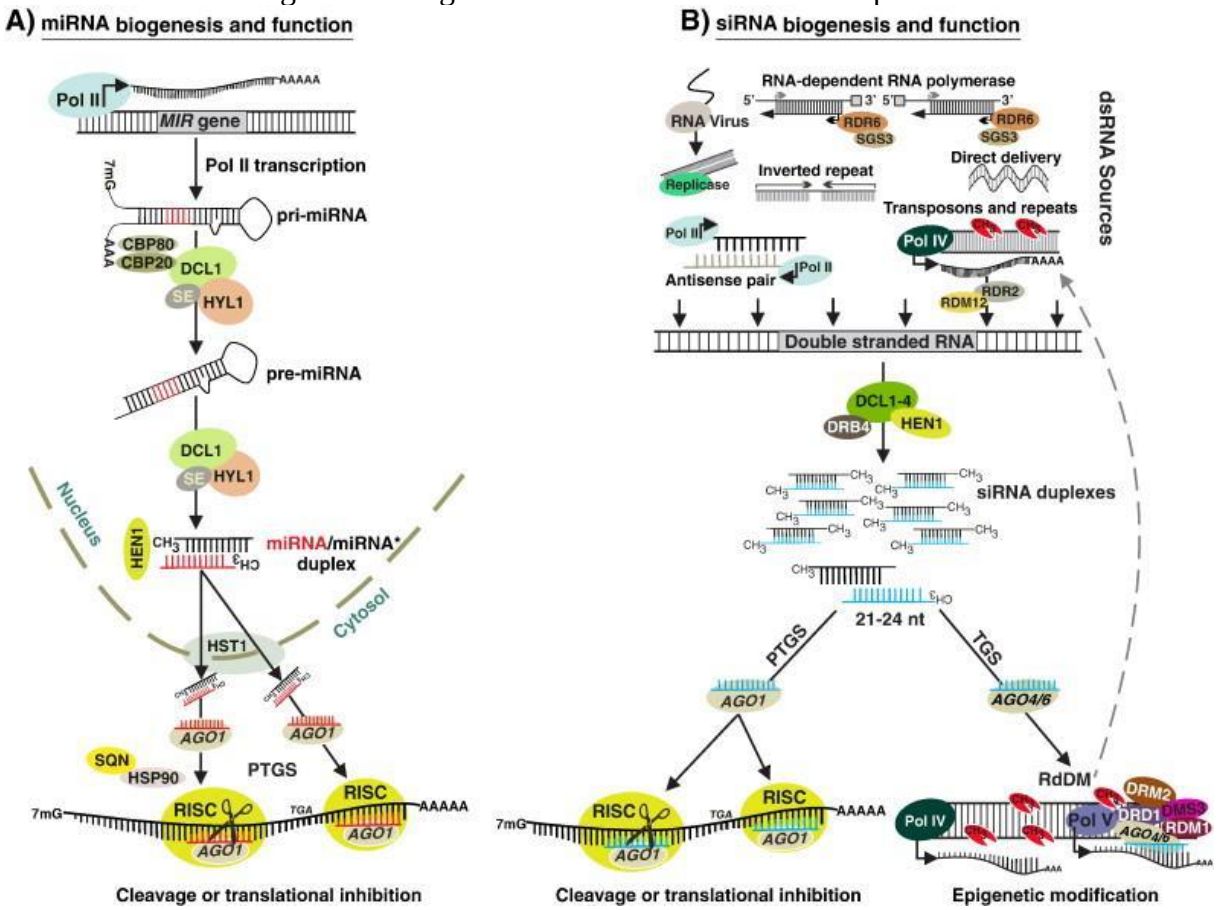
Em plantas, o transcrito primário (pri-miRNA) é estabilizado pelas proteínas de ligação DAWDLE (DDL), sendo processado ainda no núcleo pela atividade endonuclease de DICER-

LIKE1 (DCL1) com auxílio de outras proteínas, como as proteínas de ligação a RNA dupla fita (dsRNA) HYPONASTIC LEAVES 1 (HYL1) e SERRATE (SE) em precursores (pré-miRNAs) (DONG; HAN; FEDOROFF, 2008; KURIHARA et al., 2006). Em seguida, os pré-miRNAs sofrem uma clivagem subsequente pelo complexo DCL, formando uma estrutura de duplex (MARGIS et al., 2006).

O pré-miRNA em plantas varia de 49 a 900 nucleotídeos de comprimento e sofre processamento por DCL1, ou, alternativamente, por DCL2, DCL3 e DCL4, formando um duplex com dois nucleotídeos 3’ projetados lateralmente (MARGIS et al., 2006). Os miRNAs possuem em geral 21 nucleotídeos, mas o tamanho varia dependendo da DCL que faz a clivagem. A distância entre os domínios RNase III e PAZ é sugerida como sendo determinante no comprimento do miRNA (BOLOGNA; VOINNET, 2014; ROGERS; CHEN, 2013; BUDAK; AKPINAR, 2015). O duplex é então metilado na extremidadde 3’ por uma metiltransferase, denominada HUA ENCHANCER1 (HEN1), para evitar modificações posteriores e a degradação por exonucleases (RAMACHANDRAN; CHEN, 2008). A proteína HASTY (HST) realiza o transporte do duplex para o citoplasma (PARK et al., 2005), onde um dos braços do duplex é escolhido e incorporado a uma proteína da família ARGONAUTA (AGO) para formar o sistema RISC (*RNA Induced Silencing Complex*). Um dos domínios da proteína AGO, denominado PIWI possui atividade de endonuclease, capaz de clivar os alvos dos miRNAs (PARK et al., 2005; ROGERS; CHEN, 2013; CARTHEW; SONTHEIMER, 2009; DEBAT; DUCASSE, 2014).

Já os siRNAs são gerados a partir de RNAs fita dupla (dsRNA) que podem se originar de diferentes fontes, como RNAs transcritos de repetições invertidas, pares de transcrição cis - antisense naturais, a ação de RNA polimerases dependentes de RNA (RDRs) que convertem RNA de fita simples em dsRNA, a replicação de vírus de RNA e regiões do genoma ricas em retroelementos. O dsRNA é clivado em siRNAs de 21 a 24 nt por proteínas DCLs, e o tamanho dos siRNAs liberados depende da atividade catalítica específica da respectiva proteína DCL. O dsRNA é geralmente clivado por várias proteínas DCL, gerando classes de siRNA com diferentes tamanhos. Como os miRNAs, os siRNAs são carregados em RISC contendo proteína AGO que orienta a regulação do alvo no nível pós-transcricional ou no nível transcricional por meio de uma via denominada metilação do DNA dirigida por RNA (RdDM) (KHRAIWESH; ZHU; ZHU, 2012).

Figura 2 – Biogênese de miRNAs e siRNAs em plantas



Legenda: A) Os genes *MIR* são inicialmente transcritos por PolII em um RNA de fita simples que se dobra sobre si mesmo para formar uma estrutura em gancho (também chamada de pri-miRNA) que se acredita ser estabilizada pela proteína de ligação ao RNA DAWDLE (DDL). O *splicing* e o processamento adicional em corpos de corte nuclear envolvem as funções interativas de HYL1 e SE e das proteínas de ligação ao cap (CBP) CBP20 e CBP80. Pri-miRNAs e pré-miRNAs são geralmente pensados para serem processados a partir da extremidade livre oposta ao loop por DCL1 para produzir um ou vários duplexes de miRNA/miRNA* em fase. Estes são então metilados pelo HEN1 e transportados para o citoplasma pelo HST1. A fita guia de miRNA é selecionada, incorporada e estabilizada na proteína AGO1 dedicada. O RISC contendo AGO1 guiado por miRNA direciona a clivagem do mRNA ou a inibição da tradução do transcrito alvo. (B) Várias fontes de dsRNA, seu processamento em siRNAs por uma das quatro proteínas DCLs assistidas por proteínas de ligação a dsRNA, estabilização de siRNA mediada por HEN1 e fitas selecionadas de duplexes de siRNA guiam RISC contendo AGO para direcionar RNAs para clivagem endonucleolítica e para repressão de tradução, ou esses siRNAs podem então guiar AGO4 ou AGO6 para funcionar na via de metilação de DNA dirigida por RNA (RdDM) envolvendo PolIV e PolV. A metilação da citosina nestes locais específicos envolve diferentes efetores, como a metiltransferase de novo DOMÍNIOS METILTRANSFERASE 2 (DRM2); DEFEITUOSO NA METILAÇÃO DE DNA DIRIGIDA POR RNA 1 (DRD1), um membro da família de proteínas de remodelação da cromatina SWI2-SNF2; e DEFECTIVE IN MERISTEM SILENCING 3 (DMS3), uma proteína de manutenção estrutural de cromossomos (SMC).
Fonte: Khraiwesh; Zhu; Zhu (2012).

Pequenos RNAs interferentes secundários em fase (phasiRNAs) constituem uma categoria importante de pequenos RNAs em plantas, mas a maioria de suas funções ainda é pouco definida. A biogênese de phasiRNAs ocorre após a clivagem do mRNA ou RNA longo não-codificante (lncRNA) alvo, tipicamente (mas não exclusivamente) por um miRNA de 22 nucleotídeos (LIU et al., 2020). Após a clivagem, o fragmento 5' do mRNA alvo é rapidamente degradado por um

complexo exonucleolítico 3'→5'. O fragmento 3' é convertido em RNA de fita dupla (dsRNA) através da atividade de (RDR6), que pode ser recrutado por AGO1-RISC ou AGO7-RISC e assistido por SUPRESSOR DE SILENCIAMENTO DE GENES3 (SGS3), que por sua vez pode impedir a degradação do fragmento 3' de uma exoribonuclease 5'→3'. O dsRNA resultante é clivado iterativamente por uma proteína Dicer da extremidade 5' da fita "top" (derivada de Pol II) contendo o sítio de clivagem, produzindo duplexes de phasiRNAs (Fig. 3). A função do membro da família Dicer DCL4 requer a assistência de uma proteína DOUBLE-STRANDED RNA BINDING FACTOR (DRB) para produzir phasiRNAs de 21 nucleotídeos. Semelhante aos duplexes de miRNA, os duplexes de tasiRNA são classificados durante o carregamento em proteínas AGO por meio de um processo conhecido por ser dependente do nucleotídeo 5', mas esse processo não é bem descrito. O RISC contendo phasiRNA subsequentemente interage com RNAs alvo de uma maneira dependente de homologia, como com miRNAs (LIU et al., 2020).

Figura 3 – Biogênese de miRNAs e siRNAs em plantas

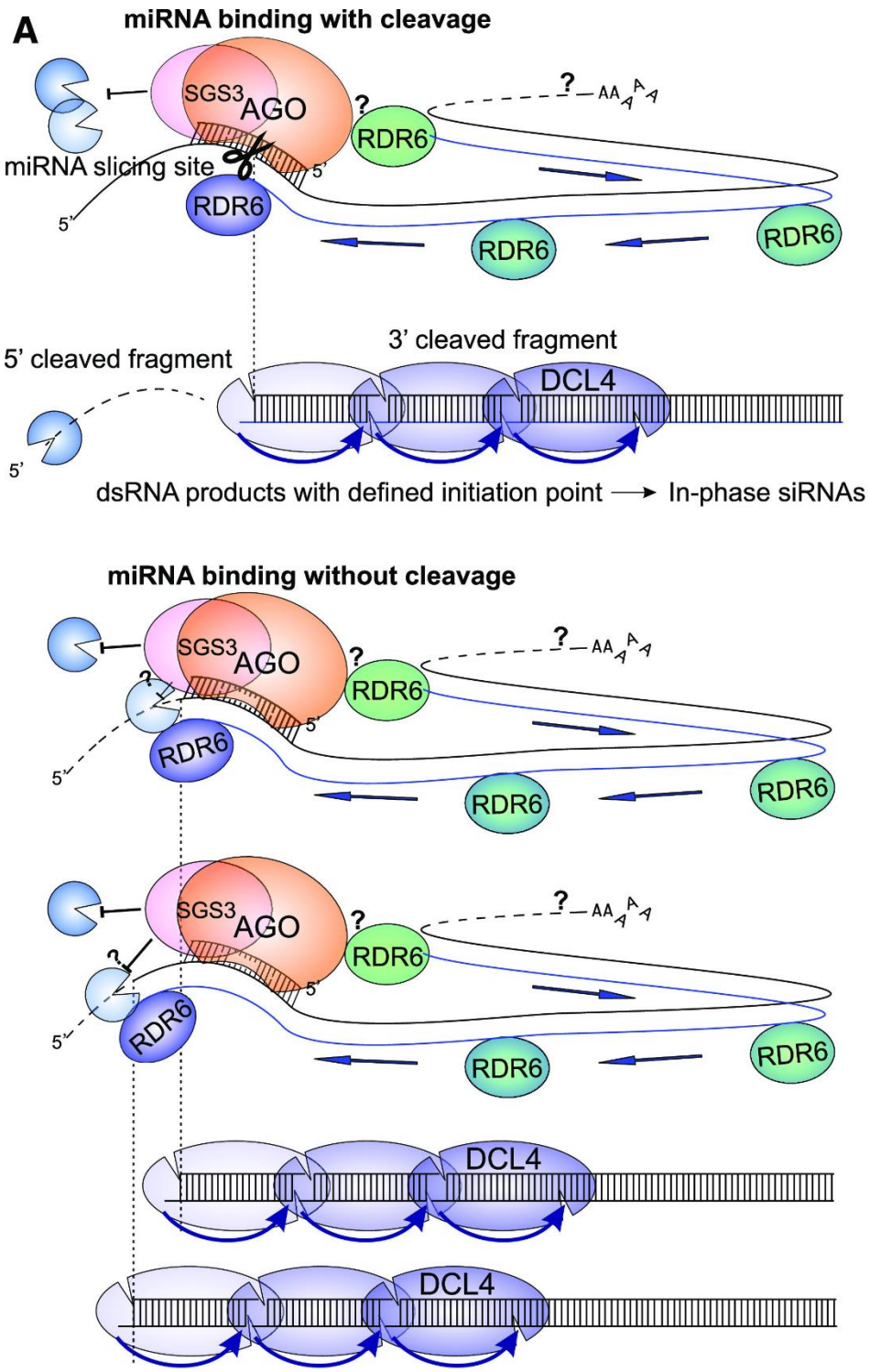


Fig. 3 - A clivagem é responsável por estabelecer o padrão de fase. Durante a biogênese secundária de siRNA mediada por miRNA, RDR6, recrutado por AGO (com a ajuda de SGS3), converte o substrato de mRNA em dsRNA, seguido de processamento por DCL4 ou DCL5. A clivagem pelo complexo miRNA-AGO em uma posição de nucleotídeo consistente marca o ponto de iniciação definido dos siRNAs resultantes, estabelecendo o padrão de fase. Por outro lado, a ausência de clivagem precisa no substrato de mRNA produz dsRNAs com pontos de iniciação indefinidos, produzindo siRNAs fora de fase. O "Pac-Man" azul-escuro representa o complexo SKI2-3-8 que direciona a exonucleólise 3′→5′, enquanto o Pac-Man azul-claro representa uma 5′→3′-exorribonuclease. O ponto de interrogação indica possível degradação.

Fonte:Liu et al. (2020).

2.6. miRNAs na resposta ao estresse térmico

Níveis elevados de CO_2 e temperatura podem afetar o crescimento e o desenvolvimento das plantas, mas as vias de sinalização que regulam esses processos e quais moléculas são reguladas pelos miRNAs nessas condições são pouco compreendidos (VOINNET, 2009; MAY et al., 2012).

Em Arabdopsis, por meio de sequenciamento de pequenos RNAs, foram identificados miRNAs cuja expressão é significativamente alterada quando a concentração de CO_2 atmosférico é dobrada ou a temperatura é aumentada de 3-6°C. Notavelmente, quase todos os miRNAs influenciados por CO_2 foram afetados inversamente pela temperatura elevada. Além disso, este estudo aponta que a via *miR156/157 - SPLs- miR172 -AP2*, relatados por atuarem na transição de fase vegetativa para reprodutiva, provavelmente tem um papel importante no florescimento precoce induzido por concentrações elevadas de CO_2 (MAY et al., 2013).

Em *Helianthus annuus* (girassol), foi estabelecida uma interação regulatória entre o fator de transcrição *HaWRKY6* e o miR396. Este miRNA que normalmente atua no desenvolvimento das plantas por meio da regulação de fatores de transcrição *GRF*, foi recrutado para proteção a altas temperaturas em girassol (45°C), uma planta particularmente bem adaptada a esse tipo de estresse. A identificação de um fator de transcrição jovem e divergente, sob o controle negativo de um elemento regulador altamente conservado, indica que é preciso cautela ao fazer inferências sobre o papel dos miRNAs conhecidos em plantas não-modelares (GIACOMELLI et al., 2012).

Um estudo em cevada demonstrou que os miRNAs podem funcionar em uma complexa rede regulatória desenvolvida para lidar com condições estressantes. Quatro miRNAs (miR160a, miR166a, miR167h e miR5175a) foram mais expressos sob estresse por aumento de temperatura (35,5°C) e o acúmulo destes miRNAs desencadeou mudanças no nível de expressão de genes alvo conservados e novos. Além disso, o *splicing* de introns do miR160a e do miR5175a também foi induzido pelo calor (KRUSZKA et al., 2014).

Em *Populus tomentosa*, bibliotecas de plântulas tratadas e não tratadas por estresse térmico (37°C) identificaram 52 miRNAs conservados sensíveis ao calor e 16 novos. Curiosamente, o total de *reads* de miRNAs conservados diminuiu sob estresse térmico e a maioria dos miRNAs conservados responsivos ao calor foram reprimidos (CHEN et al., 2012). Os miRNAs miR160, miR168 e miR169 foram responsivos ao estresse térmico nesse estudo, consistente com os dados obtidos em trigo submetidos a 40°C por 1 h. (XIN et al., 2010), sugerindo que esses miRNAs possuem um papel na reposta ao estresse térmico em diferentes plantas.

Em plantulas de *Brassica rapa* expostas a altas temperaturas (46°C por 1 h), foram identificadas cinco famílias de miRNA responsivas ao estresse térmico em bibliotecas de sRNA. Os miRNAs conservados bra-miR398a e bra-miR398b foram inibidos pelo calor, enquanto bra-

miR156h e bra-miR156g foram induzidos pelo calor e seu alvo putativo BracSPL2 teve sua expressão reduzida (YU et al, 2011).

Em plantas de *Solanum lycopersicum* L. (tomate selvagem) que receberam diferentes tratamentos térmicos (26/18°C, 33/33°C e 40/40°C), alguns miRNAs, por exemplo, miR166 e miR408, mostraram mudanças consistentes nos níveis de expressão nas temperaturas moderadamente e agudamente elevadas. Por outro lado, certos miRNAs, por exemplo o miR171, exibiram mudanças opostas a temperaturas moderadamente e agudamente elevadas, talvez porque certos miRNAs respondam diferentemente a diferentes temperaturas elevadas (ZHOU et al., 2016).

A metilação do DNA é uma importante modificação epigenética e desempenha um papel fundamental na regulação do crescimento e desenvolvimento das plantas. Em tabaco, verificou -se que a expressão de CycD3-1 e Nt-EXPA5 metilados estava alterada durante o estresse térmico (CENTOMANI et al., 2015). Em *Populus simonii*, os genes *MIR393a*, *MIR156i*, *MIR167h*, *MIR396e* e *MIR396g* foram metilados em regiões CNG em plantas tratadas termicamente por calor, enquanto foram metilados em regiões CG em sistemas tratados a frio. Estes dados demonstram que a metilação pode regular a expressão de miRNAs sob estresse térmico, afetando ainda mais o nível de expressão de seus alvos, provavelmente através da função de silenciamento gênico de miRNAs (CI et al., 2015).

REFERÊNCIAS

AGUIAR, A. T. E. et al. Diversidade química de cafeeiros na espécie *Coffea canephora*. **Bragantia**, v.64, n.4, p.577-582, 2005.

ARTEAGA-VÁSQUEZ, M.; CALLABERO-PEREZ, J.; VIELLE-CALZADA J. P. A family of miRNAs present in plants and animals. **The Plant Cell**, v. 18, n. 12 p. 3355-3369, 2006.

AVELINO, J. et al. The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions. **Food Security**, v. 7, n. 2, p. 303-321, 2015.

AXTELL, M. J.; WESTHOLM, J. O.; LAI, E. C. Vive la différence: biogenesis and evolution of microRNAs in plants and animals. **Genome biology**, v. 12, n. 4, p. 221, 2011.

BACA, M. et al. An integrated framework for assessing vulnerability to climate change and developing adaptation strategies for coffee growing families in Mesoamerica. **PloS one**, v. 9, n. 2, p. e88463, 2014.

BARROS, R. S. et al. Decline of vegetative growth in *Coffea arabica* L. in relation to leaf temperature, water potential and stomatal conductance. **Field Crops Research**, v. 54, n. 1, p. 65-72, 1997.

BERTRAND, B. et al. The greater phenotypic homeostasis of the allopolyploid *Coffea arabica* improved the transcriptional homeostasis over that of both diploid parents. **Plant and Cell Physiology**, v. 56, n. 10, p. 2035-2051, 2015.

BOLOGNA, N. G.; VOINNET, O. The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. **Annual review of plant biology**, v. 65, p. 473-503, 2014.

BUDAK, H.; AKPINAR, B. A. Plant miRNAs: biogenesis, organization and origins. **Functional & integrative genomics**, v. 15, n. 5, p. 523-531, 2015.

BUNN, C. et al. A bitter cup: climate change profile of global production of Arabica and Robusta coffee. **Climatic Change**, v. 129, n. 1-2, p. 89-101, 2015.

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

CAMARGO, A. P.; FRANCO, C. F. **Clima e fenologia do cafeeiro. In: Cultura de café no Brasil: manual de recomendações**. 5.ed. Rio de Janeiro: Instituto Brasileiro do Café, Ministério da Indústria e Comércio, p.1950, 1985.

CARR, M. K. V. The water relations and irrigation requirements of coffee. **Experimental Agriculture**, v. 37, n. 1, p. 1-36, 2001.

CARTHEW, R. W.; SONTHEIMER, E. J. Origins and mechanisms of miRNAs and siRNAs. **Cell**, v. 136, n. 4, p. 642-655, 2009.

CARVALHO, A. et al. Aspectos genéticos do cafeeiro. **Revista Brasileira de Genética**, v.14, n.1, p.135-183, 1991.

CENTOMANI, I. et al. Involvement of DNA methylation in the control of cell growth during heat stress in tobacco BY-2 cells. **Protoplasma**, v. 252, n. 6, p. 1451-1459, 2015.

CHEN, L. et al. Genome-wide identification and expression analysis of heat-responsive and novel microRNAs in *Populus tomentosa*. **Gene**, v. 504, n. 2, p. 160-165, 2012.

CHESEREK, J. J.; GICHIMU, B. M. Drought and heat tolerance in coffee: a review. **International Research Journal of Agricultural Science and Soil Science**, v. 2, n. 12, p. 498-501, 2012.

CI, D. et al. Methylation of miRNA genes in the response to temperature stress in *Populus simonii*. **Frontiers in plant science**, v. 6, p. 921, 2015.

CONAB: COMPANHIA NACIONAL DE ABASTECIMENTO. **4º Levantamento de Café - Safra 2021**. Disponível em: <https://www.conab.gov.br/info-agro/safras/cafe/boletim-da-safra-de-cafe/item/17268-4-levantamento-de-cafe-safra-2021> Acesso em: 06/08/2022.

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

CRAPARO, A. C. W. et al. *Coffea arabica* yields decline in Tanzania due to climate change: Global implications. **Agricultural and Forest Meteorology**, v. 207, p. 1-10, 2015.

DAMATTA, F. M. et al. Impacts of climate changes on crop physiology and food quality. **Food Research International**, v. 43, n. 7, p. 1814-1823, 2010.

DAMATTA, F. M. et al. Why could the coffee crop endure climate change and global warming to a greater extent than previously estimated? **Climatic change**, v. 152, n. 1, p. 167-178, 2019.

DAMATTA, F. M.; RAMALHO, J. D. C. Impacts of drought and temperature stress on coffee physiology and production: a review. **Brazilian Journal of Plant Physiology**, v. 18, n. 1, p. 55-81, 2006.

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

DAVIS, A. P. et al. The impact of climate change on indigenous arabica coffee (*Coffea arabica*): predicting future trends and identifying priorities. **PloS one**, v. 7, n. 11, p. e47981, 2012.

DEBAT, H. J.; DUCASSE, D. A. Plant microRNAs: Recent Advances and Future Challenges. **Plant Molecular Biology Reporter**, v. 32, n. 6, p. 1257–1269, 1 dez. 2014.

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

DE OLIVEIRA, R. R. et al. Elevated Temperatures Impose Transcriptional Constraints and Elicit Intraspecific Differences Between Coffee Genotypes. **Frontiers in Plant Science**, v. 11, 2020.

DONG, Z.; HAN, M.; FEDOROFF, N. The RNA-binding proteins HYL1 and SE promote accurate in vitro processing of pri-miRNA by DCL1. **Proceedings of the National Academy of Sciences**, v. 105, n. 29, p. 9970-9975, 2008.

ENGINEER, C. B. et al. CO2 sensing and CO2 regulation of stomatal conductance: advances and open questions. **Trends in plant science**, v. 21, n. 1, p. 16-30, 2016.

FRANCO, C. M. Influence of temperature on growth of coffee plant. **New York: IBEC Research Institute**, 1958.

GAY, C. et al. Potential impacts of climate change on agriculture: a case of study of coffee production in Veracruz, Mexico. **Climatic Change**, v. 79, n. 3-4, p. 259-288, 2006.

GIACOMELLI, J. I. et al. Role of recently evolved miRNA regulation of sunflower HaWRKY6 in response to temperature damage. **New Phytologist**, v. 195, n. 4, p. 766-773, 2012.

GUAN, Q. et al. Heat stress induction of mi R 398 triggers a regulatory loop that is critical for thermotolerance in Arabidopsis. **The Plant Journal**, v. 74, n. 5, p. 840-851, 2013.

GUERREIRO-FILHO, O. et al. **Origem e classificação botânica do Cafeeiro**. In: CARVALHO, C. H. S. (Ed.). Cultivares de café: origem, características e recomendações. Cultivares de café: origem, características e recomendações. Brasília: Embrapa Café, p. 23-30, 2008.

GHINI, R. et al. Coffee growth, pest and yield responses to free-air CO2 enrichment. **Climatic Change**, v. 132, n. 2, p. 307–320, 1 set. 2015.

ICO: INTERNATIONAL COFEE ORGANIZATION. **Estatísticas do comércio**. Disponível em: < https://www.ico.org/pt/trade_statisticsp.asp> Acesso em: 06/08/2022.

IPCC. **Climate change 2013**. **The physical science basis**. Cambridge: Cambridge University press, Cambrigde, UK, 2013.

IPCC. **Proceedings of the 5th assessment report, WGII, climate change 2014: impacts, adaptation, and vulnerability**. Cambridge University Press, Cambridge, UK, 2014.

KHRAIWESH, B.; ZHU, J.-K.; ZHU, J. Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. **Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms**, Plant gene regulation in response to abiotic stress. v. 1819, n. 2, p. 137–148, 1 fev. 2012.

KIMEMIA J. K. **Effect of global warming on coffee production**. In: Ugandan Coffee Traders Federation Breakfast Fellowship, Kampala Uganda, 2010.

KRUSZKA, K. et al. Transcriptionally and post-transcriptionally regulated microRNAs in heat stress response in barley. **Journal of experimental botany**, v. 65, n. 20, p. 6123-6135, 2014.

KURIHARA, Y. et al. The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. **Rna**, v.12, n. 2, p. 206-212, 2006.

LÄDERACH, P. et al. Climate change adaptation of coffee production in space and time. **Climatic change**, v. 141, n. 1, p. 47-62, 2017.

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

LAU, N. C.; LIM, L. P.; WEINSTEIN E. G.; BARTEL, D. P. An abundant class of tiny RNAs with probable regulatory roles in *Caernohabditis elegans*. **Science**, v. 294, n. 5543 p. 858-862, 2001.

LAW, J. A.; JACOBSEN, S. E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. **Nature Reviews Genetics**, v. 11, n. 3, p. 204, 2010.

LIU, F. et al. Conserved and novel heat stress-responsive micro RNAs were identified by deep sequencing in *Saccharina japonica* (Laminariales, Phaeophyta). **Plant, cell & environment**, v. 38, n. 7, p. 1357-1367, 2015.

LIU, Y. et al. PhasiRNAs in Plants: Their Biogenesis, Genic Sources, and Roles in Stress Responses, Development, and Reproduction. **The Plant Cell**, v. 32, n. 10, p. 3059–3080, out. 2020.

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

LOBELL, D. B.; SCHLENKER, W.; COSTA-ROBERTS, J. Climate Trends and Global Crop Production Since 1980. **Science**, v. 333, n. 6042, p. 616–620, 29 jul. 2011.

LONG, S. P. et al. Food for thought: lower-than-expected crop yield stimulation with rising CO2 concentrations. **Science**, v. 312, n. 5782, p. 1918-1921, 2006.

LÓPEZ, M. E. et al. An overview of the endogenous and environmental factors related to the *Coffea arabica* flowering process. **Beverage Plant Research**, v. 1, n. 1, p. 1–16, 18 nov. 2021.

MAGRACH, A.; GHAZOUL, J. Climate and pest-driven geographic shifts in global coffee production: implications for forest cover, biodiversity and carbon storage. **PloS one**, v. 10, n. 7, p. e0133071, 2015.

MAJEROWICZ, N.; SÖNDAHL, M. R. Induction and differentiation of reproductive buds in *Coffea arabica* L. **Brazilian Journal of Plant Physiology**, v. 17, n. 2, p. 247-254, 2005.

MARTINS, M. Q. et al. Protective response mechanisms to heat stress in interaction with high [CO_2] conditions in Coffea spp. **Frontiers in Plant Science**, v. 7, p. 947, 2016.

MAY, P. et al. The effects of carbon dioxide and temperature on microRNA expression in Arabidopsis development. **Nature communications**, v. 4, p. 2145, 2013.

MENDES, A. N. G.; GUIMARÃES, R.J.; SOUZA, C.A.S. **Classificação botânica, origem e distribuição geográfica do cafeeiro**. In: GUIMARÃES, R.J. et al. (Ed.). Cafeicultura. Lavras: UFLA/FAEPE, p.39-99, 2002.

MOAT, J. et al. Resilience potential of the Ethiopian coffee sector under climate change. **Nature plants**, v. 3, n. 7, p. 17081, 2017.

MORAIS, H. et al. Escala fenológica detalhada da fase reprodutiva de Cooffea arabica. **Bragantia**, v. 67, p. 257–260, 2008.

NORONHA FERNANDES-BRUM, C. et al. A genome-wide analysis of the RNA-guided silencing pathway in coffee reveals insights into its regulatory mechanisms. **PLoS ONE**, v. 12, n. 4, p. e0176333, 27 abr. 2017.

OVALLE-RIVERA, O. et al. Projected shifts in *Coffea arabica* suitability among major global producing regions due to climate change. **PloS one**, v. 10, n. 4, p. e0124155, 2015.

PARK, M. Y. et al. Nuclear processing and export of microRNAs in Arabidopsis. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 10, p. 3691-3696, 2005.

RAMACHANDRAN, V.; CHEN, X. Degradation of microRNAs by a family of exoribonucleases in Arabidopsis. **Science**, v. 321, n. 5895, p. 1490-1492, 2008.

RENA, A. B.; MAESTRI, M. **Fisiologia do cafeeiro**. In: RENA, A. B. et al. (Ed.). Cultura do cafeeiro-Fatores que afetam a produtividade. Piracicaba: Associação Brasileira para Pesquisa da Potassa e do Fosfato, p.13-106, 1986.

RODRIGUES, W. P. et al. Long-term elevated air [CO_2] strengthens photosynthetic functioning and mitigates the impact of supra-optimal temperatures in tropical *Coffea arabica* and *C. canephora* species. **Global Change Biology**, v. 22, n. 1, p. 415-431, 2016.

ROGERS, K.; CHEN, X. Biogenesis, Turnover, and Mode of Action of Plant MicroRNAs. **Plant Cell**, v. 25, n. 7, p. 2383-2399, 2013.

SAILAJA, B. et al. Prediction and expression analysis of miRNAs associated with heat stress in *Oryza sativa*. 2013. **Rice Science**, v. 21, Issue 1, p. 3-12, 2014.

SAKIYAMA, N. S.; PEREIRA, A. A.; ZAMBOLIM, L. Melhoramento de café arábica. In: BORÉM, A. (Ed.). **Melhoramento de espécies cultivadas**. Viçosa: UFV, p. 189-204, 1999.

SILVA, E. A. et al. Seasonal changes in vegetative growth and photosynthesis of Arabica coffee trees. **Field Crops Research**, v. 89, n. 2-3, p. 349-357, 2004.

SILVA, E. A. et al. The influence of water management and environmental conditions on the chemical composition and beverage quality of coffee beans. **Brazilian Journal of Plant Physiology**, v. 17, n. 2, p. 229-238, 2005.

TANG, G. Plant microRNAs: An insight into their gene structures and evolution. **Seminars in Cell & Developmental Biology**, v. 21, n. 8, p. 782-789, 2010.

TANG, J.; CHU, C. MicroRNAs in crop improvement: fine-tuners for complex traits. **Nature plants**, v. 3, n. 7, p. 17077, 2017.

TONIETTO, Â, et al. Proteomic analysis of developing somatic embryos of *Coffea arabica*. **Plant Molecular Biology Reporter**, v. 30, n. 6, p. 1393-1399, 2012.

VARSHNEY, R. K. et al. 5Gs for crop genetic improvement. **Current Opinion in Plant Biology**, Biotic interactions ● AGRI 2019. v. 56, p. 190–196, 1 ago. 2020.

VOINNET, O. Origin, biogenesis, and activity of plant microRNAs. **Cell**, v. 136, n. 4, p. 669-687, 2009.

WANG, H. L. V.; CHEKANOVA, J. A. Small RNAs: essential regulators of gene expression and defenses against environmental stresses in plants. **Wiley Interdisciplinary Reviews-Rna**, v. 7, n. 3, p. 356-381, 2016.

WANG, Q. et al. Small RNA-mediated responses to low-and high-temperature stresses in cotton. **Scientific reports**, v. 6, p. 35558, 2016.

XIN, M. et al. Diverse set of microRNAs are responsive to powdery mildew infection and heat stress in wheat (*Triticum aestivum* L.). BMC plant biology, v. 10, n. 1, p. 123, 2010.

YU, X. et al. Identification of conserved and novel microRNAs that are responsive to heat stress in Brassica rapa. **Journal of Experimental Botany**, v. 63, n. 2, p. 1025-1038, 2011.

ZHANG, B. miRNA: a new target for improving plant tolerance to abiotic stress. **Journal of Experimental Botany**, 66, 1749–1761, 2015.

ZHAO, Y. et al. NONCODE 2016: an informative and valuable data source of long non-coding RNAs. **Nucleic Acids Research**, v. 44, n. D1, p. D203D208, 2016.

ZHOU, R. et al. Identification of miRNAs and their targets in wild tomato at moderately and acutely elevated temperatures by high-throughput sequencing and degradome analysis. **Scientific reports**, v. 6, p. 33777, 2016.

ZHU, H. et al. Banana sRNAome and degradome identify microRNAs functioning in differential responses to temperature stress. **BMC genomics**, v. 20, n. 1, p. 33, 2019.

ZULLO, J. et al. Potential for growing Arabica coffee in the extreme south of Brazil in a warmer world. **Climatic Change**, v. 109, n. 3-4, p. 535-548, 2011.

SEGUNDA PARTE - ARTIGOS

ARTIGO 1: REPRODUCTIVE PHASIRNAS PATHWAY DURING THE *COFFEA ARABICA* MICROSPOROGENESIS

TITLE

Reproductive phasiRNAs pathway during the *Coffea arabica* microsporogenesis

AUTHORS

Kellen Kauanne Pimenta de Oliveira[1], Raphael Ricon de Oliveira[1], Gabriel de Campos Rume[1], Thales Henrique Cherubino Ribeiro[1], Christiane Noronha Fernandes-Brum[1], Atul Kakrana[2], Sandra Mathioni[2], Blake C. Meyers[2], Antonio Chalfun-Júnior[1*].

AFFILIATIONS

[1]Laboratory of Plant Molecular Physiology, Plant Physiology Sector, Department of Biology, Federal University of Lavras, Lavras-MG, Brazil.

[2] Donald Danforth Plant Science Center, St. Louis, MO, 63132, USA; Department of Plant and Soil Science, University of Delaware, Newark, DE, 19711, USA.

CORRESPONDING AUTHOR

Antonio Chalfun-Júnior

Email: chalfunjunior@ufla.br

Laboratory of Plant Molecular Physiology, Plant Physiology Sector, Department of Biology, Federal University of Lavras, Lavras MG, Brazil.

ABSTRACT

Reproductive phasiRNAs are a class of siRNAs that, in plants, act in the protection of the genome against mobile genetic elements, in disease resistance, and in microsporogenesis, however much of their biogenesis and mode of action are still not fully described. Recently, the presence of these molecules in eudicots has attracted attention, in addition to the proteins and miRNA triggers that are part of their processing. In this work, we identified and quantified reproductive phasiRNAs and their canonical and novel miRNA triggers in coffee anthers in pre-meiosis, meiosis, and post-meiosis of *C. arabica*. Furthermore, we identified the members of DICER and ARGONAUTE families in the coffee genome, and determined their expression profiles using RNA-seq libraries, which suggested they might be involved in the processing of reproductive phasiRNAs. We found that the pattern of abundance of reproductive phasiRNAs from *C. arabica* is unique among described eudicots and the canonical trigger car-miR2275 is involved in the processing of both 21 and 24 nt phasiRNAs. Our results increase knowledge about the reproductive phasiRNAs pathway in coffee plants, contributing to advances in studies on fertility or sterility in eudicots of economic interest.

KEYWORDS

Coffee tree; Eudicots; Epigenetics; MicroRNA (miRNA); Phased small interfering RNAs (phasiRNAs); Reproductive development.

INTRODUCTION

Phased RNAs (phasiRNAs), small RNAs generated from a long RNA precursor at intervals of 21 to 26 nucleotides (nt) (AXTELL, 2013a; KOMIYA, 2017), are found in plants and animals and have been highlighted as a relevant group in reproductive regulation (XIA et al., 2019; ZHAI et al., 2015). These phased RNAs are classified into two groups, those originating from the PHAS loci of coding and protein-coding regions (LIU et al., 2020). Systematic analysis of the main phylogenetic groups of plants, encompassing algae, mosses, gymnosperms, basal angiosperms, monocots, and eudicots, identify abundant PHAS loci among protein-coding genes, suggesting that phasiRNAs predominantly regulate these genes from which they are supplied, in contrast to the dominant trans-regulatory mode of microRNAs (miRNAs). However, there are also TAS loci, which encode long non-coding RNAs (lncRNAs) that produce trans-acting siRNAs (tasiRNAs) (FEI; XIA; MEYERS, 2013; XIA et al., 2019; ZHENG et al., 2015).

In monocots, two pathways associated with germlines are well described and produce diverse and abundant reproductive phasiRNAs (FEI et al., 2016; ZHAI et al., 2015). These phasiRNAs are produced from capped and polyadenylated non-coding precursors ("PHAS" transcripts) and transcribed by RNA polymerase II (Pol II) from non-repetitive loci. They are typically triggered by two 22-nt miRNAs, miR2118/482 for 21 nt phasiRNAs and miR2275 for 24-nt phasiRNAs. 3' mRNA fragments are converted to double-stranded RNA by RNA-DEPENDENT RNA POLYMERASE 6 (RDR6) and processed by Dicer-like 4 (DCL4) and Dicer-like 5 (DCL5) to produce 21-nt and 24-nt phasiRNAs, respectively. It is worth mentioning that DCL5 is a specific dicer of monocots and so far, has not been described in eudicots. Finally, phasiRNAs are loaded into ARGONAUTE (AGO) proteins to induce target RNA silencing (KAKRANA et al., 2018; KOMIYA, 2017; XIA et al., 2019).

Reproductive phasiRNAs have a spatiotemporal pattern in monocots. Hundreds of loci on all chromosomes produce 21-nt phasiRNAs abundant in pre-meiotic anthers and 24 nt-phasiRNAs enriched in meiotic stage anthers (FAN et al., 2016; KAKRANA et al., 2018; ZHAI et al., 2015). Disturbance of 21-PHAS loci underlies photoperiod-sensitive male sterility in rice, while mutants that do not produce 24-nt phasiRNA produce conditional male sterility in maize (ZHAI et al., 2015).

In-depth comparative analyzes show us that, contrary to what was thought, the 24-nt phasiRNA pathway is also widely present in eudicot plants. However, it is absent in Arabidopsis

history, demonstrating the evolutionary dynamics of this pathway (XIA et al., 2019). Evolutionary studies with nongrass monocots (garden asparagus, daylily, and lily) have suggested that the DCL5-dependent meiotic 24-nt phasiRNA pathway may have originated more than 117 MYA, which is when common ancestors of these species diverged (KAKRANA et al., 2018). In addition, reproductive phasiRNAs are present in female reproductive organs and therefore may function in both male and female germinal development. (KAKRANA et al., 2018). Despite advances, the evolutionary origins of reproductive phasiRNAs and their 22 nt miRNA/DCL triggers are still poorly understood in dicots. In this study, we used *Coffea arabica*, an amphiploid of the Rubiaceae family, as a model for woody species, to increase the knowledge of the role of reproductive phasiRNAs in eudicot microsporogenesis.

MATERIALS AND METHODS

Identification of microsporogenesis stages in anthers of coffee flower buds.

Flower buds of coffee tree (*C. arabica* cv. Catuaí Vermelho) were collected in the experimental field of UFLA (-21.20549, -44.98021) at all stages of development to identify pre-meiosis, meiosis, and post-meiosis phases. Flower buds were fixed in Carnoy 3:1 and then stored in the refrigerator at 4°C. Slides were prepared according to (GARDNER; RATTENBURY, 1974) protocol. The pictures were taken with a 400x magnifying glass with more than 4X digital zoom.

Scanning Electron Microscopy (SEM)

Flower buds were removed from the primary fixative and they were transferred to microtubes containing 0.05 M cacodylate buffer for ten minutes. The buffer was changed three times. Samples were immersed in 1% osmium tetroxide solution in a cacodylate buffer for two hours at room temperature. Finally, they were washed with distilled water three times and dehydrated in series with acetone (25%, 50%, 75%, 90% - 10 minutes each, and 100% three times for 10 minutes). After this step, samples were completely dried utilizing a sputtering apparatus (Balzers CPD 030). After drying, samples were mounted on stubs and submitted to gold metallization in the Balzers SCD 050 vaporizer and observed in SEM (LEO EVO 40 XVP (Carl Zeiss) with Bruker X-ray microanalysis system (Quantax EDS) and cryosystem (Gatan).

Small RNA library construction and Illumina sequencing

Anthers from flower buds were macerated in liquid nitrogen, and total RNA was extracted from 100 mg of tissue using the PureLink® RNA Mini Kit (Invitrogen) according to the manufacturer's protocol. RNA samples were treated with DNase I using the RNase-free DNase Set (Qiagen) to eliminate residual DNA contamination. RNA integrity was visually analyzed in 1 %

agarose gel, and RNA content, as well as quality, were obtained by spectrophotometry (OD260/280 and OD260/230 > 1.8) (NanoVue GE Healthcare, Munich, Germany). RNA integrity number (RIN) was obtained with the Agilent 2100 Bioanalyzer system (Agilent Technologies) according to the manufacturer's recommendations for plant RNA. Only samples with RIN greater than 7 were used in the following steps. Size selection for 20-30 nucleotide (nt) small RNAs was performed in denaturing Urea-PAGE gels, and libraries were constructed using the TruSeq Small RNA Library Preparation Kit (Illumina, cat # RS-200-0024) (MATHIONI; KAKRANA; MEYERS, 2017). All libraries were single-end sequenced with 51 cycles on an Illumina HiSeq 2500 instrument.

Small RNA library quality control

A total of 104.303.112 small-RNA-seq single-end reads were inspected for adapters with a Minion (DAVIS et al., 2013). The identified adapter sequences were removed with Trimmomatic v. 0.39 (BOLGER; LOHSE; USADEL, 2014). Only reads with more than 17 nucleotides were selected for further analysis after quality control. After sequencing and data cleanup steps, the small RNA libraries were aligned against *C. arabica* and *C. canephora* genomes databases, in which approximately 95% of the reads aligned in the first and 80% in the second.

Computational *de novo* identification of phased, siRNA-generating loci and their miRNA triggers.

A total of 102.985.416 reads from pre-meiotic, meiotic, and post-meiotic small-RNAs libraries were submitted to the PHASIS suite (KAKRANA et al., 2017). The first step was the de novo prediction of PHASI loci with the phase detect tool. Subsequently, phase merge was applied with parameter "-mode" set to "merge" to create a unified list of the PHASIS predicted in samples from the different stages. Finally, phastrigs were applied to identify miRNAs triggers for the previously identified PHAS loci, to do so we supplied to phastrigs a fasta file containing all the mature miRNAs identified in the "Conserved and novel miRNA prediction from Coffea sp" section. These steps were run twice to identify both 21 or 24-PHAS and the minimum significance p-value was set to $1e^{-6}$. Custom r-scripts were developed to plot heatmaps of the inferred abundance of each predicted PHAS.

Identification and characterization of AGO-like and DCL-like proteins sequences in *C. arabica*

In order to increase the knowledge about the biogenesis of siRNAs in *C.arabica*, we searched for the main proteins of the pathway. Predicted protein sequences of *C. arabica* were obtained from its genome annotation available at NCBI's webpage

(https://www.ncbi.nlm.nih.gov/assembly/GCF_003713225.1/#/def) and utilized in the creation of a local database. AGO and DCL protein sequences of several monocot and eudicot species were downloaded from Genbank (BENSON et al., 2013), UniProt (THE UNIPROT CONSORTIUM, 2021), and the MaizeGDB (WOODHOUSE et al., 2021). To identify AGO and DCL homologs present in the coffee tree, downloaded sequences of each family were utilized in separate alignments against the proteome of *C. arabica* using BLASTp v2.10.1 (ALTSCHUL et al., 1990) with an e-value of 10^{-3}. To verify the presence of characteristic conserved domains from each protein family, candidate sequences were then analyzed using PFAM (MISTRY et al., 2020), and filtered so that only one protein per gene locus remained.

Phylogenetic analyses

For the classification of putative AGO and DCL proteins, a global alignment of the remaining candidate sequences was carried out together with downloaded annotated sequences from multiple species, using the MAFFT v7.475 program (ROZEWICKI et al., 2019) through the Guidance v2.02 algorithm (SELA et al., 2015), and utilized in subsequent phylogenetic analyses through the use of the PHYLIP package (RETIEF, 1999). For each protein family, we generated a dataset of 1000 bootstraps with the seqboot algorithm and the similarity matrices between the different sequences were calculated with the protdist algorithm. The trees were constructed using the neighbor-joining method through the neighbor algorithm and elucidated using the consensus algorithm. The FigTree v1.4.4 program (http://tree.bio.ed.ac.uk/software/figtree/) was used to visualize the trees and generate the figures.

Transcriptional analysis of AGO and DCL genes in RNAseq libraries of different coffee tissues

To analyze the abundance of the DCLs and AGOs genes identified, we used validated RNA-seq libraries from *C. arabica* already published, available in the Sequence Read Archive database (SRA, https://www.ncbi.nlm.nih.gov/sra). Libraries were subjected to quality analyses using the program FastQC v.0.11.9 (ROBINSON; MCCARTHY; SMYTH, 2010) and trimmed as necessary with Trimmomatic v0.39 (BOLGER et al., 2014). STAR v.2.7.8a (DOBIN et al., 2013) with standard parameters was used to carry out fragment alignments against the genome of *C. arabica*. The resulting libraries were sorted and read duplicates were removed with the Picard toolkit (http://broadinstitute.github.io/picard) and mapped fragments were counted with the htseq-count algorithm (ANDERS; PYL; HUBER, 2015). Expression values were calculated in log (RPKM + 1) using the edgeR package (ROBINSON et al., 2010), from the Bioconductor R v3.13 project (Huber et al., 2015).

Conserved and novel miRNA prediction in *Coffea sp.*

To search for putative conserved miRNAs and their precursors, we applied an adapted algorithm previously described by (DE SOUZA GOMES et al., 2011) to the genome and transcriptome databases of *C. canephora*. The second prediction of conserved and new miRNAs was made from RNA-seq libraries of *C. arabica.* mapped sRNA reads from libraries deposited in (unpublished) were used as input to two different computational pipelines for the discovery of miRNAs – a stringent pipeline for de novo identification and a relaxed pipeline for identification of conserved 'known' miRNAs (JEONG et al., 2013). Steps in both pipelines involved processing using Perl scripts as described earlier (JEONG et al., 2011), with a modified version of miREAP (https://sourceforge.net/projects/mireap/) and CentroidFold (SATO et al., 2009). In the 'stringent' criteria pipeline, sRNAs of length between 20 and 24 nt, with abundance ≥ 50 TP30M in at least one library, and total genome hits ≤ 20 were assessed for the potential pairing of miRNA and miRNA* using modified miREAP optimized for plant miRNA discovery with parameters –d 400 –f 25. Strand bias for precursors was computed as the ratio of all reads mapped to the sense strand against total reads mapped to both strands. In addition to strand bias, abundance bias was computed as the ratio of the two most abundant reads against all the reads mapped to the same precursor. Candidate precursors with strand bias ≥ 0.9 and abundance bias ≥ 0.7 were selected, and the foldback structure for the precursor was predicted using CentroidFold. Each precursor was manually inspected to match the criteria as described earlier (JEONG et al., 2013). All the miRNAs identified through this stringent pipeline were then annotated by matching mature sequences to miRBASE (version - 21), and those that did not match any known miRNA were considered lineage or species-specific. In the 'relaxed' criteria pipeline, which is implemented to maximize identification of 'known' miRNAs; relaxed filters were applied – sRNA between 20 and 24nt, with hits ≤ 20 and abundance ≥ 15 TP30M; and precursors with strand bias ≥ 0.7 and abundance bias ≥ 0.4. The stem-loop structure of candidate precursors was visually inspected, the same as the 'stringent' pipeline. Mature sequences of identified miRNAs were further matched with miRBASE entries (version-21), and those with total 'variance' (mismatches and overhangs) ≤ 4 were considered conserved miRNAs. Finally, sequences of the precursors of the two predictions were aligned and all parameters were compared. Redundant sequences were eliminated.

Expression evaluation of conserved and novel miRNAs identified in the *Coffea sp.* genomes

A total of 102,985,416 reads from pre-meiotic, meiotic, and post-meiotic small-RNAs libraries were aligned against the conserved and putative novel miRNAs identified in the *Coffea sp*. The alignment was performed with ShortStack v. 3.8.5 (AXTELL, 2013b) with the parameters "--

mmap u --nohp --locifile". Custom r-scripts were developed to plot heatmaps of the inferred abundance of each putative novel miRNAs.

RESULTS

Cytogenetic analysis identified the stages of microsporogenesis in *Coffea arabica*

The anthers of coffee flowers in different stages of development were collected in three stages, G4, G5 and G6 following classification of Morais et al., (2008). Through cytogenetic analysis, it was possible to determine that anthers of flower buds up to 1.5 mm (G3) had megasporocytes at the beginning of meiosis I (Fig.1 A), with cells in interphase (INT). Nucleus (NC) and nucleolus (NO) are visible, as the nuclear envelope (CONAGIN, 1961). In this phase, the chromosomes are thin and fulfilled and the duplication of DNA and chromosomes occurs, thus forming chromatids (MA, 2005). Anthers with cells in meiosis were observed in flower buds from 1.6 mm to 6 mm approximately (G4) (Fig. 1B), in which it is possible to observe cells in two stages of meiosis, prophase (PRO) and telophase (TPH). In prophase, the chromosomes have begun to condense and the nucleoli and nuclear envelope are no longer visible. In the cell in telophase, the last phase of meiosis, the chromosomes unwound and it is possible to see two nucleoli and the nuclear envelope reappears (MA, 2005; CONAGIN, 1961).

Flower buds from 7 mm (G6) onwards present uninucleated (NV) and vacuolated (VC) post-meiosis (Fig.1C) young microspores, with very thin cell walls, uniform, and clear cytoplasm, and a large, centralized nucleus. Binucleated microspores are also observed in post-meiosis, with a generative nucleus (NG) and a vegetative nucleus (NV). Soon after mitosis, the generative nucleus can be seen attached to the microspore cell membrane and then migrates to the center of the cell, already encased in its membrane system (PARK; TWELL, 2001). Anthers measuring 4.201 mm have cells in meiosis, a phase that coincides with buds in G4 according to Morais et al., 92008) (Fig. 1 D). In this phase the buds go through a latency period, and their development stops during the winter season, which coincides with the lack of rains in the main coffee-producing areas of Brazil (LÓPEZ et al., 2021). Mature pollen grains are not observed in anthers at this stage (Fig. 1 E), but they occur in anthers in post-meiosis (Fig. 1 F and G).

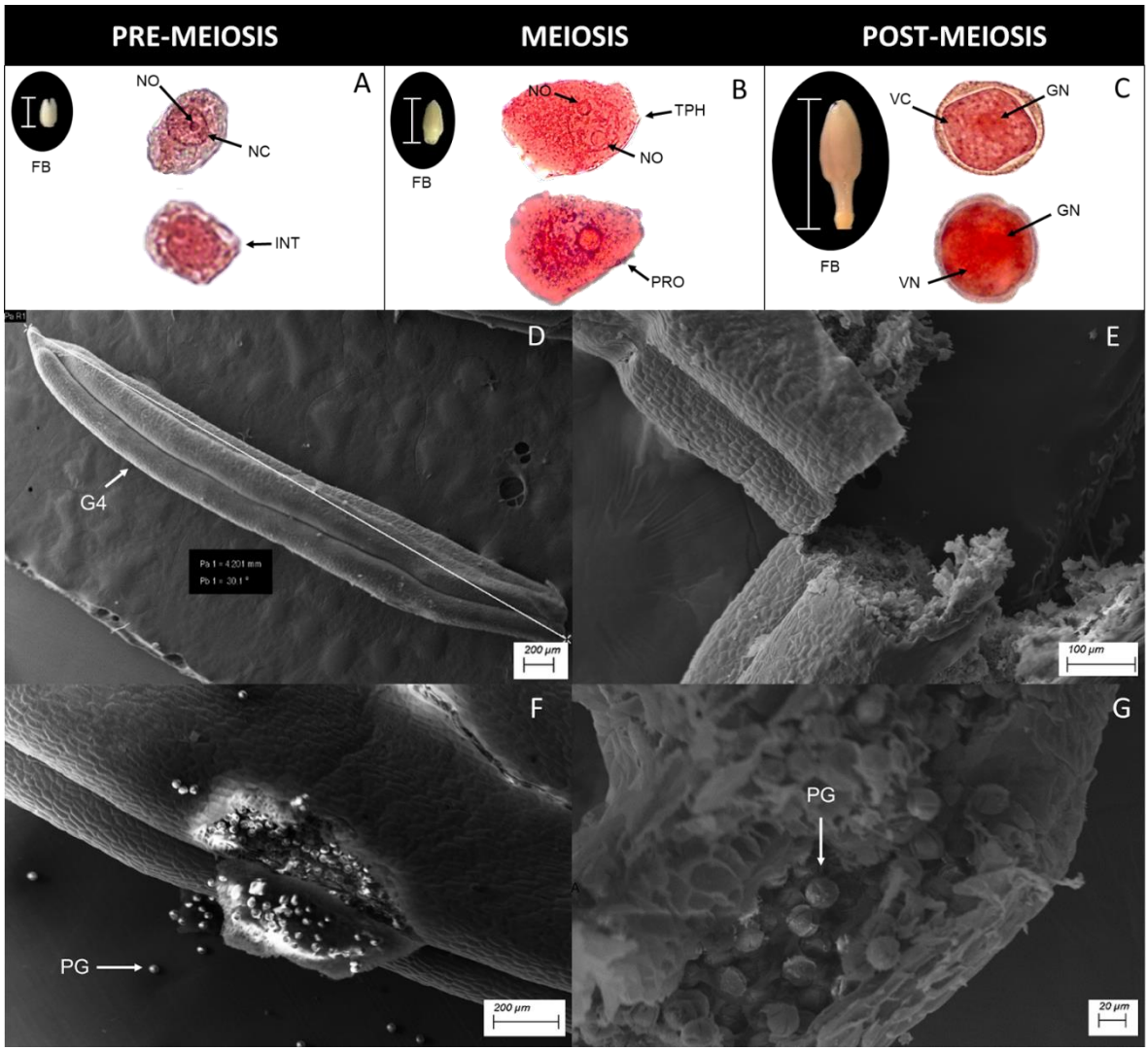


Fig. 1. Flower buds (FB) of coffee plants, light microscopy of coffee anthers cells in different stages with 400X magnification more than 4X digital zoom, and, scanning electron microscopy of coffee anthers. A: FB with pre-meiotic cells, megasporocytes in interphase (INT) with a well-defined nucleus (NC) and nucleolus (NO). Bar measuring 1,5 mm. B: FB with cells in meiosis, and coffee anther cells in prophase I, the beginning of meiosis and, in telophase (TPH), with two nucleoli visible. Bar measuring 4 mm. C: FB with post-meiotic cells, one young vacuolated (VC) and mononucleated microspore, and one binucleated microspore with a generative nucleus (GN) and a vegetative nucleus (VG). Bar measuring 15 mm. D: SEM of an anther from G4 buds measuring 4.2 mm. E: SEM of another fragment without mature pollen grains (PG). F: SEM of mature PG extravasated from the pollen sac from coffee anther. G: Scanning electron microscopy of mature PG in post-meiosis another fragment.

Characterization of DCLs and AGOs proteins in the *Coffea arabica* genome

DICER-likes (DCLs) and ARGONAUTE proteins (AGOs) are two families of proteins related to the production of siRNAs and phasiRNAS which are generated by specific and divergent members in plants (BAUMBERGER; BAULCOMBE, 2005; LIU; FENG; ZHU, 2009). Thus, to identify DCLs and AGOs and the possibility to generate phasiRNAs during microsporogenesis, we characterized the two families in the *C. arabica* genome.

As a result of searches in the *C. arabica* genome, proteins homologous to the DCLs-like of other monocots and eudicots were found. After multiple alignments of protein sequences, phylogenetic analyses were performed to assign the *C. arabica* DCLs their respective clades and

determine the evolutionary relationship between species. Three CaDCL1s, six CaDCL2s, two CaDCL3s, and one CaDCL4, were identified. In total, twelve DICER-like proteins were found in the genome of *C. arabica* and distributed among all four distinct clades of the phylogenetic tree (Fig 2), more than the five CaDCLs described in a previous study (MOSHARAF et al., 2019). *C. arabica* is an amphidiploid from the hybridization between *C. eugenioides* and *C. canephora* (LASHERMES et al., 1999). We found copies in the two parentals subgenomes for most of the identified *DCLs* genes, except for *CaDCL4* gene, which we identified as a copy from chromosome 6 of *C. canephora* (Table S1). A study in *C. canephora*, one of the parents of *C. arabica*, identified 9 DCLs, one CaDCL1, six CcDCL2s, one CcDCL3, and one CcDCL4 proteins (NORONHA FERNANDES-BRUM et al., 2017). In our work, we identified two more copies of *DCL1* gene from *C. eugenioides* and one new *DCL3* gene, from *the C. eugenioides* subgenome.

Dicer and DCL proteins are a unique class of ribonucleases that generally contain six types of domains, including the DEAD-box, helicase-C, DUF283, PAZ, RNase III, and dsRBD domains. However, for some species, one or more of these domains may be absent (LIU; FENG; ZHU, 2009; MARGIS et al., 2006). All identified CaDCLs presented the 3 Ribonuclease III domains, which have been considered one of the key domains for the DCL proteins (NICHOLSON, 2014), except for CaDCL1.3 which only presented two (Fig. S1). In addition, CaDCL1.1, CaDCL1.2, CaDCL2.1, CaDCL3.1, CaDCL3.2, and CaDCL4 had the highest number of domains in common, Ribonuclease III, DEAD/RES III, Helicase_C, Dicer_dimer, PAZ, and DSRM. Last one is a divergent double-stranded RNA-binding domain coinciding with the DUF283 of Dicer (DLAKIĆ, 2006).

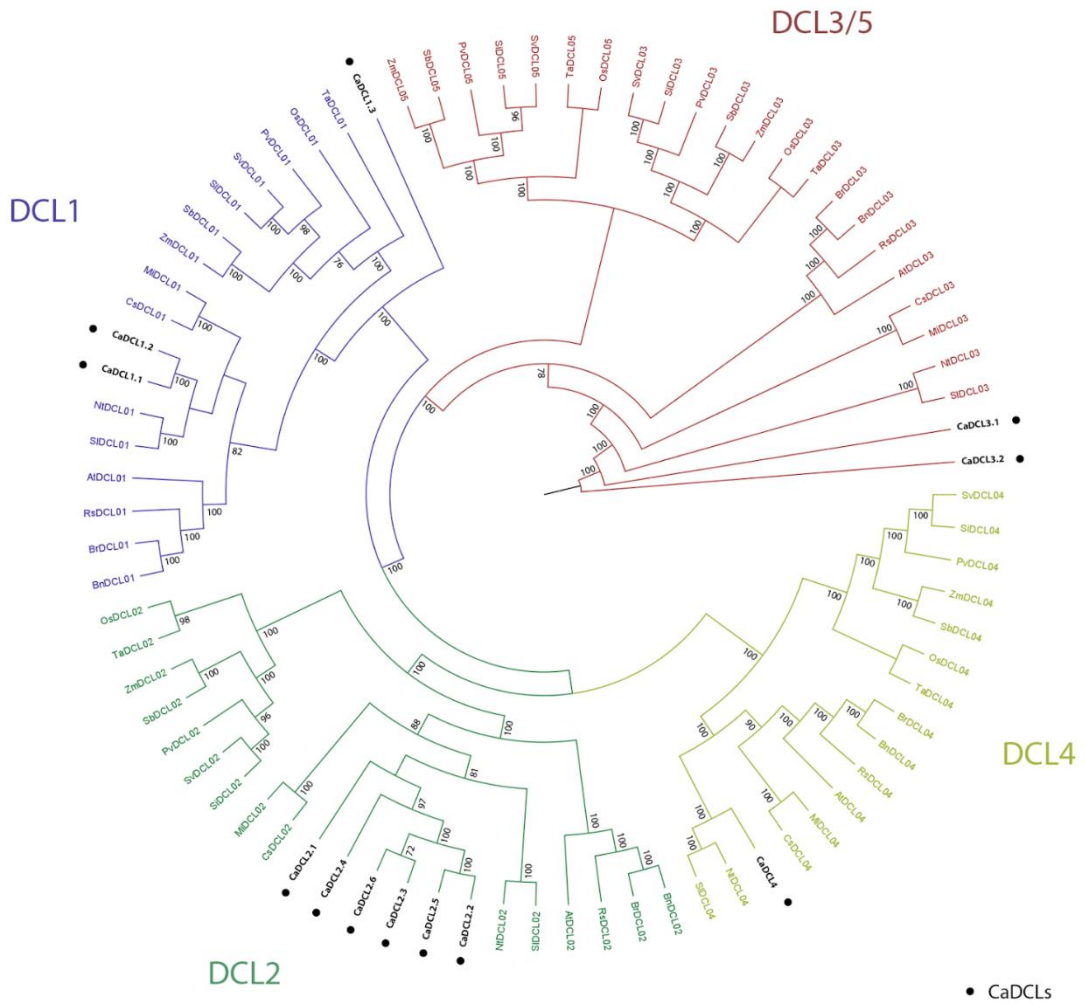


Fig. 2. Phylogenetic analysis of DCL proteins from *C. arabica*. The phylogenetic tree illustrating the protein family's evolutionary history was inferred using the neighbor-joining method and contains both monocot and eudicot sequences. The major DLC clades, in which the sequences were clustered, are indicated around the tree. Sequence details are available in table S1. Nodes with bootstrap values higher than 70% are indicated in the figure. *C. arabica* proteins are marked with black dots. A total of 12 *C. arabica* proteins were identified, three clustered in clade 1, six in clade 2, two in clade 3/5, and one in clade 4.

Flowering plants have ten or more members of AGO proteins (ZHAI et al., 2015). In total, 20 AGO-like proteins were found in the genome of *C. arabica* and distributed in three distinct clades of the phylogenetic tree (Fig 3), one CaAGO1, four CaAGO2s, five CaAGO4s, two CaAGO5s, two CaAGO6s, two CaAGO7s, and four CaAGO10s were identified. CaAGO1, CaAGO5, and CaAGO10 proteins clustered in clade 1 which comprises the AGO1/5/10 proteins. Proteins CaAGO2 and CaAGO7 were grouped into clade II, which comprises AGO2/3/7 proteins. CaAGO4 and CaAGO 6 proteins were clustered in clade III, which comprises the AGO4/6/8/9 proteins. Among the AGO family members identified, all had copies in both subgenomes of *C. arabica*, except for CaAGO1, for which we identified only one copy in chromosome 4 from *C.*

canephora (Table S1). In addition to the AGOs previously identified in the genome of *C. canephora* (NORONHA FERNANDES-BRUM et al., 2017), we identified two more copies of AGO2, one copy of AGO5, and three AGO4 from *C. eugenioides* chromosomes. We also identified one new AGO6 and two new copies of AGO10, one from each of the two subgenomes.

AGOs proteins are characterized by ArgoN (amino-terminal N), PAZ (PIWI–ARGONAUTE–ZWILLE), MID (middle), and PIWI domains (MEISTER, 2013). The conserved domains of the CaAGO proteins were analyzed, and all of them presented ArgoN, PAZ, PIWI, and ArgoL1 domains (Fig. S2). CaAGO1, CaAGO4.1, CaAGO4.2, CaAGO10.1, CaAGO10.2, CaAGO10.3, and CaAGO10.4 had all 6 domains ArgoN, PAZ, Piwi, ArgoL1, ArgoL2 and ArgoMID. The AGO1 proteins have an additional N-glycine-rich region (Gly-rich Ago1) domain that was present both in the CaAGO1 identified in this work and in the CcAGO1 identified by Noronha Fernandes-Brum et al. (2017). The MID domain is present in eight identified CaAGOs (CaAGO1, CaAGO 4.1, CaAGO 4.2, CaAGO 5.1, CaAGO 10.1, CaAGO 10.2, CaAGO 10.3, CaAGO 10.4). The MID domain anchors the 5′ ends of the small RNA by providing a binding pocket in which the 5′ terminal base engages in stacking interactions with a conserved tyrosine. In addition, several hydrogen bonds coordinate correct 5′ end binding (MEISTER, 2013).

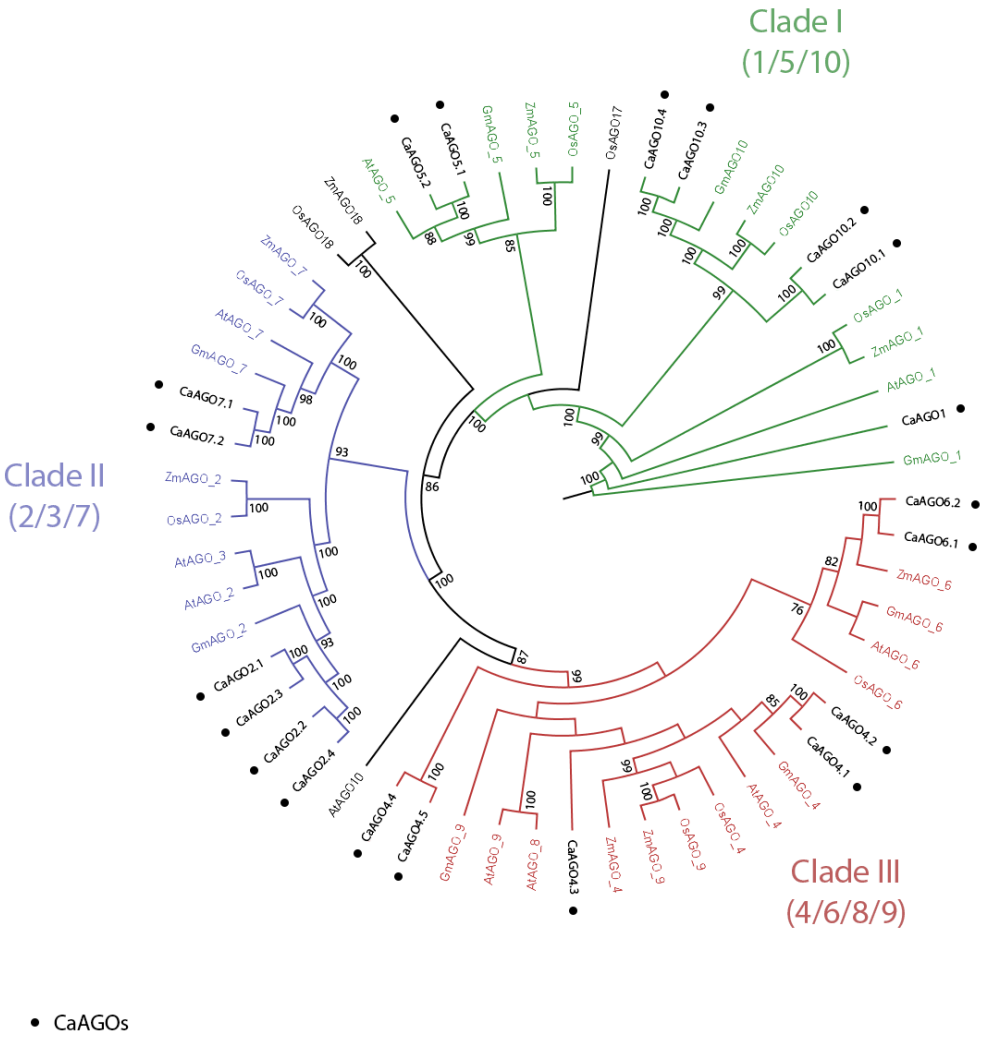


Fig. 3. Phylogenetic analysis of AGO proteins from *Coffea arabica*. The phylogenetic tree illustrating the protein family's evolutionary history was inferred using the neighbor-joining method and contains both monocot and eudicot sequences. The major AGO clades, in which the sequences were clustered, are indicated around the tree. Sequence details are available in table S1. Nodes with bootstrap values higher than 70% are indicated in the figure. *C. arabica* proteins are marked with black dots. A total of 20 *C. arabica* proteins were identified, seven clustered in clade I (AGO1/5/10), six in clade II (AGO2/3/7), and seven in clade III (AGO4/6/8/9).

Expression analyses of *DCLs* and *AGOs* genes in different RNAseq libraries

To analyze the expression profile of *CaDCLs* and *CaAGOs* identified in this work, we aligned them to RNA-seq libraries from *C. arabica* tissues under different conditions: leaves from plants under optimal temperature conditions and warm temperature conditions (cv. Acauã and Catuaí); floral buds at G4 and G5 stages (cv. Catuaí and Siriema), embryos, stems, roots, meristems, floral buds, green drupes, yellow drupes, red drupes, fruits at multiple stages of development, fruits 30 days after flowering, and fruits 90 days after flowering. Of the twenty AGO

proteins identified in this work, eleven were expressed in the analyzed libraries (Fig. 4a): CaAGO1, CaAGO2.1, CaAGO2.2, CaAGO2.3, CaAGO2.4, CaAGO4.2, CaAGO4.3, CaAGO5.2, CaAGO7.1, CaAGO7.2, and CaAGO10.2. Of these, CaAGO1 was the most expressed in all libraries. As for the DCL proteins, of the 12 proteins identified, only 3 were expressed, CaDCL2.3, CaDCL2.6, and CaDCL3.2 (Fig. 4b). Of these, CaDCL3.2 was the most expressed, in the reproductive tissues G4_CA, G5_CA, G4_SE, and G5_SE.

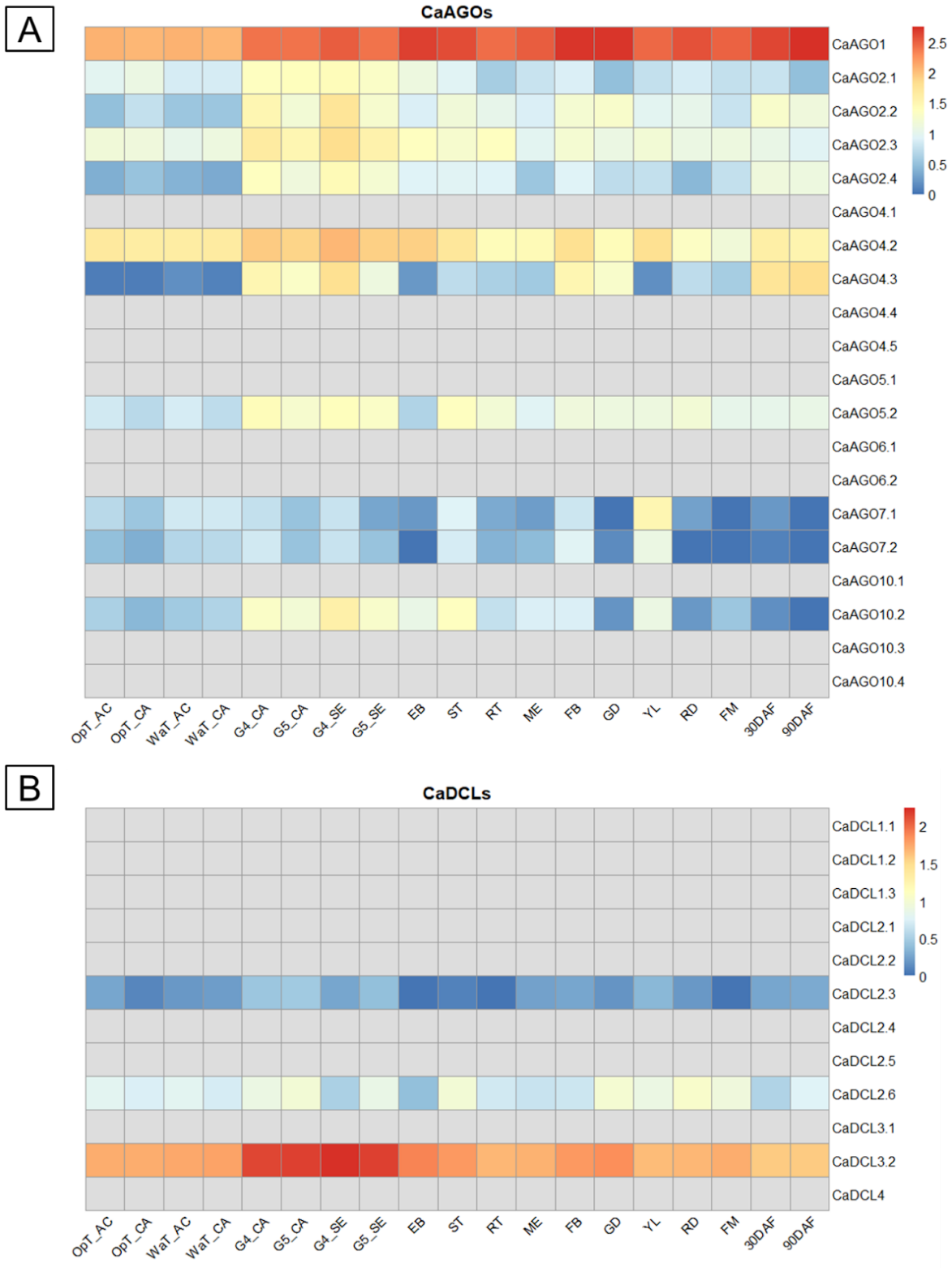


Fig. 4. Heatmap representing coffee DCLs (A) and AGOs (B) proteins (lines) abundance in different RNA-seq libraries (columns). OpT_AC: *C. arabica* (cv. Acauã) leaves from plants in optimal temperature conditions; OpT_CA: *C. arabica* (Catuaí cultivar) leaves from plants in optimal temperature conditions; WaT_AC: *C. arabica* (Acauã cultivar) leaves from plants in warm temperature conditions; WaT_CA: *C. arabica* (Catuaí cultivar) leaves from plants in warm temperature conditions; G4_CA: floral buds at G4 stage from cv. Catuaí; G5_CA: floral buds at G5 stage from Catuaí cultivar; G4_SE: floral buds at G4 stage from Siriema cultivar; G5_SE: floral buds at G5 stage from Siriema cultivar; EB: embryos; ST: stems; RT: roots; ME: meristems; FB: floral buds; GD: green drupes; YL: yellow drupes; RD: red drupes; FM: fruits at

multiple stages of development; 30DAF: fruits 30 days after flowering; 90DAF: fruits 90 days after flowering. Colors are coded as a function of logarithmic base 10 of expression values normalized in FPKM +1 (Fragments per million mapped reads).

Identification of miRNAs and reproductive phasiRNAs and their quantification in sRNA-seq libraries

Our analyses allowed the identification of 138 miRNA precursors and 149 gender-specific mature miRNAs, 80 of which were miRNA-5p and 69 miRNA-3p. We also identified 259 conserved miRNA precursors that gave rise to 349 mature ones, 98 miRNA-5p and 251 miRNA-3p (Table S2). The phasiRNAs identified in *C. arabica* sRNA-seq libraries are classified in the Table S3. In our analyses, 21-nt and 24-nt reproductive phasiRNAs were identified. Different from the data reported for many monocots (FAN et al., 2016; ZHAI et al., 2015), the 21-nt phase is more enriched in the post-meiotic stage and the 24nt-phasiRNAs in the pre-meiotic one (Fig.5a and b). It is noteworthy that recently, (BÉLANGER et al., 2020) also identified 24-nt phasiRNAs enriched in wheat pre-meiosis, in which they represented more than 50% of all 24-PHAS loci that were annotated. The meiotic condition has the lowest number of phasiRNAs of both classes. We found 10 24-nt phasiRNAs and 7 21-nt phasiRNAs. In the literature, miR2118/482 is described as responsible for processing 21-nt phasiRNAs while miR2275 processes 24-nt phasiRNAs (KAKRANA et al., 2018; KOMIYA, 2017; XIA et al., 2019). In our results, car-miR482 is one of the main triggers for 21-nt phasiRNAs. In addition, car-miR2275 appears as a trigger for both 21-nt and 24-nt phasiRNAs (Table S4). We have not identified miR2118 in our data, so in addition to car-miR2275, other miRNAs can cleave the 24-nt phasiRNAs. We also quantified the conserved and lineage-specific miRNAs in the sRNA libraries, and it was possible to observe that one member of the miR482 family was one of the most abundant in all the stages of microsporogenesis in *C. arabica* (Fig. 5c).

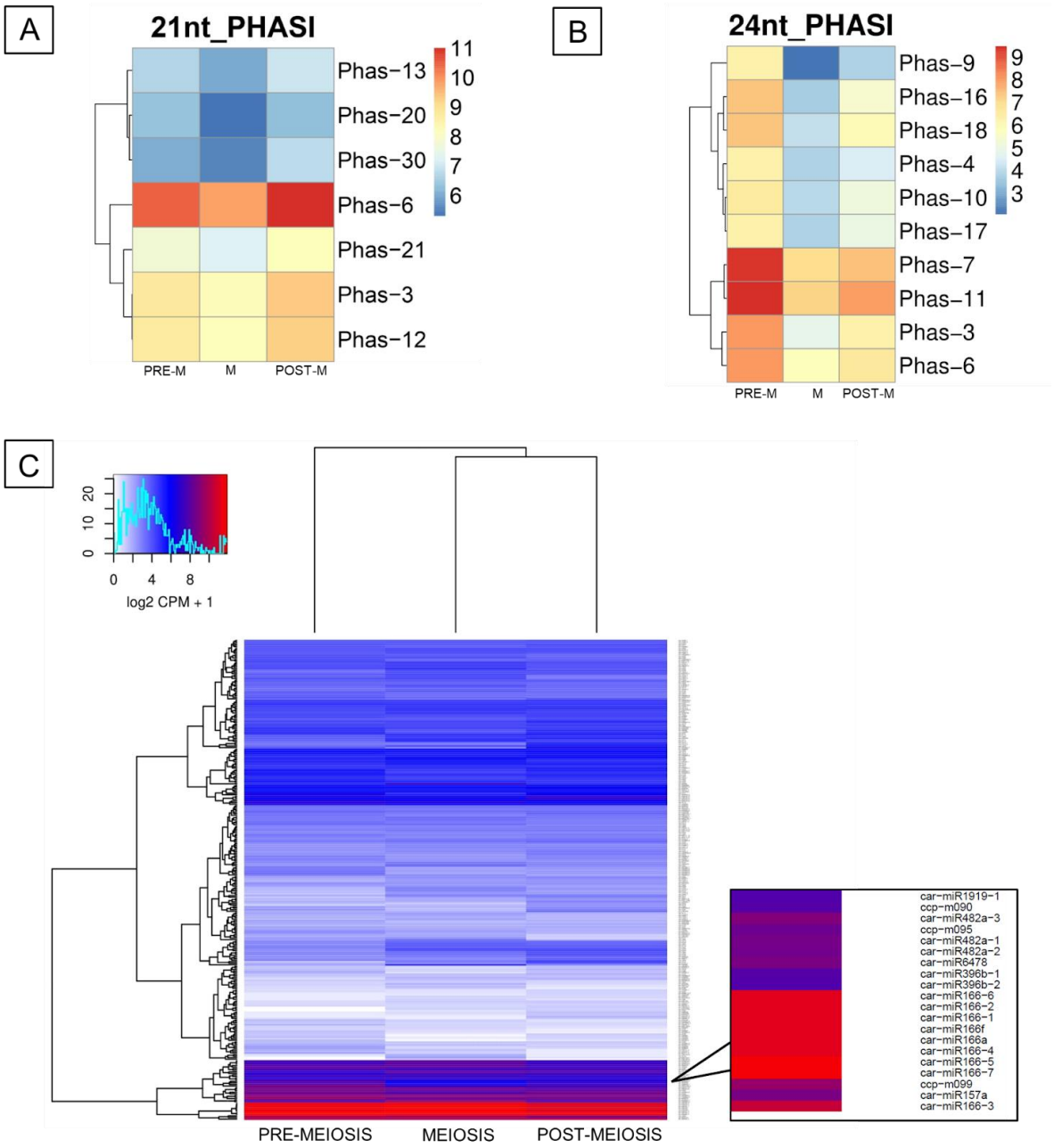


Fig. 5. A: Heatmap of 21-nt phasiRNAs (lines) identified in libraries (columns) of anthers in pre-meiosis (PRE-M), meiosis (M), and post-meiosis (POST-M). B: Heatmap of 24-nt phasiRNAs (lines) identified in libraries (columns) of anthers in pre-meiosis (PRE-M), meiosis (M), and post (POST-M). C: Heatmap of conserved and novel miRNAs (lines) identified in libraries (columns) of anthers in pre-meiosis, meiosis, and post-meiosis. The dendrogram on top of the heatmaps reflects the relationship of different conditions while the dendrogram on the left reflects the relationship of different sRNAs across conditions. Colors are coded as a function of logarithmic base 10 of expression values normalized in FPKM +1 (Fragments per million mapped reads) in A e B. Colours are coded as a function of logarithmic base 2 of expression values normalized in CPM+1 (Counts per million mapped reads) in C.

DISCUSSION

To comprehend the molecular pathways that govern the development of anthers, a detailed understanding of their morphological development is necessary. Our study allowed us to identify a crucial phase of microsporogenesis, meiosis. Meiosis in coffee anthers coincides with the G4 phase of bud development (MORAIS et al., 2008) and with the stage 4 considering optical microscopy

analysis (DE OLIVEIRA et al., 2014). At this stage, buds measure from 3.1 to 6 mm in length and, as mentioned earlier, are in a latent state but already occurred differentiation and elongation of the reproductive floral organs, such as the anthers (DE OLIVEIRA et al., 2014). However, there are no studies demonstrating the physiological basis of this type of dormancy (LÓPEZ et al., 2021). In this work, we show that there is activity at the cellular level at this stage of development, despite the visible latency.

Reproductive phasiRNAs are crucial for the microsporogenesis of many species, especially monocots (LIU; WANG, 2021; LIU et al., 2020). It was previously believed that only this group had 24-nt phasiRNAs, because their processing depends on a specialized DCL5 not previously reported in eudicots. However, recent studies have demonstrated the presence of these phasiRNAs in eudicots (XIA et al., 2019) and now we have also identified this group in *C. arabica*. Compared to monocots and other eudicots, we identified a smaller number of reproductive phasiRNAs, and a specific temporal pattern, with the 21 nt-phasiRNAs being more expressed in post-meiosis, and not in pre-meiosis, and the 24-nt phasiRNAs in pre-meiosis and not in meiosis (KAKRANA et al., 2018; POKHREL; HUANG; MEYERS, 2021; XIA et al., 2019).

The studies followed the opposite course of evolutionary history. The recent discovery of 24-nt phasiRNAas in eudicots drew attention to which DCL was involved in this process, and until now, this was uncovered in eudicots (LIU et al., 2020; POKHREL; HUANG; MEYERS, 2021; XIA et al., 2019). The emergence of DCL5 in monocots is believed to be explained by a subfunctionalization of the ancestral DCL3 eudicot protein, which is suggested to function in the production of heterochromatic siRNAs (hc-siRNAs) and phasiRNAs (LIU et al., 2020). In rice, OsDCL3 and OsDCL5 have completely different substrate specificities, while the eudicot AtDCL3 has an intermediate preference for dsRNAs with an overhanging 5' triphosphate and 3' structure (CHEN et al., 2022). This implies that monocot DCL5 and DCL3 proteins were optimized for cognate substrates after duplication of the old 'eudicot-like' DCL3 protein and this functional specialization appears to have been achieved through the accumulation of mutations in the PAZ domain (CHEN et al., 2022). Therefore, it is hypothesized that DCL3 has dual functions in some eudicots, processing distinct substrates (Pol II vs Pol IV RNA) for the biogenesis of 24-nt phasiRNAs or 24-nt hc-siRNAs (XIA et al., 2019), and future works have to explore this possibility.

Interestingly, of the three *CaDCLs* genes expressed in the tissues analyzed in this study, *CaDCL3.2* was more abundant in the libraries of reproductive tissues, flower buds in the G4 and G5 stages, precisely those corresponding to meiosis and post meiosis during coffee anthers development, supporting hypothesis of a role of this enzyme for siRNAs biogenesis of during the microsporogenesis. DCL4, described as the responsible for cleaving the 21-nt phasiRNAs, although

identified in this work did not show expression in any of the analyzed tissues. Each class of DCL has evolved to participate in a primary pathway, however, DCLs2-4 can also function redundantly, because defects in one class of DCL may be compensated by other classes in some cases (FUKUDOME; FUKUHARA, 2017; POKHREL; HUANG; MEYERS, 2021). Thus, it is possible that other DCLs, as CaDCL3 protein may be involved in the biogenesis of 21-phasiRNAs in *C.arabica*.

In our study we identified twenty members of the CaAGO-like protein family. The high number of AGO protein family members in flowering plants points to functional diversification, presumably reflecting regulatory pathways targeting the expanding sRNA class of molecules (ZHANG et al., 2015). Members of the AGO4/6/8/9 clade have been described as effectors of 24 nt hc-siRNAs and are known to function in the regulation of transcriptional genes (ZHANG et al. 2015). In some eudicots, works based on phylogenetic and expression data, suggested that AGO5 and AGO1 could act as effector proteins for 21-nt phasi and AGO6 and AGO9 for 24-nt phasiRNA (POKHREL; HUANG; MEYERS, 2021). In our data, at least one copy of each member of the *CaAGO*-like gene family was expressed in the libraries, with CaAGO4.2 and CaAGO4.3 being more abundant in the reproductive libraries, suggesting the role of the codified enzymes in the biogenesis of 24-nt phasiRNAs. *CaAGO6* gene was not expressed in the analyzed libraries and we did not identify AGO9 protein in our data.

Another relevant element in the biogenesis of reproductive phasiRNAs are miRNAs that act as triggers of them. In several species, as solanaceas (XIA et al., 2019b), barley and wheat (BÉLANGER et al., 2020), other triggers for 24-nt phasi RNAs have already been described for 24-nt phasiRNAs, in addition to the canonical miRNA miR2275. Interestingly, miR2118/482, known to trigger 21-nt phasiRNAs in eudicots and monocots, also has a role in generating 24-nt phasiRNAs in two eudicots as columbine and wild strawberry. (POKHREL; HUANG; MEYERS, 2021). This suggests that miR2118/482 members present in gymnosperms may initiate 24-nt phasiRNA expression in angiosperms/basal eudicots, but miR2275 may have arisen later and therefore assumed this role later. In *C. arabica*, miR2118 seems to have been lost, while the miR482 family acts as a trigger for 21-nt phasiRNAs, possibly together the miR2275 family that supposed to have assumed a role in the processing of both reproductive phasiRNAs. Further experiments are necessary to demonstrate this hypothesis, which has never been reported before.

CONCLUSIONS

In this work, we reported siRNAs realated to the microsporogenesis of *C. arabica* and molecular pathways related to its biosynthesis. We identified two families of the most important proteins involved in this process, CaAGOs and CaDCLs, in addition to canonical and new miRNA

triggers for reproductive phasiRNAs of 21 and 24-nt. The presence of reproductive phasiRNAs in *C. arabica* presents a temporal pattern different from most of the eudicots already studied, with 21-phasi RNAs enriched in post-meiosis and 24-nt phasi-RNAs in pre-meiosis, which shows the different pattern of this class of phasiRNAs in a woody eudicot such as coffee tree. The triggers identified here reinforce data that canonical microRNAs such as miR2275 may be involved in the processing of both 21 and 24-nt reproductive phasiRNAs. Despite advances in knowledge about the role of reproductive phasiRNAs in eudicots microsporogenesis, functional experiments are needed to fully characterize this pathway, in addition to understanding the targets of these molecules, which will allow for the modulation of sterility in species of commercial interest such as coffee tree.

ACKNOWLEDGEMENTS

We thank the Federal University of Lavras (UFLA/Brazil) and members of the Laboratory of Plant Molecular Physiology (LFMP, UFLA/Brazil) for structural support of the experiments and analysis. We also thank professor Vania Helena Techio of the Biology Institute (UFLA/Brazil) for technical support in the cytogenetic analyses. Finally, we thank Instituto Brasileiro de Ciência e Tecnologia do Café (INCT/Café), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG) for the financial support.

AUTHORS' CONTRIBUTIONS

R.R.O., C.N.F.B., B.C.M, and A.C-J. conceptualized the project; K.K.P.O. conducted all data research and analyses; G.C.R.; T.H.C.R. A.K and S.M. supported for bioinformatic analyses; R.R.O. and A.C-J supervised experiments and analyses. K.K.P.O wrote the manuscript; R.R.O. and G.C.R contributed to writing.

STATEMENTS AND DECLARATIONS

The authors declare that there are no conflicts of interest.

REFERENCES

ALTSCHUL, S. F. et al. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, n. 3, p. 403–410, 5 out. 1990.

ANDERS, S.; PYL, P. T.; HUBER, W. HTSeq—a Python framework to work with high-throughput sequencing data. **Bioinformatics**, v. 31, n. 2, p. 166–169, 15 jan. 2015.

ANDREWS, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

AXTELL, M. J. Classification and Comparison of Small RNAs from Plants. **Annual Review of Plant Biology**, v. 64, n. 1, p. 137–159, 29 abr. 2013a.

AXTELL, M. J. ShortStack: comprehensive annotation and quantification of small RNA genes. **RNA (New York, N.Y.)**, v. 19, n. 6, p. 740–751, jun. 2013b.

BÉLANGER, S. et al. Premeiotic, 24-Nucleotide Reproductive PhasiRNAs Are Abundant in Anthers of Wheat and Barley But Not Rice and Maize. **Plant Physiology**, v. 184, n. 3, p. 1407–1423, 1 nov. 2020.

BENSON, D. A. et al. GenBank. **Nucleic Acids Research**, v. 41, n. Database issue, p. D36–D42, jan. 2013.

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 1 ago. 2014.

CHEN, S. et al. Functional specialization of monocot DCL3 and DCL5 proteins through the evolution of the PAZ domain. **Nucleic Acids Research**, v. 50, n. 8, p. 4669–4684, 5 abr. 2022.

CONAGIN, C. H. T. M. Microsporogênese, incompatibilidade e esterilidade masculina em Coffea congensis froehner. **Bragantia**, v. 20, p. 669–677, 1961.

DAVIS, M. P. A. et al. Kraken: A set of tools for quality control and analysis of high-throughput sequence data. **Methods**, Diversity of the non-coding transcriptomes revealed by RNA-seq technologies. v. 63, n. 1, p. 41–49, 1 set. 2013.

DE OLIVEIRA, R. R. et al. Flower development in Coffea arabica L.: new insights into MADS-box genes. **Plant Reproduction**, v. 27, n. 2, p. 79–94, 1 jun. 2014.

DE SOUZA GOMES, M. et al. Genome-wide identification of novel microRNAs and their target genes in the human parasite Schistosoma mansoni. **Genomics**, v. 98, n. 2, p. 96–111, ago. 2011.

DLAKIĆ, M. DUF283 domain of Dicer proteins has a double-stranded RNA-binding fold. **Bioinformatics**, v. 22, n. 22, p. 2711–2714, 15 nov. 2006.

DOBIN, A. et al. STAR: ultrafast universal RNA-seq aligner. **Bioinformatics (Oxford, England)**, v. 29, n. 1, p. 15–21, 1 jan. 2013.

FAN, Y. et al. PMS1T, producing phased small-interfering RNAs, regulates photoperiod-sensitive male sterility in rice. **Proceedings of the National Academy of Sciences of the United States of America**, v. 113, n. 52, p. 15144–15149, 27 dez. 2016.

FEI, Q. et al. Dynamic changes of small RNAs in rice spikelet development reveal specialized reproductive phasiRNA pathways. **Journal of Experimental Botany**, v. 67, n. 21, p. 6037–6049, nov. 2016.

FEI, Q.; XIA, R.; MEYERS, B. C. Phased, Secondary, Small Interfering RNAs in Posttranscriptional Regulatory Networks. **The Plant Cell**, v. 25, n. 7, p. 2400–2415, 1 jul. 2013.

FUKUDOME, A.; FUKUHARA, T. Plant dicer-like proteins: double-stranded RNA-cleaving enzymes for small RNA biogenesis. **Journal of Plant Research**, v. 130, n. 1, p. 33–44, 1 jan. 2017.

GARDNER, R. O.; RATTENBURY, J. A. Staining pollen nuclei with Snow's alcoholic hydrochloric acid-carmine. **Stain Technology**, v. 49, n. 5, p. 257–259, set. 1974.

JEONG, D.-H. et al. Massive analysis of rice small RNAs: mechanistic implications of regulated microRNAs and variants for differential target RNA cleavage. **The Plant Cell**, v. 23, n. 12, p. 4185–4207, dez. 2011.

JEONG, D.-H. et al. Comprehensive Investigation of MicroRNAs Enhanced by Analysis of Sequence Variants, Expression Patterns, ARGONAUTE Loading, and Target Cleavage1[W][OA]. **Plant Physiology**, v. 162, n. 3, p. 1225–1245, jul. 2013.

KAKRANA, A. et al. **PHASIS: A computational suite for de novo discovery and characterization of phased, siRNA-generating loci and their miRNA triggers**. bioRxiv, , 3 jul. 2017. Disponível em: <https://www.biorxiv.org/content/10.1101/158832v1>. Acesso em: 22 ago. 2022

KAKRANA, A. et al. Plant 24-nt reproductive phasiRNAs from intramolecular duplex mRNAs in diverse monocots. **Genome Research**, v. 28, n. 9, p. 1333–1344, set. 2018.

KOMIYA, R. Biogenesis of diverse plant phasiRNAs involves an miRNA-trigger and Dicer-processing. **Journal of Plant Research**, v. 130, n. 1, p. 17–23, jan. 2017.

LASHERMES, P. et al. Molecular characterisation and origin of the Coffea arabica L. genome. **Molecular and General Genetics MGG**, v. 261, n. 2, p. 259–266, 1 mar. 1999.

LIU, L.; WANG, T. Male gametophyte development in flowering plants: A story of quarantine and sacrifice. **Journal of Plant Physiology**, v. 258–259, p. 153365, 1 mar. 2021.

LIU, Q.; FENG, Y.; ZHU, Z. Dicer-like (DCL) proteins in plants. **Functional & Integrative Genomics**, v. 9, n. 3, p. 277–286, 1 ago. 2009.

LIU, Y. et al. PhasiRNAs in Plants: Their Biogenesis, Genic Sources, and Roles in Stress Responses, Development, and Reproduction. **The Plant Cell**, v. 32, n. 10, p. 3059–3080, out. 2020.

LÓPEZ, M. E. et al. An overview of the endogenous and environmental factors related to the *Coffea arabica* flowering process. **Beverage Plant Research**, v. 1, n. 1, p. 1–16, 18 nov. 2021.

BAUMBERGER, N.; BAULCOMBE, D. C. Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 33, p. 11928–11933, 16 ago. 2005.

CHEN, S. et al. Functional specialization of monocot DCL3 and DCL5 proteins through the evolution of the PAZ domain. **Nucleic Acids Research**, v. 50, n. 8, p. 4669–4684, 5 abr. 2022.

DAMATTA, F. M. et al. Why could the coffee crop endure climate change and global warming to a greater extent than previously estimated? **Climatic Change**, v. 152, n. 1, p. 167–178, jan. 2019.

DAMATTA, F. M.; RAMALHO, J. D. C. Impacts of drought and temperature stress on coffee physiology and production: a review. **Brazilian Journal of Plant Physiology**, v. 18, n. 1, p. 55–81, mar. 2006.

DE OLIVEIRA, R. R. et al. Flower development in Coffea arabica L.: new insights into MADS-box genes. **Plant Reproduction**, v. 27, n. 2, p. 79–94, 1 jun. 2014.

DE OLIVEIRA, R. R. et al. Elevated Temperatures Impose Transcriptional Constraints and Elicit Intraspecific Differences Between Coffee Genotypes. **Frontiers in Plant Science**, v. 11, p. 1113, 21 jul. 2020.

LIU, Q.; FENG, Y.; ZHU, Z. Dicer-like (DCL) proteins in plants. **Functional & Integrative Genomics**, v. 9, n. 3, p. 277–286, 1 ago. 2009.

MA, H. Molecular genetic analyses of microsporogenesis and microgametogenesis in flowering plants. **Annual Review of Plant Biology**, v. 56, p. 393–434, 2005.

NORONHA FERNANDES-BRUM, C. et al. A genome-wide analysis of the RNA-guided silencing pathway in coffee reveals insights into its regulatory mechanisms. **PLoS ONE**, v. 12, n. 4, p. e0176333, 27 abr. 2017.

SHARMA, D. K. et al. Wheat cultivars selected for high Fv/Fm under heat stress maintain high photosynthesis, total chlorophyll, stomatal conductance, transpiration and dry matter. **Physiologia Plantarum**, v. 153, n. 2, p. 284–298, 2015.

MARGIS, R. et al. The evolution and diversification of Dicers in plants. **FEBS Letters**, v. 580, n. 10, p. 2442–2450, 1 maio 2006.

MATHIONI, S. M.; KAKRANA, A.; MEYERS, B. C. Characterization of Plant Small RNAs by Next Generation Sequencing. **Current Protocols in Plant Biology**, v. 2, n. 1, p. 39–63, 2017.

MEISTER, G. Argonaute proteins: functional insights and emerging roles. **Nature Reviews Genetics**, v. 14, n. 7, p. 447–459, jul. 2013.

MISTRY, J. et al. Pfam: The protein families database in 2021. **Nucleic Acids Research**, v. 49, n. D1, p. D412–D419, 30 out. 2020.

MORAIS, H. et al. Escala fenológica detalhada da fase reprodutiva de Cooffea arabica. **Bragantia**, v. 67, p. 257–260, 2008.

MOSHARAF, M. P. et al. Genome-wide identification, characterization and phylogenetic analysis of Dicer-like (DCL) gene family in Coffea arabica. **Bioinformation**, v. 15, n. 11, p. 824, 2019.

NORONHA FERNANDES-BRUM, C. et al. A genome-wide analysis of the RNA-guided silencing pathway in coffee reveals insights into its regulatory mechanisms. **PLoS ONE**, v. 12, n. 4, p. e0176333, 27 abr. 2017.

PARK, S.K., TWELL, D., 2001. Novel patterns of ectopic cell plate growth and lipid body distribution in the Arabidopsis gemini pollen1 mutant. Plant Physiol 126, 899–909. https://doi.org/10.1104/pp.126.2.899

POKHREL, S.; HUANG, K.; MEYERS, B. C. **Conserved and non-conserved triggers of 24-nt reproductive phasiRNAs in eudicots**. [s.l.] Plant Biology, 21 jan. 2021. Disponível em: <http://biorxiv.org/lookup/doi/10.1101/2021.01.20.427321>. Acesso em: 5 ago. 2022.

RETIEF, J. D. Phylogenetic Analysis Using PHYLIP. Em: MISENER, S.; KRAWETZ, S. A. (Eds.). **Bioinformatics Methods and Protocols**. Methods in Molecular Biology™. Totowa, NJ: Humana Press, 1999. p. 243–258.

ROBINSON, M. D.; MCCARTHY, D. J.; SMYTH, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. **Bioinformatics (Oxford, England)**, v. 26, n. 1, p. 139–140, 1 jan. 2010.

ROZEWICKI, J. et al. MAFFT-DASH: integrated protein sequence and structural alignment. **Nucleic Acids Research**, v. 47, n. W1, p. W5–W10, 2 jul. 2019.

SATO, K. et al. CENTROIDFOLD: a web server for RNA secondary structure prediction. **Nucleic Acids Research**, v. 37, n. Web Server issue, p. W277-280, jul. 2009.

SELA, I. et al. GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. **Nucleic Acids Research**, v. 43, n. W1, p. W7–W14, 1 jul. 2015.

THE UNIPROT CONSORTIUM. UniProt: the universal protein knowledgebase in 2021. **Nucleic Acids Research**, v. 49, n. D1, p. D480–D489, 8 jan. 2021.

WOODHOUSE, M. R. et al. A pan-genomic approach to genome databases using maize as a model system. **BMC Plant Biology**, v. 21, n. 1, p. 385, 20 ago. 2021.

XIA, R. et al. 24-nt reproductive phasiRNAs are broadly present in angiosperms. **Nature Communications**, v. 10, n. 1, p. 627, 7 fev. 2019.

ZHAI, J. et al. Spatiotemporally dynamic, cell-type–dependent premeiotic and meiotic phasiRNAs in maize anthers. **Proceedings of the National Academy of Sciences**, v. 112, n. 10, p. 3146–3151, 10 mar. 2015.

ZHANG, H. et al. Evolution, functions, and mysteries of plant ARGONAUTE proteins. **Current Opinion in Plant Biology**, Cell signalling and gene regulation. v. 27, p. 84–90, 1 out. 2015.

ZHENG, Y. et al. A dynamic evolutionary and functional landscape of plant phased small interfering RNAs. **BMC Biology**, v. 13, n. 1, p. 32, 16 maio 2015.

ARTIGO 2: NOVEL AND CONSERVED MICRORNA ACT IN RESPONSE TO TEMPERATURE ELEVATION IN *COFFEA* REGULATING GABA AND ANTHOCYANIN.

TITLE

Novel and conserved microRNA act in response to temperature elevation in *Coffea* regulating GABA and anthocyanin.

AUTHORS

Kellen Kauanne Pimenta de Oliveira[1], Raphael Ricon de Oliveira[1], Jacqueline Oliveira dos Santos[2], Rafael Moreira Oliveira[1], Dawyson de Lima[1], Gabriel de Campos Rume[1], Iasminy Silva Santos[1], Robert Márquez Gutiérrez[1], Thales Henrique Cherubino Ribeiro[1], Christiane Noronha Fernandes-Brum[1], Antonio Chalfun-Júnior[1*].

AFFILIATIONS

[1]Laboratory of Plant Molecular Physiology, Plant Physiology Sector, Department of Biology, Federal University of Lavras, 3037, 37200-900, Lavras MG, Brazil.

[2]Laboratory of Plant Biochemistry and Metabolism, Department of Biology, Federal University of Lavras, Lavras, Brazil.

CORRESPONDING AUTHOR

Antonio Chalfun-Júnior

Email: chalfunjunior@ufla.br

Laboratory of Plant Molecular Physiology, Plant Physiology Sector, Department of Biology, Federal University of Lavras, 3037, 37200-900, Lavras MG, Brazil.

crops

ABSTRACT

Currently, predicted changes to the climate scenario are alarming, and the search for technologies to mitigate the effects of heat stress in crops is a necessity to maintain the global food supply. Thus, understanding the underlying mechanisms that act in the regulation of heat stress in large crops such as coffee plants is relevant and may be useful for this purpose, and then expanding the knowledge to other crops. MiRNAs are small molecules that act in the post-transcriptional regulation of their target genes and have already been reported to act in the thermal regulation in many species. In this context, we use two species of coffee plants that are mainly cultivated, *C. arabica* and *C. canephora*, to investigate the mechanisms of response to heat stress, in this genre. We subjected four genotypes (*C. arabica* cv. Mundo Novo, Catuaí Vermelho, and Arara, and also *C. canephora* cv. Robusta Tropical) to temperature ramp in a growth chamber (T1: 15 days at 19°C/23°C; T2: 15 days at 23°C/27°C, T3: 15 days at 27°C/31°C and T4: 15 days at 19°C/23°C). Physiological and

biochemical parameters were evaluated and revealed that *C. canephora* responds differently from other cultivars from T2 to T4. Through RNASeq analysis we identified eight temperature-responsive miRNAs and explored their expression patterns to correlate with possible functions in thermoregulation. The miRNAs expression varied according to the cultivar and heat treatment, being induced or repressed by the temperature elevation. Between these miRNAs, we selected one novel and another conserved, *car-miR060-5p* and *car-miR828a-5p* respectively, to explore their potential roles hypothesized to be related to temperature tolerance in coffee. In silico analysis of these miRNAs targets, suggested the regulation in particular pathways, such as the γ-aminobutyric acid (GABA) for car-*miR060*-5p and anthocyanin for *car-miR828a-5p*. The novel miRNA was related to the reduction in GABA levels in coffee plants under moderate stress, through the regulation of a related transporter. Whereas, the overexpression of the conserved miRNA, car-miR828a-5p, in Arabidopsis, increases the tolerance of Arabidopsis under severe heat stress. In conclusion, our results increase the knowledge about siRNAs in response to warmth and possibly related to tolerance in coffee plants, which will help the development of new technologies useful to adequate crop genotypes to climate changes.

KEYWORDS

Climate change; Coffee plants; miRNAs; post-transcriptional regulation; protective molecules, siRNAs.

INTRODUCTION

The current scenario of climate change has worried researchers and producers around the world, due to the possible damage that the increase in temperature can cause to the quality and production of crops. Coffee is an important global commodity (ICO, 2022) that has been increasingly studied in relation to its management and development of more tolerant cultivars (TONIETTO et al., 2012), due to its sensitivity to high temperatures, more specifically of *C. arabica*, a species responsible for about 60% of world production (DAMATTA; RAMALHO, 2006; ICO, 2022).

In coffee plants, high temperatures associated with an intense water deficit can cause the death of pollen tubes due to dehydration, causing flower abortion (CAMARGO E CAMARGO, 2001) or the early opening of flower buds (DAMATTA ET AL., 2019), which ends up causing production losses. In addition, exposure of *Coffea arabica* to temperatures close to 30°C can result not only in growth suppression, but also in anomalies such as yellowing of leaves (DAMATTA; RAMALHO, 2006), changes in the leaf cell wall, with alterations in the organization of cell wall polysaccharides (LIMA et al., 2013), and growth of tumors at the base of the stem (FRANCO,

1958). Despite advances in studies on damage caused by heat stress, there are still gaps about which mechanisms act on tolerance to this type of stress in this culture, especially when it comes to molecular mechanisms and gene expression regulators.

MicroRNAs (miRNAs), for example, are an extensive class of endogenous small RNA molecules that act as post-transcriptional regulators of gene expression, degrading or blocking the target mRNA, regulating different processes such as development, differentiation, flowering, metabolism, and plant defense (ZHANG, 2015; VOINNET, 2009). Several studies have been characterized a number of microRNAs responsive to heat stress by high temperatures in different species, such as in *Populus tomentosa* (CHEN et al., 2012), *Arabidopsis thaliana* (GUAN et al., 2013), *Oryza sativa* (SAILAJA et al., 2014), *Saccharina japonica* (LIU et al., 2015) and *Solanum lycopersicum* (ZHOU et al., 2016) However, the stress response mechanisms and how these microRNAs act in a specific metabolic pathway have not been explored in coffee plants yet.

Considering the climate change panel and the current state of the art on the role of siRNAs in coffee plants, an in-depth study of these molecules in this genus is necessary. Furthermore, understanding the mechanisms and molecules that act in the regulation of heat stress can manipulate agronomic traits of commercial interest with more efficiency and precision. This work aimed to evaluate whether coffee miRNAs can play a role in heat stress tolerance in four different coffee cultivars, covering the two main commercially interesting species *Coffea arabica* and *Coffea canephora*, which vary in relation to heat stress sensitivity.

MATERIAL AND METHODS

Plant material and experimental design

The first experiment was conducted at the Laboratory of Molecular Physiology of Plants at the Federal University of Lavras. Four cultivars (cv.) of two coffee species were selected, one of *C. canephora* (Robusta tropical-RT), and three of *C. arabica* (Mundo Novo-MN, Arara-AR, and Catuaí Vermelho 144-CV). *C. canephora* is reported to be a more tolerant species than *C. arabica* under hea and CO_2 stress (RODRIGUES et al., 2016). Mundo Novo is reported as a cultivar sensitive to different types of stress, while Catuaí (most cultivated in Brazil) and Arara would be more tolerant (ALMEIDA et al., 2018; MEGURO; MAGALHÃES, 1983; PENNACCHI, 2022). The characteristics and origin of these cultivars are shown in table S1. 8-month-old seedlings were grown in 19x32 cm bags with soil fertilized with NPK (4:14:8) and acclimated for two months in a greenhouse. During this period, the plants underwent foliar and soil analyses and received an application of the fungicide Orkestra® (BASF) to control rust and one fertilization according to recommendations by Malavolta et al. (1997). After acclimatization, twenty plants of each cultivar were distributed in a growth chamber (CMP6050 – Conviron®) in randomized blocks and

submitted to temperature ramp with a 4:4:5 factorial, totaling 80 plants. The growth conditions (Fig. S1) were T1: 15 days at 19°C/23°C (dark/light); T2: 15 days at 23°C/27°C, and T3: 15 days at 27/31°C. Considering the age difference of the plants from the beginning to the end of the experiment was 45 days, T4 heat treatment was repeated at the end of the experiment (T4). Temperatures used in the treatments were defined based on the annual temperature averages of Brazil and in three cities representing three coffee producing regions in Minas Gerais - BR in 2018 (Table S2), in addition to data from the IPCC (2014), which warn that the average global temperature could rise by up to 4°C by the end of the century. During the experiment, the only variable between treatments was temperature. The light was fixed at 300 $\mu molm^{-2}s^{-1}$, CO_2 was turned off, the humidity was maintained at 60% and the photoperiod was 12/12 h. The plants were kept well hydrated throughout the experiment.

To evaluate the effect of temperature on transgenic and mutant plants (described below), we conducted a second experiment with seedlings in 0,5 Murashige and Skoog culture medium with phytagel and 1 % sucrose. Seeds of wild-type (WT) plants, knocked out Ath-miR828 (KO-miR828 - SALKseq_021292), and 35S-Car-miR828-T3 (developed in this work), were cultivated for 15 days at 22°C/17°C (T1), followed by 5 days at 30°C/25°C (T2) and another 3 days at 35°C/30°C (T3). During the experiment, the only variable between treatments was temperature. The light was fixed at 150 µmol m-2s-1, CO2 was turned off, the humidity was turned off and the photoperiod was 16h/8h. At the end of each treatment, three biological replicates were collected for gene expression analysis. In addition, at the end of the experiment, the survival rate was calculated and expressed in percentages. The expression analysis was performed according to the item “Analysis of gene expression”, and RNA extraction was performed according to (OLIVEIRA et al., 2015) with reference genes TIP41 and PUX7 (ŠKILJAICA et al., 2022). A. thaliana car-miR828 targets were predicted according to the item “Targets prediction”, using GCF_000001735.4_TAIR10.1_rna file, available at https://www.ncbi.nlm.nih.gov/assembly/GCF_000001735.4/.

Physiological parameters and plant material

Two fully expanded leaves from the middle third of the plants were collected from four biological replicates of each cultivar, at the end of each heat treatment, at three different times (06:00h, 10:00h, and 18:00h). The plant material was immediately frozen in liquid N_2 and stored in an ultrafreezer at - 80°C until use. Leaf temperature (LT) and stomatal conductance measurements were also conducted on two fully expanded leaves of the middle third of the plants at the end of each heat treatment at the same times as the sampling. To measure leaf temperature, we used a digital infrared laser thermometer, measurements were made at 20cm from each leaf. Stomatal conductance (g_s) was measured with an LI-1600 Steady porometer (LI-COR).

Biochemical analyzes

Macro and micromolecules extraction were performed according to (ZANANDREA et al., 2009), with 1g of fresh matter macerated in liquid N_2. Total soluble sugars quantification was performed by the Antron method, according to (YEMM; WILLIS, 1954). Proline quantification was measured according to (BATES; WALDREN; TEARE, 1973). Lipid peroxidation (MDA) quantification was performed by the thiobarbituric acid method according to (BUEGE; AUST, 1978). Hydrogen peroxide (H_2O_2) quantification was performed according to (VELIKOVA; YORDANOV; EDREVA, 2000) and total proteins quantification was measured by the (BRADFORD, 1976) method. Anthocyanin quantification was performed according to (LANGE; SHROPSHIRE; MOHR, 1971) with minor changes. 1 mL of buffer (propanol-HCL-H2O = 18:1:81 vol %) was added to 100 mg of fresh frozen material. After vortexing for 1 min, the mixture was incubated at 99 °C for 3 min, followed by 25 °C for 24 h, and centrifuged for 20 min at 13000g. The supernatant was collected and the absorbance was measured at 535 nm and 650 nm and normalized to dry weight. γ-aminobutyric acid (GABA) quantification was performed according to Watchararparpaiboon; Laohakunjit; Kerdchoechuen, (2010).

Selection of miRNAs responsive to warm, RNAseq expression analysis and coffee genotyping

For gene expression analysis by RT-qPCR, novel and conserved miRNAs identified in *C. arabica* (unpublished) were selected. The precursors from the miRNAs were used as a reference for the alignment of single-end reads RNA-seq libraries from young plants of *C. arabica* (cv. Acauã and Catuaí) cultivated under different temperature conditions (SRA under BioProject ID PRJNA609253) (DE OLIVEIRA et al., 2020), using the Bowtie 2 program (LANGMEAD; SALZBERG, 2012) with the parameter very sensitive local. The quantification of the aligned precursors was performed with the Express program (ANDERS; PYL; HUBER, 2015). The total number of reads aligned was extracted using in-house scripts. Then the data were processed by CPM (counts per million) using the edgeR package (ROBINSON; MCCARTHY; SMYTH, 2010) from the R statistical environment (R CORE TEAM, 2022). Figures of the presence and absence of these microRNAs were generated in the RNA-seq libraries. All precursors that had more than 10 reads aligned were considered present and those that had less than 10 were considered absent. The figures were generated in R, with the Ggplot package (GINESTET, 2011). Some groups of microRNAs that were present only in libraries of interest were selected. Subsequently, a literature search was conducted in studies with heat stress in different species, to cross them with the RNA-seq data. The selected miRNAs were tested to the efficiency of the primers to proceed to the gene

expression analysis. The sequence of the novel miRNA was confirmed in the genome of *C. arabica* and its parents (*C. eugenioides* and *C. canephora*) by PCR.

Targets prediction

MicroRNAs targets selected in the previous step were predicted using the psRNATarget tool (http://plantgrn.noble.org/psRNATarget/) with default parameters. The microRNAs were aligned to the Johns Hopkins University *C. arabica* genome RNA library (PRJNA506972) available at the NCBI (https://www.ncbi.nlm.nih.gov/genome/?term=Coffea+arabica).

Gene ontology analysis for miRNA targets

Predicted sequences of *C. arabica* proteins were downloaded from its genome annotation available at the National Center for Biotechnology Information (NCBI: https://www.ncbi.nlm.nih.gov/assembly/GCF_003713225.1/#/def). To obtain information about the respective proteins families, domains, and functional annotations, these sequences were loaded into the Blast2GO suite (GÖTZ et al., 2008) and utilized in alignments against the RefSeq protein database (https://www.ncbi.nlm.nih.gov/refseq/) using BLASTp (ALTSCHUL et al., 1990) with 1 e^{-5} threshold e-value, as well as in InterProScan analyses. Posteriorly, were obtained and annotated GO classifications for each blast hit using Blast2GOs Mapping and Annotation functions with default parameters. The results merged from both analyses were used in the subsequent enrichment analysis. The topGO package (ALEXA; RAHNENFUHRER, 2022) from the R Bioconductor project (GENTLEMAN et al., 2004) was used to perform the enrichment analysis of retrieved GO terms, with p-values calculated using Fisher's exact test and corrected for multiple testing with the appliance of a false discovery rate (FDR). GO structure figures were generated from the topGO package and the WEGO online tool (YE et al., 2018).

Gene expression analysis

For RNA extraction, 200 mg of plant material macerated in liquid N2 was used. RNA extraction was performed using the PureLink® Plant RNA Reagent (Thermo Fisher Scientific) following the manufacturer's recommendations. RNA integrity was analyzed visually on 1 % agarose gel, and RNA content, as well as the presence of contaminants, were assessed by spectroscopy (OD260/280 and OD260/230 > 1.8) (NanoVue GE Healthcare). To eliminate residual DNA, 5 µg of RNA was treated with DNA-free™ DNA Removal Kit (Invitrogen). Primers for stem-loop RT-PCR and RT-qPCR were designed for each miRNA according to (VARKONYI-GASIC et al., 2007) The sequences of the designed primers (Table S4) were analyzed for the probability of secondary structure formation using the OligoAnalyzer™ Tool

(https://www.idtdna.com/pages/tools/oligoanalyzer). The cDNA was synthesized by the pulsed stem-loop RT-PCR method (VARKONYI-GASIC et al., 2007). The first strand of cDNA was synthesized by reverse transcription using 1 µg of RNA treated with the Improm-II Reverse Transcriptase kit (Promega) following the manufacturer's instructions. The analysis of primer efficiency and gene expression by RT-qPCR was performed by the SYBR Green detection system in the Rotor-Gene Q Real-Time PCR equipment (Qiagen). A reaction volume of 15 µL was used: 7.5 µL of SYBR-green (QuantiNova SYBR Green PCR Kit - Qiagen), 1.5 µL of cDNA at 10 ng/µL, primers ranging from 1-2 nmol, and RNase-DNase free water. Amplification was performed according to the manufacturer's recommendations and completed by a melting curve analysis to assess the specificity of the reaction. Relative bend differences were calculated based on the $\Delta\Delta^{CT}$ method (PFAFFL, 2001). The samples were analyzed in technical duplicates with four biological replicates. *MDH* and *UBQ* were used as reference genes (MARTINS et al., 2016). This experiment followed the *Minimum Information for the Publication of Quantitative Real-Time PCR Experiments* (MIQE) (BUSTIN et al., 2009), information is available in table S5.

Plant transformation and obtaining transgenic

After analyzing the data from the experiment, we studied the role of *car-R828-5p* heat stress through the resistance of *A. thaliana* transgenic overexpressing this miRNA. DNA for transformation was obtained from PCR of 300 bp upstream and 300 bp downstream fragments of the miRNA precursor (Supplemental Data Set 1). Overexpression was performed using the Gateway Recombination System (Invitrogen™). The DNA sequence was cloned into the pENTR™/D-TOPO™ Cloning Kit (Invitrogen™) entry vector. Chemically competent *Escherichia coli* cells One Shot™ TOP10 (Invitrogen™) were used for transformation. Confirmed sequences were recombined with the destination vector (pK7WG2.0) driven by the cauliflower mosaic virus 35S promoter, with the Gateway™ LR Clonase™ II Enzyme mix kit (Invitrogen™). Destination vector was used to transform chemically competent cells of *A. tumefaciens* GV3101. *A. thaliana* plants were transformed with *Agrobacterium tumefaciens* GV3101 by the floral dip method according to Zhang et al. (2006). Transgenic plants were selected in 0,5 Murashige and Skoog medium with phytagel, 1 % sucrose, and 50 ng/L kanamycin antibiotic. Plants that grew in the medium were also confirmed by PCR, the DNA extraction was performed according to (OLIVEIRA et al., 2015). Plants growing on the medium were transferred to a mixture of substrate and vermiculite (2:1). Homozygous lines were obtained by self-fecundation until the T3 generation.

Statistical analysis

For physiological and biochemical analysis, the data obtained were subjected to normality assessment using the Shapiro-Wilk ($P \geq 0.05$) and Bartlett's ($P \geq 0.05$) (BERGMAN; ROEDIGER, 1999). Interactions between treatments (temperature and time) and cultivars (Mundo Novo, Catuaí Vermelho, Robusta Tropical, and Arara) were included in the model. When differences were detected by ANOVA, the means were compared by Tukey ($P < 0.05$), using the R software (R CORE TEAM, 2022) and the ExpDes (FERREIRA; CAVALCANTI; NOGUEIRA, 2021) and Ggplot (GINESTET, 2011) packages. When the data residuals did not show a normal distribution, the Kruskal-Wallis ($P \leq 0.05$) test was used. Multivariate methods have numerous appealing properties for data integration. To verify the overall behaviours of the cultivars in relation to the full set of variables, we used the reduction of data dimensionality using Partial Least Squares – Discriminant Analysis (PLS-DA) (LÊ CAO; BOITARD; BESSE, 2011). We elaborated a graphic biplot, using the Factoextra (MUNDT, 2017) and ggfortify packages (TANG; HORIKOSHI; LI, 2016). For the gene expression statistical analysis, the expression rate and the confidence intervals were calculated according to the method proposed by (STEIBEL et al., 2009), which considers the linear mixed model. given by the following equation:

$$y_{ijklm} = \mu + TG_{ijk} + I_l + e_{ijklm}$$

where y_{ijklm} is the Cq (quantification cycle) obtained from the thermocycler software for the k_{th} gene (reference or target) from the m_{th} well, corresponding to the l_{th} plant subject to the i_{th} treatment at the j_{th} cultivar for the experiment I; TG_{ijk} is the effect of the combination of the i_{th} treatment (TR1 and TR3) at the j_{th} cultivar (MN, RT, and CV), in the expression of the gene k (reference or target). The analysis of data generated in the second experiment (using Arabidopsis transgenic and mutant) followed the same model as the first experiment, differing in the number of treatments (TR1, TR2, and TR3), and instead of cultivars, we have lines (WT, KO-Ath-miR828, and 35S-car-miR828). Graphics were generated using the SigmaPlot software version 12.3.

RESULTS

Physiological and biochemical parameters of coffee plants under different heat treatments

To understand the behavior of contrasting coffee cultivars in response to different heat treatments and discriminate tolerant and sensitive groups together which components influence more in this discrimination, leaf temperature (LT), stomatal conductance (g_s), total soluble sugars (TSS), reducing sugars, starch, Lipid peroxidation (MDA), hydrogen peroxide (H_2O_2), total proteins, and proline data, were analyzed at different times. Based on a sPLS-DA multivariate analysis, it is observed that from T2, *C. canephora* RT begins to separate from the other cultivars of *C. arabica* (Fig. 1). Even in T4, when the temperature returns to the initial condition, the variables still presented variation that contributed to this discrimination.

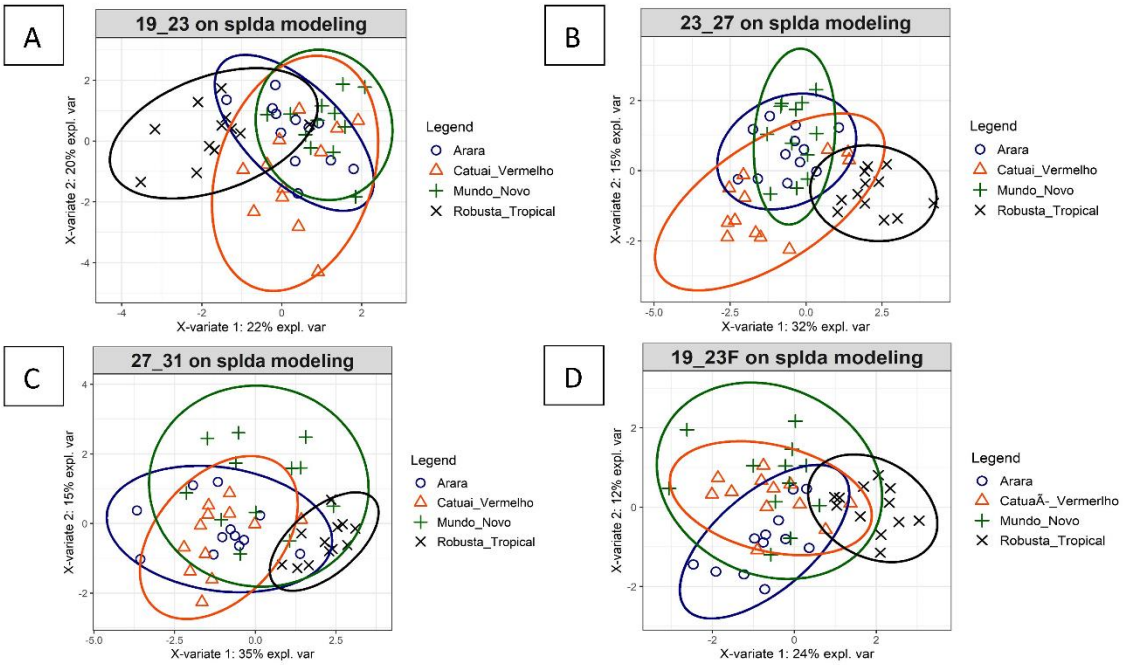


Fig. 1. Confidence ellipses for coffee cultivars (Arara, Mundo Novo, Catuaí Vermelho, and Robusta Tropical) after basic PLS-DA operate on the data leaf temperature, g_s, proline, MDA, hydrogen peroxide, total soluble proteins, starch, total soluble sugars, and reducing sugars in three collection times (6 a.m., 10 a.m., and 18 p.m.) cultivated under four heat treatments: T1 - 15 days at 19°C/23°C (A); T2 - 15 days at 23°C/27°C (B); T3 - 15 days at 27/31°C (C); and T4 - 15 days at 19°C/23°C (D). The first two components were used as axes.

For univariate analysis of the cited data, we selected the time of 10h due to the higher metabolic activity at this time (Fig 2). However, a univariate analysis for all times and treatments is available in Fig S3. Regarding LT (Fig 2A), in T1, T2, and T4 there was no significant difference between cultivars. At T3, CV and RT showed higher LT values, followed by MN and AR, which showed no differences between them. The highest LT values were observed in T3 for all cultivars. For g_s (Fig 2B), in T1 and T2 there were no significant differences between cultivars. At T3, AR and CV showed the highest g_s values, followed by RT and MN, which did not show any difference between them. The highest values of g_s for all cultivars were observed in T2 and T3.

For MDA (Fig 2C), in T2 the *C. arabica* cultivars showed higher concentrations of MDA, while RT had a higher concentration in T1. However, there was no significant difference between the cultivars considering the treatments. As for H_2O_2 (Fig 2D), in T1, RT showed the lowest concentration when compared to the other cultivars, while in T2 and T3 there were no significant differences between the cultivars. At T4, CV showed the highest concentration of H_2O_2, followed by AR and MN. RT showed the lowest H_2O_2 values at T2. And again, in T2 the *C. arabica* cultivars showed their peak of H_2O_2, while for RT the highest concentration was in T3.

Considering total protein content, MN and AR showed protein higher concentration at T2, while CV and RT at T1. In T1, there is no significant difference between cultivars. In T2, AR, MN,

and RT present higher concentrations of proteins than CV. In T3, MN, CV, and RT present higher concentrations than AR. Finally, in T4 there was no difference between the cultivars in protein concentration. For proline, MN peaked at T4, while CV, AR, and RT peaked at T1. There was no difference between cultivars in any HT.

Regarding the sugar contents, for TSS, RT shows a decrease along the treatments, with a higher concentration in T1. CV has its higher concentration in T3, while MN and AR are in T2. At T2, AR has the highest concentration of TSS, followed by CV and MN, which do not differ from each other, and RT, which has the lowest concentration. In T3 and T4, RT has the lowest concentrations of RT, while there is no difference between the other cultivars in this HT. About RT , AR and CV present their peak concentration at T2, followed by MN and RT at T1. There is no difference between the cultivars in T1, T2, and T4, while in T3, CV has the highest concentration, followed by the other cultivars that do not differ from each other. For starch content, MN, AR, and AR show higher concentration at T4, while for RT, the peak occurs at T1, followed by a drop at T2 and constant until T4. In T2 and T3, RT has a lower starch concentration when compared to other cultivars. Already in T4 AR has the highest concentration of starch, followed by CV and MN that do not differ, and RT, with the lowest concentration.

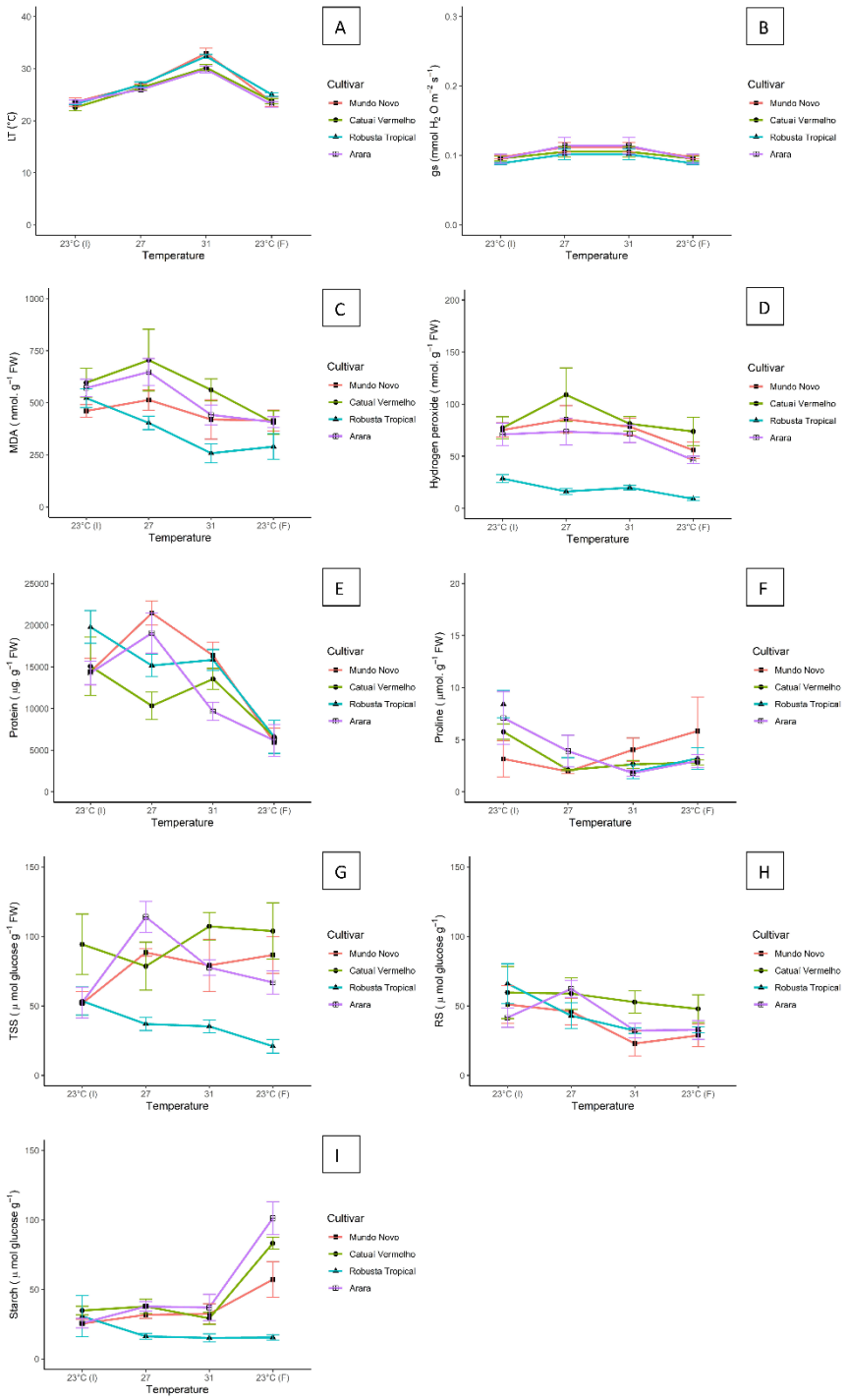


Fig. 2. Physiological and biochemical parameters of coffee leaves from four cultivars (Arara, Mundo Novo, Catuaí Vermelho, and Robusta Tropical) cultivated under four heat treatments (T1: 15 days at 19°C/23°C; T2: 15 days at 23°C/27°C; T3: 15 days at 27/31°C; and T4: 15 days at 19°C/23°C). Leaf temperature (A), g_s (B), MDA (C), hydrogen peroxide (D), protein (E), proline (F), total soluble sugars (G), reducing sugars (H), and starch (I). Differences between means were calculated by Tukey test at $P<0.05$ and are described in the text.

Identification of miRNAs responsive to warm and their expression profiles in coffee plants under temperature elevation

Alignment of new and conserved coffee miRNA precursors against RNA-seq libraries (Fig S2) allowed classifying some groups as candidates for gene expression analysis, based on the fact that these were present only in libraries of optimal temperature (Op) or only in heating (Wa), for one cultivar or both (Acauã and Catuaí). These data were crossed with information found in the literature for miRNAs responsive to different heat treatments in different species (Table S3) and we have norrowed to eight candidates for miRNAs possibly responsive to temperature elevation in coffee plants: *car-miR396a-5p, car-828a- 5p, car-miR172d-1-3p, car-156f-3p, car-miR 166-1-3p, car-miR 393a-5p, car-miR 393a-5p, car-miR157-5p* and a novel miRNA: *ca-miRNA-060-5p*. Gene expression analysis was performed comparing two most contrasting treatments, T1 and T3. In addition, AR was excluded from the analysis from now on, as it does not differ from other cultivars of *C. arabica* based on the parameters evaluated until now.

For *car-miR396a-5p* (Fig. 3A), we observed the expression of this miRNA was lower in RT independent of treatment. However, there were no significant differences in T1 and T2 in each cultivar. For *car-miR828-5p* (Fig 3B), we observed there is a reduction in the expression of miRNA in T2 in relation to T1, being reduced 1,3 times in MN in T1 in relation to T2 and 3 times in CV in T2 in relation to T1. When evaluating the expression of *car-miR060-5p* (Fig. 3C), we noticed that its expression was at least 600 times higher in *C. arabica* cultivars, MN and CV than in RT. Furthermore, in CV, there was 600 times increase in the expression of this miRNA in T3 compared to T1. For *car-miR166-3p* (Fig. 3D), we observed that it was 2 times more expressed in T3, in general, this miRNA is also less expressed in RT compared to other cultivars independent of treatment. Regarding *car-miR172-d-1-3p* (Fig 3E), we observed an increase in the expression of this miRNA about 1.8 times in RT on T2 in relation to T1 and 4 times in CV on T2 in relation to T1. For *car-miR156f-3p* (Fig. 3F), we noticed a similar behavior to *car-miR172d-1-3p,* with an increase in expression from miRNA by about 3 times for both in RT and CV. As for *car-miR393a-5p* (Fig. 3G), T3 induced 4 times the expression of this miRNA in RT and 1.2 times in CV. It was also observed that this miRNA is higher expressed in MN independent of TH, when compared to other cultivars. The expression of *car-miR157-5p* (Fig. 3H) was 9 times increased in the T2 compared to T1 in RT, while in CV there was a reduction of about 0.8 times in T2 in relation to T1.

Overall, we observed that all miRNAs studied here were responsive to the temperature increases. Interestingly, this expression profile varies according to the species and cultivar studied. We also observed that most of the studied miRNAs were induced by temperature increase, except the *car-miR828a-5p,* which was repressed. Furthermore, *car-miR060-5p,* showed much lower expression values in RT, a *C. canephora* cultivar, a species described as more tolerant to heat stress, than *C. arabica* cultivars described as more sensitive to heat stress.

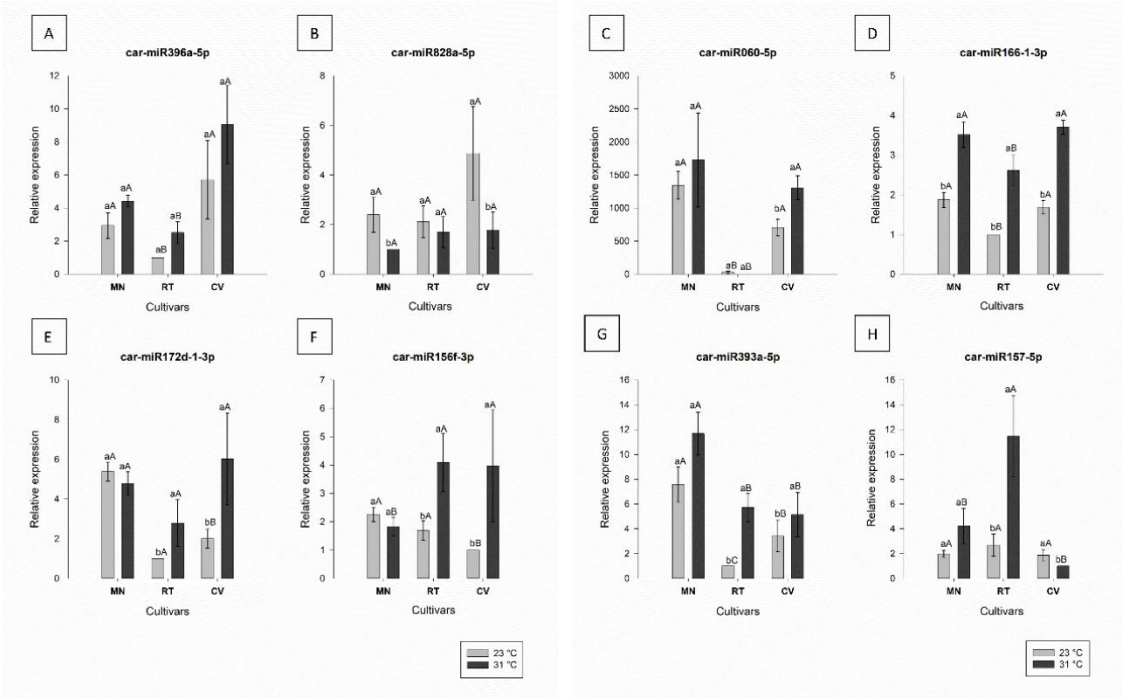


Fig. 3. Relative expression of *car-miR396a-5p* (A), *car-miR828a-5p* (B), *car-miR060-5p* (C), *car-miR166-1-3p* (D), *car-miR172d-1-3p* (E), *car-miR156f-3p* (F), *car-miR393a-5p* (G), and *car-miR157-5p* (H) from coffee leaves of cultivars Mundo Novo (MN), Robusta tropical (RT), and Catuaí Vermelho (CV) under two heat treatments (T1: 15 days at 19°C/23°C and T3: 15 days at 27/31°C). Lowercase letters differ the means for each cultivar under different treatments, and uppercase letters differ the means for cultivars under the same treatment at $P<0.05$ of significance.

Based on these data, we decided to evaluate the presence of this miRNA in *C. arabica* and its parents' genomes, *C. canephora* and C. *eugenioides*. Bearing in mind that *C. arabica* is an amphidiploid and presents the set of chromosomes from both parents (LASHERMES et al., 1999). Searches in the databases for the precursor of this miRNA show that it is not present in the genome of *C. canephora*, only in the genome of *C. arabica*, in the chromosome of *C. eugenioides*. However, PCR analysis for the precursor of this miRNA shows that it is present in the three genomes (Fig 4). The lack of this precursor in *C. canephora* genome can be due to an error in the assembly of the genome, since it is very hard to cover repetitive regions.

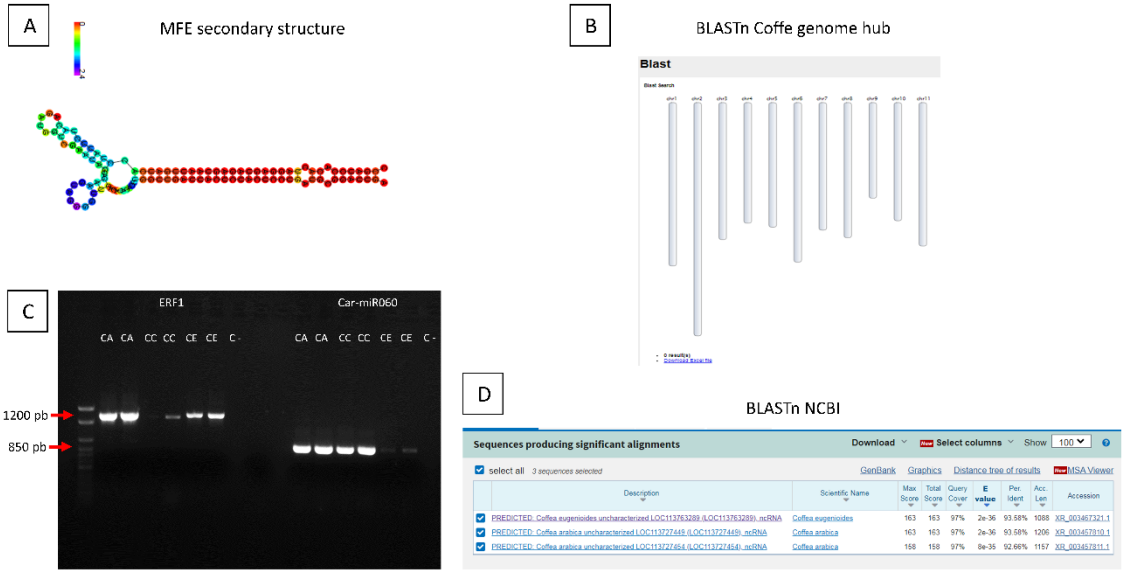


Fig. 4. Secondary structure of *car-miR060-5p* precursor (A). Not found hits for BLASTn of *car-miR060-5p* in the *C. canephora* genome available from the Coffee Genome Hub (http://www.coffee-genome.org/) (B). PCR of *car-miR060-5p*, with *CaERF1* as positive control from *C. arabica* (CA), C. *canephora* (CC), and *C. eugenioides* (CE) DNA at 50 ng/µl (C). Found hits for BLASTn of *car-miR060-5p* in the *C. arabica* genome available in NCBI (https://www.ncbi.nlm.nih.gov/assembly/GCF_003713225.1) (D).

Prediction of miRNAs targets and GO term enrichment analysis

Target prediction allowed us to confirm the canonical nature of conserved coffee plants' miRNA targets (Table S4). Based on the gene expression, we decided to explore *car-miR828a-5p* and *car-miR060-5p* regarding their role in the response to temperature increases in coffee plants. 164 targets were identified for *car-miR828a-5p*, with *Coffea arabica transcription factor WER-like* (LOC113741846) being the target with the lowest Expectation (-1.0). For *car-miR060*-5p, 165 targets were predicted, being *C. arabica probable GABA transporter 2 transcript variant X1*(LOC113703510), the target with the lowest expectation (-1.0). To get an idea of what biological processes these miRNAs were involved in, we performed an enrichment analysis of gene ontology (GO) terms (Supplemental DATA 2). Among the enriched GO terms for biological processes, we highlight the *alpha-amino acid biosynthetic process* for *car-miR060-5p* targets and *negative regulation of secondary metabolite biosynthetic process* and *flavonol biosynthetic process* for *car-miR828-5p*, which strengthens the hypothesis that the targets of these miRNAs are involved in these processes.

Previous studies demonstrate *WER-like* transcription factor, a transcription factor of the *MYB* family, is related to anthocyanin production in other species. In *H. macrophylla*, the WER-like transcription factor, together MYB114 and WDR68, act as potential regulators of flower color formation, by inhibiting or promoting the synthesis of anthocyanins, negatively regulating *DFR* expression (PENG et al., 2021). In addition, rice mutants that produce more anthocyanin show

greater resistance in seed germination under heat stress. Heat-stressed alfalfa plants accumulated more anthocyanin (MARTINS et al., 2016), as did sugarcane (WAHID, 2007).

γ-aminobutyric acid (GABA) comprises a significant fraction of the free amino acid pool in plant cells and has been reported as a molecule that can mitigate the effect of heat stress in different species, such as rice, (NAYYAR et al., 2014), bentgrass (LI et al., 2019) and sunflower (ABDEL RAZIK et al., 2021). Based on this information, we decided to quantify the anthocyanin and GABA amounts in coffee cultivars MN, AR, and RT under treatments T1 and T3 in order to understand the relationship between miRNAs and the production of these molecules in response to temperature elevation.

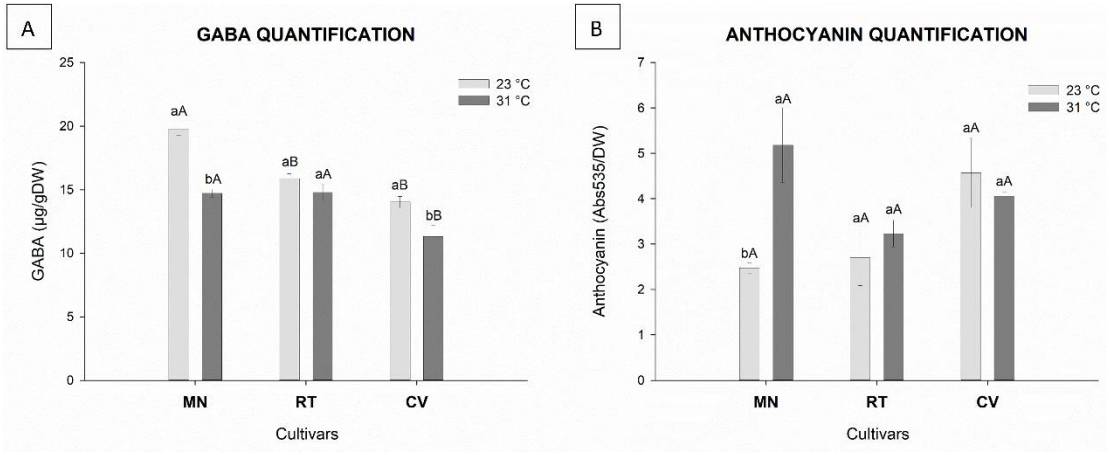


Fig. 5. GABA (A) and anthocyanin (B) quantification in coffee leaves of cultivars Mundo Novo (MN), Robusta tropical (RT), and Catuaí Vermelho (CV) under two heat treatments (T1: 15 days at 19°C/23°C and T3: 15 days at 27/31°C). Lowercase letters differ the means for each cultivar under different treatments, and uppercase letters differ the means for cultivars under the same treatment at P<0.05 of significance.

The GABA concentration was reduced in MN plants by about 5 times and about 3 times in the CV in T3 compared to T1, while in RT there was no significant difference between the THs (Fig. 5A). For anthocyanins, there was 2,5 times increase in MN in T2 compared to T1. For RT and CV, there were no significant differences between the treatments (Fig 5B).

Mutants overexpressing car-*miR828a*-5p showed higher survival rates in warm temperatures

To prove that *car-miR060-5p* and *car-miR828a-5p* act on heat stress tolerance, we sought to develop *A. thaliana* mutants overexpressing these miRNAs. However, for car-miR060-5p, we found that this miRNA cleaves the *C. arabica probable GABA transporter 2* target in the 5' UTR region of the mRNA. In *A. thaliana*, targets are not conserved due to the lack of conservation of the UTR region and, therefore, lack of homology with target genes in another species. The miR828 is conserved and the analysis of *car-miR828a-5p* targets identified 158 targets for this miRNA in the transcriptome of A. thaliana (Table S6). Among the predicted targets, *A. thaliana production of*

anthocyanin pigment 1 (*PAP1*), an important gene of the anthocyanin biosynthesis pathway was identified.

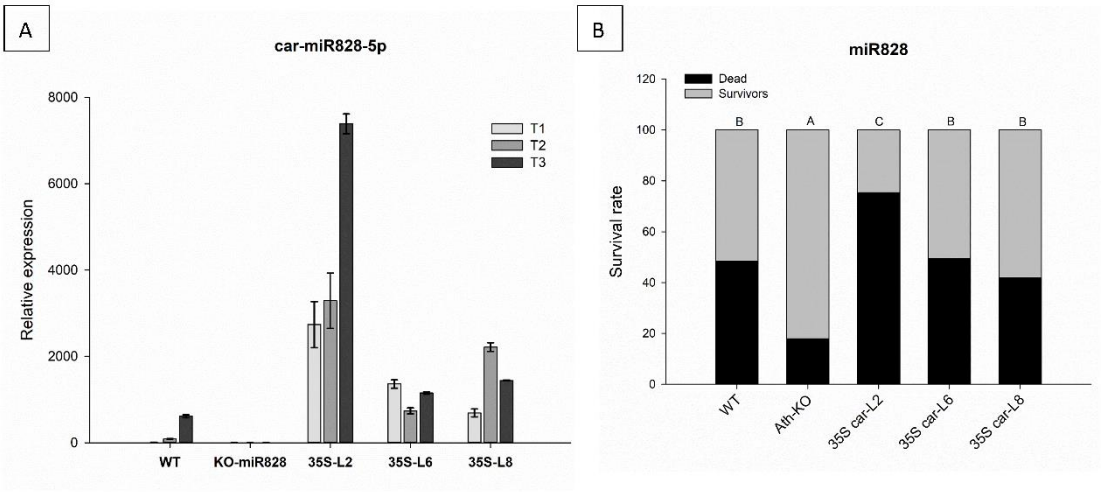


Fig. 6. Relative expression of *car-miR828a*-5p from *A. thaliana* seedlings wild type (WT), *ath-miR828* knocked out (KO), and three lineages overexpressing *car-miR828a-5p* (35S-L2, 35S-L6, and 35S-L8) under three heat treatments (T1: 15 days at 22°C/17°C; T2: 5 days at 30°C/25°C; and T3: 3 days at 35°C/30°C) (A). Lineages survival rate at the end of the experiment under the T3 treatment. Uppercase letters differ the survival rate for each lineage at the end of the experiment at $P<0.05$ of significance.

Gene expression analysis (Fig. 6A) showed that KO mutants for ath-miR828 did not express *car-miR828-5p*, while mutant strains that overexpressed *car-miR828a-5p* (35S-L2, 35S-L6, and 35S- L8) showed higher expression levels for this miRNA, with 35S-L2 being the strain that showed the highest expression in all treatments. At T1, L2 showed an expression about 2700 times greater than WT. Regarding the treatments, WT and L2 showed a similar expression profile, since there was an increase in miRNA expression in T2, followed by T3, in relation to T1. In KO there was no difference in expression between treatments, while in L8 a reduction in expression was observed in T2, followed by an increase in T3. L8 showed an expression profile opposite to L6, with an increase in expression in T2 followed by a decrease in T3. Regarding the death rate (Fig 6B), miR828 mutants that do not express *Ath-miR828* had a higher survival rate (82%), followed by WT (51%) and 35S-car-miR828-L6 strains (50) and 35S-car-miR828-L6 (58%), which did not differ, and finally, 35S-car-miR828-L2 (24%) had the lowest survival rate.

DISCUSSION

We observed that the physiological parameters LT and gs were higher for all the genotypes at higher temperatures than observed in other species such as coffee (DE OLIVEIRA et al., 2020) and wheat (SHARMA et al., 2015), but not so much for these parameters as responses are not

genotype-specific. Lower values of MDA and H_2O_2, were observed for RT, when compared to all other genotypes of *C. arabica*, in all treatments, suggesting less oxidative response for this species at high temperatures, as reported in previous studies (DAMATTA et al., 2019; reviewed by DAMATTA; RAMALHO, 2006). In general, for all the parameters evaluated, *C. canephora* behaves differently from *C arabica* cultivars as demonstrated by the multivariate analysis at higher temperatures.

Elevated temperature affects metabolic pathways mainly through oxidative damage to cells, affecting the levels of primary and secondary metabolites, which are of great biological importance (WAHID, 2007). γ-aminobutyric acid (GABA), a non-protein α-amino acid, has been gaining prominence as an important regulator of growth and stress tolerance, including heat tolerance in animals and plants (KINNERSLEY; TURANO, 2000).

Primary metabolites, originating from the initial reactions of carbon, participate directly in osmotic adjustment and the formation of cellular structures (WAHID, 2007). Studies in rice demonstrate that exogenously applied GABA can protect rice seedlings from heat stress through increased leaf turgor and antioxidant defense (NAYYAR et al., 2014). In Bermudagrass (*Cynodon dactylon*), the high concentration of ambient CO2 significantly promoted GABA shunt metabolism, contributing to better growth and heat tolerance (YU et al., 2017). Furthermore, the increase in endogenous GABA level by foliar application of GABA significantly alleviated heat-induced damage in undergrowth associated with improved osmotic adjustment, antioxidant capacity, and metabolic homeostasis (PAN et al., 2016).

Lower *car-miR060-5p* expression levels, regardless of the treatments imposed on a cultivar described as more tolerant to heat stress and increased expression of this miRNA in plants under temperature increase, demonstrates that this miRNA may be related to the sensitivity of *C. arabica* to thermal stress. When we relate the miRNA expression data with the quantification of GABA we noticed that there is an inverse relationship. Stressed plants of *C. arabica* showed lower levels of GABA under temperature increase, which suggests that when plants are under stress there is increased expression of car-miR060-5p that downregulates its target, a GABA transporter, influencing the transport of this molecule and the stress response.

Secondary metabolites are synthesized from the intermediates of primary carbon metabolism (WAHID, 2007). The formation of anthocyanins in plants, is involved in different physiological functions, such as alteration of hormonal signals, defense responses against injuries caused by ultraviolet radiation, and biotic and abiotic stresses (AKHTER et al., 2019). In Arabidopsis, miR828 targets the anthocyanin activator *AtMYB113*, a closely related paralog of *AtPAP1* and *AtPAP2*. Furthermore, miR828 also targets the non-coding transcript of *TAS4* and triggers the production of secondary siRNAs in phases. One of these siRNAs, TAS4-siR81(-), further silences

AtMYB113, *AtPAP1*, and *AtPAP2*. Consequently, overexpression of miR828 in Arabidopsis led to inhibition of anthocyanin production. However, in potatoes and grapes, miR828 and TAS4-siR81(-) activate anthocyanin biosynthesis by targeting MYB repressors. In this way, miR828 can act in the activation or repression of anthocyanins depending on which MYB target genes are co-expressed with them (LAFOUNTAIN; YUAN, 2021).

Our data suggest that *car-miR828a-5p* regulates a transcriptional factor *WER-like* gene, which acts in the positive regulation of anthocyanin biosynthesis (WAHID, 2007). We hypothesized that plants overexpressing this miRNA accumulate less anthocyanin due to increased target cleavage (WER-like). In sensitive coffee cultivars under temperature elevation, we observed an increase in anthocyanin content concomitant with a reduction in the expression level of *car-miR828a-5p*. Thus, the role of this miRNA in high-temperature tolerance was confirmed by the high death rate in plants that overexpressed *car-miR828a-5p* in *A. thaliana*.

CONCLUSIONS

In this study, we identified eight miRNAs responsive to temperature elevation in coffee plants using described sensitive and tolerant cultivars. From this, six conserved miRNAs and a novel one, were induced (*car-miR396a-5p*, *car-miR172d-1-3p*, *car-156f-3p*, *car-miR 166-1-3p*, *car-miR 393a-5p*, *car-miR 393a-5p*, *car-miR157-5p* and *car-miR060-5p*), while only one conserved miRNA (*car-828a-5p*) was repressed by temperature elevation. GABA transporter is a putative target of *car-miRNA-060-5p*, and increased expression of this miRNA in heat-stress sensitive plants could lead to the downregulation of this protective molecule. In addition, *A. thaliana* plants that overexpress *car-miR828a-5p* are more sensitive to heat when compared to KO and WT plants, which suggests that this miRNA acts on heat stress tolerance in coffee through the regulation of WER transcription factors, that act in the anthocyanin biosynthesis pathway. These results, unprecedented for coffee, may contribute to the development of new technologies based on the control of post-transcriptional regulation of protective molecules in a plant of global economic importance and whose production can be affected by climate change. In addition, such results can be extrapolated for other crops through stimulus of conserved pathways or even testing the exogenous application of synthetic miRNAs to action in interest targets.

ACKNOWLEDGEMENTS

We thank the Federal University of Lavras (UFLA/Brazil) and members of the Laboratory of Plant Molecular Physiology (LFMP, UFLA/Brazil), in particular Gabriel Lasmar dos Reis, Júlia de Carvalho Costa and Lillian Magalhães Azevedo for support in experiments and analysis. We thank Beatriz Gonçalves Pereira Costa for support in biochemical analysis. We also thank Instituto

Brasileiro de Ciência e Tecnologia do Café (INCT/Café), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG) for the financial support.

AUTHORS' CONTRIBUTIONS

R.R.O. and A.C-J. conceptualized the project. K.K.P.O. conducted all data research and analyses. J.O.S., I.S.S., and R. M.G. supported K.K.P.O in the biochemical and statistical analyzes. R.O.M and D.L. supported K.K.P.O in obtaining the mutants. G.C.R. and T.H.C.R. supported bioinformatic analyses. R.R.O. and C.N.F-B supervised the experiments and data analyses. K.K.P.O wrote the first version of the manuscript and R.R.O contributed to writing.

STATEMENTS AND DECLARATIONS

The authors declare that there are no conflicts of interest.

REFERENCES

ABDEL RAZIK, E. S. et al. γ-Aminobutyric acid (GABA) mitigates drought and heat stress in sunflower (Helianthus annuus L.) by regulating its physiological, biochemical and molecular pathways. **Physiologia Plantarum**, v. 172, n. 2, p. 505–527, 2021.

AKHTER, D. et al. A rice gene, OsPL, encoding a MYB family transcription factor confers anthocyanin synthesis, heat stress response and hormonal signaling. Gene, v. 699, p. 62–72, 30 maio 2019.

ALEXA, A.; RAHNENFUHRER, J. topGO: **Enrichment Analysis for Gene Ontology**. Bioconductor version: Release (3.15), 2022. Disponível em: <https://bioconductor.org/packages/topGO/>. Acesso em: 5 set. 2022

ALMEIDA, J. A. S. DE et al. Water stress in germination, growth and development of coffee cultivars. **Journal of Seed Science**, v. 40, p. 82–89, mar. 2018.

ALTSCHUL, S. F. et al. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, n. 3, p. 403–410, 5 out. 1990.

BATES, L. S.; WALDREN, R. P.; TEARE, I. D. Rapid determination of free proline for water-stress studies. **Plant and Soil**, v. 39, n. 1, p. 205–207, 1 ago. 1973.

BERGMAN, E. T.; ROEDIGER, H. L. Can Bartlett's repeated reproduction experiments be replicated? **Memory & Cognition**, v. 27, n. 6, p. 937–947, nov. 1999.

BRADFORD, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical Biochemistry**, v. 72, p. 248–254, 7 maio 1976.

BUEGE, J. A.; AUST, S. D. Microsomal lipid peroxidation. **Methods in Enzymology**, v. 52, p. 302–310, 1978.

BUSTIN, S. A. et al. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. **Clinical Chemistry**, v. 55, n. 4, p. 611–622, 1 abr. 2009.

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

CHEN, L. et al. Genome-wide identification and expression analysis of heat-responsive and novel microRNAs in Populus tomentosa. **Gene**, v. 504, n. 2, p. 160-165, 2012.

DAMATTA, F. M. et al. Why could the coffee crop endure climate change and global warming to a greater extent than previously estimated? **Climatic change**, v. 152, n. 1, p. 167-178, 2019.

DAMATTA, F. M.; RAMALHO, J. D. C. Impacts of drought and temperature stress on coffee physiology and production: a review. **Brazilian Journal of Plant Physiology**, v. 18, n. 1, p. 55–81, mar. 2006.

DE OLIVEIRA, R. R. et al. Elevated Temperatures Impose Transcriptional Constraints and Elicit Intraspecific Differences Between Coffee Genotypes. **Frontiers in Plant Science**, v. 11, 2020.

FERREIRA, E. B.; CAVALCANTI, P. P.; NOGUEIRA, D. A. **ExpDes.pt: Pacote Experimental Designs** (Portugues). Disponível em: <https://CRAN.R-project.org/package=ExpDes.pt>. Acesso em: 6 set. 2022

FRANCO, C. M. Influence of temperature on growth of coffee plant. New York: **IBEC Research Institute**, 1958.

GENTLEMAN, R. C. et al. Bioconductor: open software development for computational biology and bioinformatics. **Genome Biology**, v. 5, n. 10, p. R80, 15 set. 2004.

GINESTET, C. ggplot2: Elegant Graphics for Data Analysis. Journal of the Royal Statistical **Society Series** A, v. 174, n. 1, p. 245–246, 2011.

GÖTZ, S.; GARCÍA-GÓMEZ, J. M.; TEROL, J.; et al. High-throughput functional annotation and data mining with the Blast2GO suite. **Nucleic acids research**, v. 36, n. 10, p. 3420–3435, 2008. Oxford University Press.

GUAN, Q. et al. Heat stress induction of mi R 398 triggers a regulatory loop that is critical for thermotolerance in Arabidopsis. **The Plant Journal**, v. 74, n. 5, p. 840-851, 2013.

KINNERSLEY, A. M.; TURANO, F. J. Gamma Aminobutyric Acid (GABA) and Plant Responses to Stress. **Critical Reviews in Plant Sciences**, v. 19, n. 6, p. 479–509, 1 nov. 2000.

LAFOUNTAIN, A. M.; YUAN, Y.-W. Repressors of anthocyanin biosynthesis. **New Phytologist**, v. 231, n. 3, p. 933–949, 2021.

LANGE, H.; SHROPSHIRE, W.; MOHR, H. An Analysis of Phytochrome-mediated Anthocyanin Synthesis. **Plant Physiology**, v. 47, n. 5, p. 649–655, maio 1971.

LANGMEAD, B.; SALZBERG, S. L. Fast gapped-read alignment with Bowtie 2. **Nature Methods**, v. 9, 2012.

LASHERMES, P. et al. Molecular characterisation and origin of the Coffea arabica L. genome. **Molecular and General Genetics** MGG, v. 261, n. 2, p. 259–266, 1 mar. 1999.

LIMA, R. B. et al. Heat stress causes alterations in the cell-wall polymers and anatomy of coffee leaves (Coffea arabica L.). **Carbohydrate Polymers, 2nd European Polysaccharide Network of Excellence International Conference**, August 2011. v. 93, n. 1, p. 135–143, 1 mar. 2013.

LIU, F. et al. Conserved and novel heat stress-responsive micro RNAs were identified by deep sequencing in Saccharina japonica (Laminariales, Phaeophyta). **Plant, cell & environment**, v. 38, n. 7, p. 1357-1367, 2015.

MARTINS, M. Q. et al. Protective Response Mechanisms to Heat Stress in Interaction with High [CO2] Conditions in Coffea spp. **Frontiers in Plant Science**, v. 7, 2016.

MEGURO, N. E.; MAGALHÃES, A. C. Water stress affecting nitrate reduction and leaf diffusive resistance in Coffea arabica L. cultivars. **Journal of Horticultural Science**, v. 58, n. 1, p. 147–152, 1 jan. 1983.

NAYYAR, H. et al. γ-Aminobutyric Acid (GABA) Imparts Partial Protection from Heat Stress Injury to Rice Seedlings by Improving Leaf Turgor and Upregulating Osmoprotectants and Antioxidants. **Journal of Plant Growth Regulation**, v. 33, n. 2, p. 408–419, 1 jun. 2014.

OLIVEIRA, R. et al. An efficient method for simultaneous extraction of high-quality RNA and DNA from various plant tissues. **Genetics and Molecular Research**, v. 14, p. 18828–18838, 28 dez. 2015.

PAN, L. et al. Transcriptional Profiles of Drought-Related Genes in Modulating Metabolic Processes and Antioxidant Defenses in Lolium multiflorum. **Frontiers in Plant Science**, v. 7, p. 519, 25 abr. 2016.

PENG, J. et al. Exploring the Molecular Mechanism of Blue Flower Color Formation in Hydrangea macrophylla cv. “Forever Summer”. **Frontiers in Plant Science**, v. 12, 2021.

PENNACCHI, J. P. Café Arara: tudo sobre a variedade de café de alta qualidade. **Blog da Aegro**, 13 jul. 2022. Disponível em: <https://blog.aegro.com.br/cafe-arara/>. Acesso em: 6 set. 2022

ROBINSON, M. D.; MCCARTHY, D. J.; SMYTH, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. **Bioinformatics** (Oxford, England), v. 26, n. 1, p. 139–140, 1 jan. 2010.

SAILAJA, B. et al. Prediction and expression analysis of miRNAs associated with heat stress in Oryza sativa. 2013. **Rice Science**, v. 21, Issue 1, p. 3-12, 2014.

SHARMA, D. K. et al. Wheat cultivars selected for high Fv/Fm under heat stress maintain high photosynthesis, total chlorophyll, stomatal conductance, transpiration and dry matter. **Physiologia Plantarum**, v. 153, n. 2, p. 284–298, 2015.

STEIBEL, J. P. et al. A powerful and flexible linear mixed model framework for the analysis of relative quantification RT-PCR data. **Genomics**, v. 94, n. 2, p. 146–152, ago. 2009.

TANG Y, HORIKOSHI M, LI W. GGFORTIFY: Unified interface to visual-ize statistical results of popular R packages. **The R Journal**.; v.8, n.2, p.474, 2016.

TONIETTO, Â, et al. Proteomic analysis of developing somatic embryos of Coffea arabica. **Plant Molecular Biology Reporter**, v. 30, n. 6, p. 1393-1399, 2012.

VARKONYI-GASIC, E. et al. Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. **Plant Methods**, v. 3, n. 1, p. 12, 12 out. 2007.

VELIKOVA, V.; YORDANOV, I.; EDREVA, A. Oxidative stress and some antioxidant systems in acid rain-treated bean plants. **Plant Science - PLANT SCI**, v. 151, p. 59–66, 7 fev. 2000.

VOINNET, O. Origin, biogenesis, and activity of plant microRNAs. **Cell**, v. 136, n. 4, p. 669–687, 20 fev. 2009.

WAHID, A. Physiological implications of metabolite biosynthesis for net assimilation and heat-stress tolerance of sugarcane (Saccharum officinarum) sprouts. **Journal of Plant Research**, v. 120, n. 2, p. 219–228, 1 mar. 2007.

YE, J. et al. WEGO 2.0: a web tool for analyzing and plotting GO annotations, 2018 update. Nucleic Acids Research, v. 46, n. **Web Server issue**, p. W71–W75, 2 jul. 2018.

YEMM, E. W.; WILLIS, A. J. The estimation of carbohydrates in plant extracts by anthrone. **Biochemical Journal**, v. 57, n. 3, p. 508–514, jul. 1954.

YU, J. et al. Metabolic Pathways Involved in Carbon Dioxide Enhanced Heat Tolerance in Bermudagrass. **Frontiers in Plant Science**, v. 8, p. 1506, 2017.

ZANANDREA, I. et al. Tolerance of Sesbania virgata plants to flooding. **Australian Journal of Botany** - AUST J BOT, v. 57, 1 jan. 2009.

ZHANG, B. miRNA: a new target for improving plant tolerance to abiotic stress. **Journal of Experimental Botany**, 66, 1749–1761, 2015.

ZHOU, R. et al. Identification of miRNAs and their targets in wild tomato at moderately and acutely elevated temperatures by high-throughput sequencing and degradome analysis. **Scientific Reports**, v. 6, n. 1, p. 33777, set. 2016.

CONCLUSÕES E PERSPECTIVAS FUTURAS

O presente trabalho contribuiu com o aumento do conhecimento sobre siRNAs em cafeeiro, identificando miRNAs conservados e novos, bem como phasiRNAs que estão envolvidos em dois importantes processos biológicos para esta espécie, a reprodução e a tolerância ao estresse térmico. Um dos grandes problemas do cafeeiro está relacionado ao seu florescimento e maturação dos frutos que ocorre de forma desuniforme, bem como o estresse térmico que prejudica o florescimento e promove flores anormais. Com a identificação da fase que ocorre a meiose e quais siRNAs reprodutivos estão envolvidos nesse processo, assim como as enzimas envolvidas na biossíntese de siRNAs, objetivamos contribuir com as pesquisas no controle deste processo.

Da mesma forma, ao identificar miRNAs responsivos ao calor que estão envolvidos nos mecanismos de tolerância ao estresse térmico em cafeeiro, regulando moléculas do metabolismo primário e secundário, como GABA e antocianina, esperamos contribuir com a caracterização de funções dos miRNAs de cafeeiro e no desenvolvimento de novas tecnologias que mitigam os efeitos do aquecimento global em crops, como no desenvolvimento de produtos à base de miRNAs sintéticos que regulam moléculas da fisiologia do estresse, ou produtos que regulam a expressão deste miRNAs.