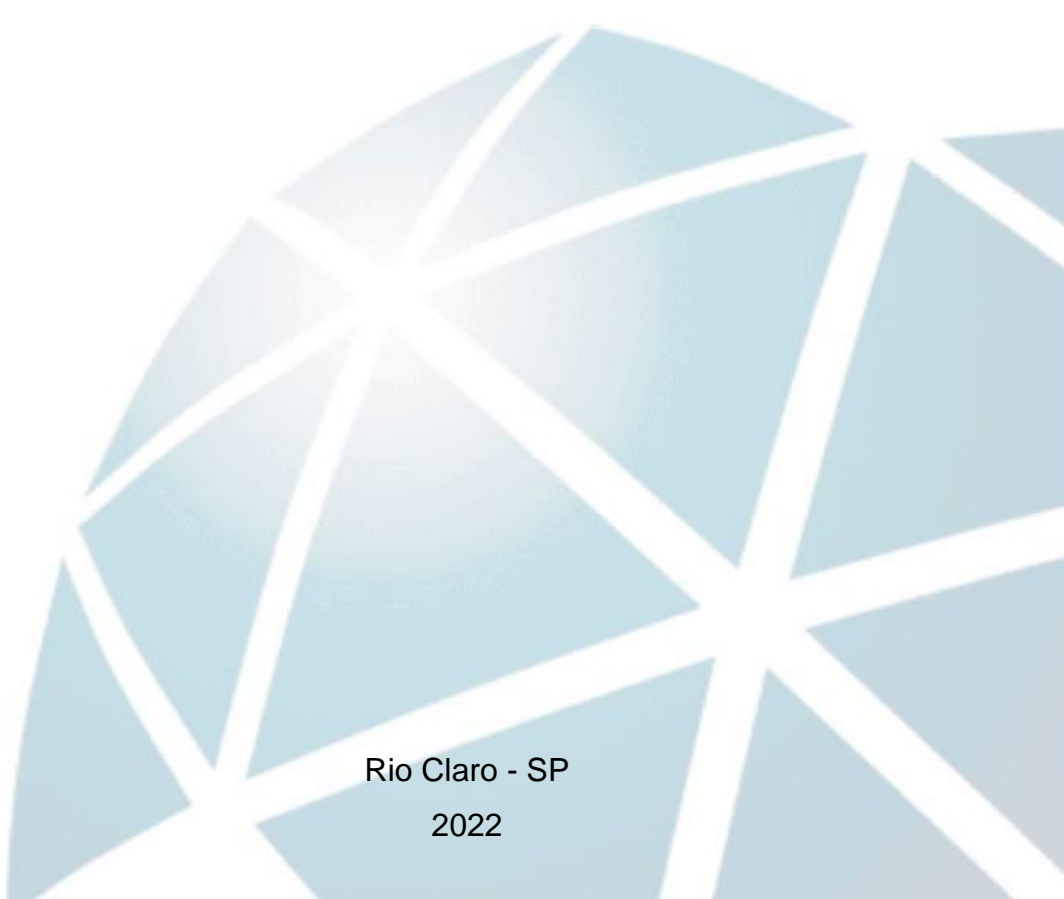

CIÊNCIAS BIOLÓGICAS

MATHEUS OZILIERO TAGLIAFERRO

**DECIFRANDO A BASE MOLECULAR DA
PRODUÇÃO DE COMPOSTOS BIOATIVOS EM
COFFEA ARABICA: PERFIL TRANSCRICIONAL
DE DITERPENOSINTASES EM TANDEM**



Rio Claro - SP
2022

MATHEUS OZILIERO TAGLIAFERRO

**DECIFRANDO A BASE MOLECULAR DE COMPOSTOS BIOATIVOS
EM *COFFEA ARABICA*:
PERFIL TRANSCRICIONAL DE DITERPENOSINTASES EM
TANDEM**

Trabalho de Conclusão de Curso apresentado ao Instituto de Biociências – Câmpus de Rio Claro, da Universidade Estadual Paulista “Júlio de Mesquita Filho”, para obtenção do grau de Bacharel e Licenciado em Ciências Biológicas

Orientador: Douglas Silva Domingues

Rio Claro – SP

2022

T126d Tagliaferro, Matheus Oziliero
Decifrando a base molecular de compostos bioativos em Coffea
arabica : perfil transcricional de diterpeno sintases em tandem /
Matheus Oziliero Tagliaferro. -- Rio Claro, 2022
46 p. : tabs.

Trabalho de conclusão de curso (Bacharelado e licenciatura -
Ciências Biológicas) - Universidade Estadual Paulista (Unesp),
Instituto de Biociências, Rio Claro

Orientador: Douglas Silva Domingues

1. Genética vegetal. 2. Diterpenos. 3. Sequenciamento de nucleotídeos
em larga escala. 4. Café. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Biociências,
Rio Claro. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

MATHEUS OZILIERO TAGLIAFERRO

**DECIFRANDO A BASE MOLECULAR DE COMPOSTOS BIOATIVOS EM
COFFEA ARABICA:
PERFIL TRANSCRICIONAL DE DITERPENOSINTASES EM TANDEM**

Trabalho de Conclusão de Curso apresentado ao Instituto de Biociências – Câmpus de Rio Claro, da Universidade Estadual Paulista “Júlio de Mesquita Filho”, para obtenção do grau de Bacharel e Licenciado em Ciências Biológicas

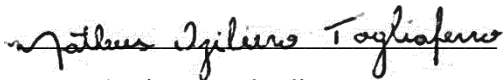
BANCA EXAMINADORA:

Prof. Dr. Douglas Silva Domingues

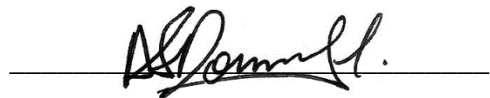
Profa. Dra. Karen Cristiane Martinez de Moraes

Prof. Dr. Gustavo Habermann

Aprovado em: 01 de dezembro de 2022



Assinatura do discente



Assinatura do orientador

AGRADECIMENTOS

Em primeiro lugar, gostaria de agradecer a meus pais pela oportunidade e incentivo em concluir um curso de graduação. Para eles, a educação sempre deteve reconhecida importância e afirmavam, com frequência e veemência, que “o melhor que podemos te dar é uma educação de qualidade”. Assim sendo, agradeço todo o suporte financeiro e emocional que viabilizaram essa conquista, a qual tenho orgulho de dizer que não é minha, mas sim nossa. Também gostaria de agradecer a meus avós, que estiveram sempre presentes, demonstrando apoio e contribuindo para tornar minha estadia em Rio Claro mais aconchegante, com muitos mimos.

Agradeço, também, aos meus colegas de curso, com quem convivi intensamente durante os últimos anos, pelo companheirismo e pela troca de experiências que me permitiram crescer não só como pessoa, mas também como formando. Em especial, agradeço à Laura, Letícia, Felipe e Daniel por todas as vivências e aprendizados que me proporcionaram ao longo dessa jornada. Nossa amizade, cultivada com muito carinho nesses últimos seis anos, certamente transpassará a graduação. Também agradeço a meus colegas de república pelo acolhimento, companheirismo e pelos conselhos em momentos de incertezas no decorrer da graduação.

Aos professores, agradeço os valiosos ensinamentos que contribuíram para meu aprimoramento enquanto profissional e pesquisador. Agradeço ao GECEA – Grupo de Estudos em Citogenômica e Evolução Animal, liderado pela Profa. Dra. Patricia Pasquali Parise Maltempi, do qual fiz parte por longo período de tempo durante a graduação e que contribuiu grandemente para minha formação. E em especial, agradeço ao grupo de pesquisa GeTransPlant – Grupo de Genômica e Transcriptômica de Plantas, liderado pelo Prof. Dr. Douglas Silva Domingues, que despertou meu interesse em Biotecnologia aplicada a plantas e me recebeu de braços abertos para o desenvolvimento do presente estudo.

Por fim, mas não menos importante, agradeço à UNESP, não apenas por minha formação enquanto profissional plenamente habilitado ao exercício da profissão, mas também por minha formação pessoal. Afinal, a Universidade não se resume ao mero cumprimento créditos acadêmicos, mas compreende a participação em diferentes espaços que contribuem para a formação humana, crítica e consciente de nosso papel no mundo.

A todos, meus mais sinceros sentimentos de gratidão!

RESUMO

Terpenos constituem uma classe de compostos químicos amplamente diversa em plantas, desempenhando funções relacionadas ao metabolismo primário e especializado desses organismos, estando, muitas vezes, restritos a linhagens ou espécies específicas. A biossíntese de terpenos é catalisada por terpeno sintases (TPS), uma família de enzimas cujas propriedades variam entre si. Tal variação relaciona-se a eventos de expansão aos quais os genes codificantes das TPS foram submetidos ao longo da evolução de plantas terrestres, os quais encontram-se frequentemente organizados em *clusters* como resultado de episódios de duplicação e neofuncionalização genética. Análises recentes detectaram relações de homologia entre *clusters* de diterpeno sintases em *Nicotiana tabacum* (Solanaceae), *Solanum lycopersicum* (Solanaceae) e nas espécies de cafeeiro *Coffea canephora* (Rubiaceae) e *Coffea arabica* (Rubiaceae), com variações quanto ao padrão de expressão e ao produto sintetizado. *C. arabica*, de histórica relevância econômica, carece das estruturas vegetativas nas quais os genes em questão são majoritariamente expressos nessas espécies de Solanaceae, o que avança indagações acerca do perfil de expressão e função desses genes nessa espécie de cafeeiro. Ademais, dada sua importância econômica, é substancial investigar os possíveis efeitos das mudanças climáticas previstas sobre elementos do metabolismo especializado de *C. arabica*. Neste estudo, o objetivo foi investigar o perfil transcricional de sete diterpeno sintases dispostas em tandem no genoma de *C. arabica*. Para apurar o perfil transcricional desses genes relacionados à síntese de diterpenos na espécie *Coffea arabica*, foram reanalisados dados de 48 bibliotecas de RNA-seq de *C. arabica*, oriundas de quatro estudos relacionados à resposta da espécie a estressores ambientais e ativação de defesas, bem como estágios da maturação do fruto. Análises de expressão diferencial revelaram que cinco genes apresentaram baixos níveis transcricionais, e dois mostraram expressão diferencial entre experimentos. Foram detectadas variações em relação aos órgãos em que esses genes são transcritos, bem como a modulação do nível transcricional em resposta às condições às quais as plantas foram submetidas. Esses resultados representam um passo inicial no delineamento do perfil transcricional e entendimento funcional desses genes na produção de metabólitos especializados em *C. arabica*, que podem deter interesse aplicado.

Palavras-chave: genética vegetal; diterpenos; sequenciamento de nucleotídeos em larga escala; café

ABSTRACT

Terpenes constitute a widely diverse class of chemical compounds in plants, performing functions related to the primary and specialized metabolism of these organisms, being, many times, restricted to specific lineages or species. Terpene biosynthesis is catalyzed by terpene synthases (TPS), a family of enzymes whose properties vary among themselves. Such variation is related to expansion events to which TPS coding genes were submitted throughout the evolution of land plants, which are often organized in clusters as a result of episodes of genetic duplication and neofunctionalization. Recent analyzes have detected homology relationships between clusters of diterpene synthases in *Nicotiana tabacum* (Solanaceae), *Solanum lycopersicum* (Solanaceae) and in the coffee species *Coffea canephora* (Rubiaceae) and *Coffea arabica* (Rubiaceae), with variations in expression pattern and product. *C. arabica*, of historical economic importance, lacks the vegetative structures in which the genes in question are mostly expressed in these Solanaceae species, which raises questions about the expression profile and function of these genes in this coffee species. Furthermore, given its economic importance, it is substantial to investigate the possible effects of predicted climate change on elements of the specialized metabolism of *C. arabica*. In this study, the objective was to evaluate the transcriptional profile of seven diterpene synthases arranged in tandem in the genome of *C. arabica*. In order to determine the transcriptional profile of these genes related to the synthesis of diterpenes in *Coffea arabica*, data from 48 RNA-seq libraries of *C. arabica* were reanalyzed, from four studies related to the response of the species to environmental stressors and activation of defenses, as well as as stages of fruit maturation. Differential expression analyzes revealed that five genes showed low transcriptional levels, and two showed differential expression between experiments. Variations were detected in relation to the organs in which these genes are transcribed, as well as the modulation of the transcriptional level in response to the conditions to which the plants were submitted. These results represent an initial step in delineating the transcriptional profile and functional understanding of these genes in the production of specialized metabolites in *C. arabica*, which may be of applied interest.

Keywords: plant genetics; diterpenes; high-throughput nucleotide sequencing; coffee

SUMÁRIO

1	INTRODUÇÃO.....	p. 6
2	OBJETIVOS.....	p. 15
3	MATERIAIS E MÉTODOS.....	p. 16
3.1	Dados genômicos – genoma de referência.....	p. 16
3.2	Análise do perfil transcricional em <i>Coffea arabica</i> por RNA-seq.....	p. 16
3.3	Material vegetal para análise por RT-PCR semi-quantitativo.....	p. 21
3.4	Extração de RNA total.....	p. 22
3.5	Desenho de iniciadores para TPS de interesse.....	p. 22
4	RESULTADOS E DISCUSSÃO.....	p. 23
4.1	Perfil transcricional das TPS de interesse em <i>Coffea arabica</i>	p. 23
4.2	RT-PCR semi-quantitativa: ensaios iniciais.....	p. 33
5	CONCLUSÃO.....	p. 38
6	REFERÊNCIAS.....	p. 39

1 INTRODUÇÃO

Terpenos – ou terpenoides – constituem uma classe de compostos químicos amplamente diversa em plantas, a qual conta com mais de 80.000 já identificados e muitos a serem ainda descritos (EVERT; EICHHORN, 2018; CHRISTIANSON, 2017). Parte desses compostos são comuns à maioria das plantas e desempenham funções relacionadas ao metabolismo primário desses organismos, como é o caso das giberelinas, citocininas, auxinas, esteróis de membrana e outros. Todavia, a maioria dos terpenos mostra-se restrita a determinadas linhagens ou espécies de plantas nas quais participam do metabolismo especializado, desempenhando funções ecológicas que contribuem para o aumento do *fitness* desses organismos, tais como a atração de polinizadores, de dispersores de sementes e de microrganismos benéficos, bem como a defesa contra patógenos e herbívoros (PICHERSKY; RAGUSO, 2018; TETALI, 2019). Essa miríade de compostos tem sido empregada pela humanidade visando diferentes finalidades, que perpassam o uso de terpenos enquanto produtos farmacêuticos, nutricionais, flavorizantes, defensivos químicos de plantas, biocombustíveis, entre outros (ZERBE *et al.*, 2015).

Embora amplamente diversos, todos os terpenos derivam das mesmas unidades básicas compostas por cinco carbonos: os isômeros isopentenil difosfato (IPP) e dimetilalil difosfato (DMAPP). Estes podem ser gerados no citosol, pela via do Ácido Mevalônico (MVA), ou em plastídeos, pela via do Fosfato de Metileritritol (MEP). A partir da atividade de *cis*- e *trans*-prenil transferases (CPTs/TPTs), as unidades são condensadas para formar intermediários de diferentes comprimentos, os *cis*- e *trans*-prenildifosfatos. Tais intermediários servem, então, como substrato para a ação de uma família de enzimas, denominadas terpeno sintases (TPS), cuja atividade catalítica remove o grupamento difosfato e conduz à síntese do esqueleto de carbono de isoprenos (com cinco carbonos), monoterpenos (com 10 carbonos), sesquiterpenos (com 15 carbonos), diterpenos (com 20 carbonos) ou sesterterpenos (com 25 carbonos) (CHEN *et al.*, 2011; THOLL, 2015; PICHERSKY; RAGUSO, 2018; ZHOU; PICHERSKY, 2020).

Durante a biossíntese de terpenos, diferentes fatores contribuem para a geração da característica diversidade estrutural desses compostos. Dentre eles, cabe salientar: (i) a capacidade das TPS em produzir diferentes produtos a partir de um único substrato como consequência de alterações estocásticas de carga no sítio ativo da enzima; (ii) a modificação de aminoácidos nas TPS resultando em drástica alteração na síntese de terpenos com a produção de novos compostos, a qual é intensificada pela ampla plataforma de 30 a 100 genes de TPS em cada genoma de forma a favorecer a seleção e evolução de terpenos; e (iii) a possibilidade de modificação dos esqueletos de terpenos gerados por reações enzimáticas secundárias,

perpassando hidroxilação, peroxidação, metilação, clivagem, ou outras, de maneira a ampliar o número de compostos produzidos (THOLL, 2015; PICHERSKY; RAGUSO, 2018).

Uma vez que as TPS desempenham papel primordial na síntese e diversificação de terpenos em plantas, estas enzimas constituem objeto de interesse central da engenharia metabólica *in planta* e heteróloga para a produção de terpenóides de interesse econômico (THOLL, 2015; PICHERSKY; RAGUSO, 2018; TETALI, 2019). Compreendendo um elevado e crescente número de enzimas presentes em quase todos os táxons do reino vegetal, análises da família de TPS revelaram que esta divide-se nas subfamílias TPS-a, -b, -c, -d, -e/f, -g e -h, de acordo com relações filogenéticas e determinado grau de conservação funcional (CHEN *et al.*, 2011; HESKES *et al.*, 2018; THOLL, 2015). Ademais, TPS podem ser classificadas de acordo com sua estrutura e mecanismo de atividade em dois grupos: TPS de classe I, que detêm um domínio DDxxD rico em aspartato e altamente conservado na extremidade C-terminal, relacionado à remoção do grupamento difosfato para criar um intermediário carbocátion, o qual pode passar por diferentes reações e originar compostos distintos; e TPS de classe II, que detêm um domínio DxDD na extremidade N-terminal relacionado à ionização de seu substrato via protonação, geralmente formando diterpenos. (CHEN *et al.*, 2011; THOLL, 2015; ZHOU; PICHERSKY, 2020). Existem, ainda, casos de TPS bifuncionais, nas quais se observa a ocorrência de ambos os domínios ativos das classes I e II. Essas foram descritas apenas em plantas sem sementes e Gimnospermas, e são consideradas primitivas em relação à TPS monofuncionais de classe I e II (CHEN *et al.*, 2011; ZERBE; BOHLMANN, 2014; 2015).

Os genes que codificam as TPS passaram por significativos eventos de expansão ao longo da evolução das plantas terrestres. Tais genes são frequentemente encontrados em *clusteres* – ou agrupamentos – nos cromossomos, cuja origem sugere eventos de duplicação por *crossing-over* desigual e neofuncionalização combinados à realocação no genoma. Frequentemente, estão também inclusos outros genes que codificam enzimas modificadoras que têm como substrato o produto das TPS, como as enzimas citocromo P450s. Caracteristicamente, genes organizados em tandem apresentam grande similaridade entre si. Tal configuração facilita a regulação dos múltiplos genes envolvidos em uma mesma via ao nível da cromatina, além de auxiliar na prevenção do acúmulo de possíveis produtos citotóxicos (CHEN *et al.*, 2011; THOLL, 2015).

O rápido desenvolvimento da biologia de sistemas vegetais – que combina a metabolômica, genômica, transcriptômica, bioquímica e bioinformática – tem se mostrado um poderoso aliado para o setor industrial, uma vez que possibilita a identificação de genes de

interesse em espécies de plantas não-modelo. Essa, combinada à engenharia metabólica, propicia a produção sustentável de compostos vegetais de interesse em grande escala (ZERBE; BOHLMANN, 2015). Neste cenário, tem-se atribuído grande atenção a diterpenos. Utilizados pela humanidade desde a Era Neolítica visando as mais diversas finalidades, hoje, tais compostos contam com ampla gama de aplicações industriais e movimentam vários bilhões de dólares anualmente (ZERBE; BOHLMANN, 2015). Dentre os diterpenos de maior relevância para o setor industrial, cabe destacar o *cis*-abienol. Este composto pode ser empregado industrialmente como precursor na produção do Ambrox®, fixador utilizado em itens de perfumaria de elevado valor (ZERBE; BOHLMANN, 2014; 2015).

O *cis*-abienol foi identificado em espécies de plantas pertencentes a famílias de Gimnospermas e Angiospermas (ZERBE; BOHLMANN, 2014). Diferenças observadas no processo de biossíntese do Z-abienol entre esses grupos sugerem surgimento evolutivo independentemente do diterpeno, representando um caso de evolução convergente do metabolismo especializado em plantas (SALLAUD *et al.*, 2012; ZERBE; BOHLMANN, 2014). Na espécie *Abies balsamea* (Pinaceae), na qual foi inicialmente identificado, a biossíntese do composto realiza-se pela atividade enzimática da diterpeno sintase bifuncional *cis*-abienol sintase (CAS), tendo como resultado o acúmulo do *cis*-abienol na periderme da espécie (ZERBE; BOHLMANN, 2014; 2015). Já em *Nicotiana tabacum* (Solanaceae), a biossíntese do diterpeno processa-se em duas etapas sequenciais catalisadas por enzimas distintas, respectivamente codificadas pelos genes NtCPS2 e NtABS expressos em tricomas glandulares, onde se acumula o Z-abienol nessa espécie. Em *N. tabacum*, o diterpeno detém particular interesse como precursor de compostos que influenciam a qualidade e o odor característico do fumo (SALLAUD *et al.*, 2012).

Análises filogenéticas preliminares indicam que os genes de *N. tabacum* NtCPS2 e NtABS são prováveis ortólogos dos genes codificantes das TPS18 e TPS41 no tomateiro - *Solanum lycopersicum* (Solanaceae) (MATSUBA *et al.*, 2015). Ambas são diterpeno sintases, sendo a TPS18 (locus Solyc08g005720) do tipo TPS-e/f, e a TPS41 (locus Solyc08g005710) do tipo TPS-c. Há, ainda, diferenças em relação aos órgãos e tecidos em que estes genes são expressos, bem como ao produto da proteína codificada pelos genes entre as espécies. Tais diferenças compreendem a expressão do gene codificante da TPS18 em raízes e botões florais de tomateiro, com biossíntese de um diterpeno desconhecido, e expressão do gene codificante da TPS41 em tricomas, com papel fisiológico indefinido em *S. lycopersicum* (MATSUBA *et al.*, 2013; ZHOU; PICHERSKY, 2020). Ademais, nessa espécie, os genes em questão

encontram-se organizados em um *cluster* localizado no cromossomo 8. Nesse *cluster*, estão também inclusos os genes correspondentes às TPS19, TPS20 e CPT1, altamente expressos em tricomas e envolvidos na biossíntese e acúmulo de monoterpenos nesses apêndices epidérmicos, bem como os genes codificantes da TPS21 (Solyc8g005640, diterpeno sintase do tipo TPS-e/f), CPT2 e da citocromo P450 oxidorreductase (CYP71BN1), altamente expressos em peciólulos e relacionados à biossíntese de um diterpeno desconhecido até aquele momento, então batizado lycosantalol (MATSUBA *et al.*, 2015; SCHILMILLER *et al.*, 2009).

Por meio de análises de Bioinformática precedentes, detectou-se homologia entre genes descritos em *S. lycopersicum* e genes encontrados nas espécies de cafeeiro *Coffea canephora* (Rubiaceae) e *Coffea arabica* (Rubiaceae), também organizados em *cluster* (dados não publicados). Esta organização genômica similar para a disposição dessas TPS em espécies de Solanaceae e Rubiaceae suscitou interesse na obtenção de maior grau de detalhamento do perfil genômico e transcricional dos genes codificantes para estas TPS. Em análises preliminares, foram determinadas, com melhor exatidão, as relações de ortologia (por meio de BLAST recíproco na plataforma Galaxy, AFGAN *et al.*, 2018) e paralogia entre as TPS do cromossomo 8. Na Figura 1, estão indicadas as relações observadas entre ortólogos e parálogos das diterpeno sintases Solyc08g005640, Solyc08g005710 e Solyc08g005720 de tomateiro e *C. canephora*. Na tabela 1 e Figura 2, essas mesmas relações encontram-se detalhadas no genoma de *C. arabica*.

Curiosamente, *C. arabica* possui folhas opostas, glabras e carece de foliólulos. Dessa forma, a expressão das TPS na espécie não poderá corresponder ao padrão transcricional observado nas TPS21 e TPS41 de *S. lycopersicum*, nem a NtABS e NtCPS2 de *N. tabacum*. Isso aventa indagações acerca do perfil de expressão e função dos genes em questão nessa espécie, os quais podem estar relacionados à biossíntese de diterpenos distintos daqueles descritos em tomateiro e tabaco, embora homólogos. Ademais, cabe salientar que *C. arabica* é uma espécie alotetraploide, cujo genoma resulta da fusão dos genomas das espécies parentais *C. canephora* e *C. eugenioides*, e na qual considera-se que cada gene tem sua própria regulação de expressão homeóloga, com regulação global mais complexa que a de espécies diploides (CLARINDO; CARVALHO, 2008; VIEIRA *et al.*, 2019).

No cenário atual, avanços recentes nas tecnologias de sequenciamento têm beneficiado grandemente a compreensão de como os genomas são expressos, em especial, com o desenvolvimento da técnica de *RNA-sequencing* (RNA-seq) (WIMBERLEY *et al.*, 2020). Esta vale-se de métodos de sequenciamento de alto rendimento combinados a recursos de

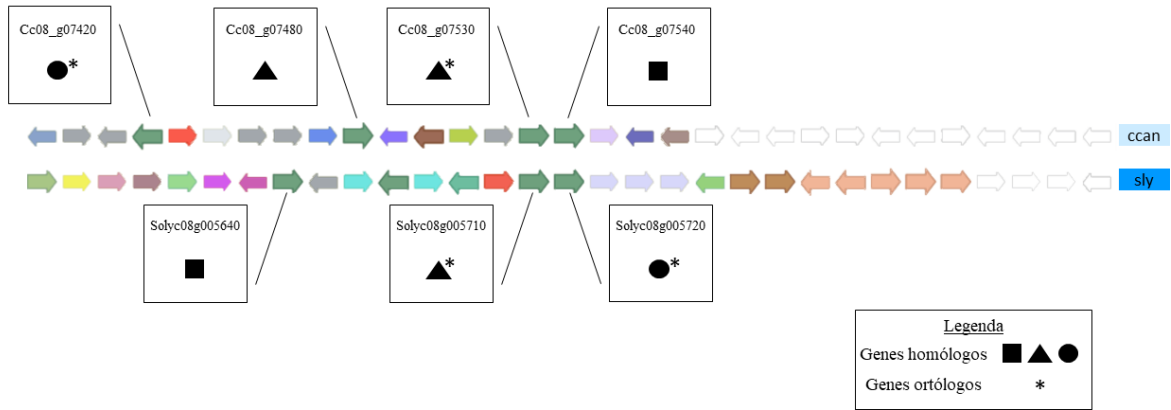
Bioinformática para identificar genes presentes em um genoma específico, genes transcricionalmente ativos em um determinado momento ou sob determinada condição, e identificar, com precisão, o nível de expressão gênica. Dentre as inovações proporcionadas pela técnica em questão, cabe destacar a possibilidade de realizar análises em espécies não-modelo, que carecem – ou não – de um genoma de referência para comparação. Posteriormente, os resultados podem ser validados por meio de uma técnica independente, como a RT-qPCR (LOWE et al., 2017). Tal estratégia tem se mostrado promissora no estudo de diversos *taxa* de plantas, como *Vitis vinifera* (Vitaceae) (MATARESE et al., 2014), *Ocimum* spp. (Lamiaceae) (ANAND et al., 2019), *Grindelia robusta* (Asteraceae) (ZERBE et al., 2015), bem como no estudo de diferentes famílias gênicas em *C. arabica* (IVAMOTO et al., 2017ab).

A produção de diterpenos em *C. arabica* tem recebido especial atenção nos últimos anos. Esses encontram-se entre os principais compostos químicos sintetizados pela espécie, em especial nos frutos, e mostraram deter generosa gama de aplicações de cunho nutricional, industrial, farmacêutico e cosmético (ESQUIVEL; JIMÉNEZ, 2012; MOEENFARD; ALVES, 2020). Isso tem incrementado a histórica relevância econômica da espécie, tradicionalmente apreciada enquanto bebida, acarretando implicações favoráveis nos aspectos físico e social de países tropicais e subtropicais, principais produtores e exportadores desse *commodity* no contexto mundial, nos quais inclui-se o Brasil (LASHERMES et al., 2008; VIEIRA, 2008).

Dada a importância econômica da espécie, é substancial investigar possíveis efeitos das tendências climáticas previstas sobre elementos do metabolismo especializado em *C. arabica*, as quais compreendem, principalmente, estresse hídrico, térmico e elevação da concentração de CO_2 ($[CO_2]$) atmosférico (SANTOS et al., 2022). Conhecer o perfil genético e metabólico no que tange a síntese de diterpenos em *C. arabica* torna-se primordial para a obtenção de informações visando ampliar as possibilidades de aplicação desses compostos bioativos em prol do bem-estar humano (LASHERMES et al., 2008; TETALI, 2019).

Figura 1. Organização de genes ortólogos e parálogos de diterpeno sintases de *S. lycopersicum* e *C. canephora*

Homólogos *Coffea canephora* – *Solanum lycopersicum*



Fonte: elaborado pelo autor

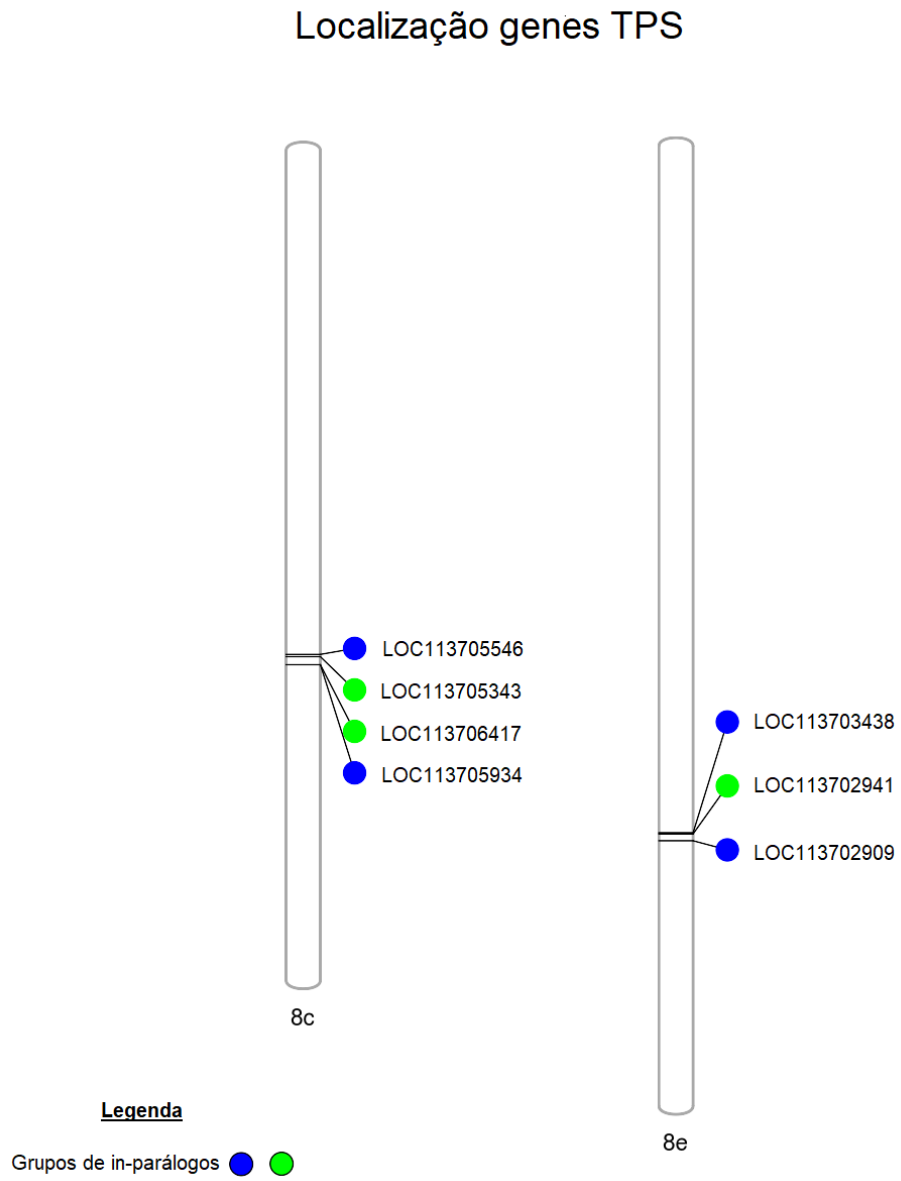
Tabela 1. Correspondência entre genes codificantes de TPS de *C. arabica* (informações do NCBI) e *S. lycopersicum* (informações da Sol Genomics Network)
 C = Cromossomo; ID = Identificação do gene; IT = Identificação do transcrito; IP = Identificação da proteína; I = Isoforma; L = Localização; TPS = Terpeno sintase

<i>Coffea arabica</i>						<i>Solanum lycopersicum</i>	
C	ID	IT	IP	I	L	TPS	ID
8c	LOC113705343	XM_027227161.1	XP_027082962.1	X1	23814473:23824970	TPS41	Solyc08g005710
		XM_027227162.1	XP_027082963.1	X1	23814731:23824970		
		XM_027227163.1	XP_027082964.1	X2	23814473:23824970		
		XM_027227164.1	XP_027082965.1	X3	23817857:23824970		
LOC113705934	XM_027227835.1	XP_027083636.1	-	24162064:24178611	TPS21	Solyc08g005640	
LOC113705546	XM_027227438.1	XP_027083239.1	X1	23711335:23719503	TPS18	Solyc08g005720	
	XM_027227439.1	XP_027083240.1	X2	23711463:23719432			
LOC113706417	XM_027228320.1	XP_027084121.1	X1	24145754:24157315	TPS41	Solyc08g005710	
	XM_027228321.1	XP_027084122.1	X2	24148103:24157315			
LOC113702909	XM_027224085.1	XP_027079886.1	X1	32672602:32680763	TPS18	Solyc08g005720	
	XM_027224086.1	XP_027079887.1	X2	32672608:32680763			
8e	LOC113703438	XM_027224798.1	XP_027080599.1	X1	32295395:32303338	TPS21	Solyc08g005640
		XM_027224799.1	XP_027080600.1	X2	32295395:32303338		
		XM_027224800.1	XP_027080601.1	X3	32295395:32304534		
		XM_027224801.1	XP_027080602.1	X4	32295395:32303338		
		XM_027224802.1	XP_027080603.1	X5	32295395:32303337		
LOC113702941	XM_027224124.1	XP_027079925.1	X1	32308944:32318536	TPS41	Solyc08g005710	

XM_027224125.1	XP_027079926.1	X2	32309962:32318536
----------------	----------------	----	-------------------

XM_027224126.1	XP_027079927.1	X3	32310316:32318536
----------------	----------------	----	-------------------

XM_027224127.1	XP_027079928.1	X3	32309980:32318536
----------------	----------------	----	-------------------

Figura 2. Disposição dos genes codificantes de TPS nos cromossomos 8c e 8e *C. arabica*

Fonte: Elaborado pelo autor

2 OBJETIVOS

O presente estudo tem por objetivo analisar o perfil transcricional de sete genes relacionados à síntese de diterpenos na espécie *Coffea arabica*, os quais são homólogos a diterpeno sintases caracterizadas no genoma de *S. lycopersicum* e *N. tabacum*. Busca-se, com isso, investigar a diversidade transcricional e resposta a estressores abióticos desses genes na espécie *C. arabica*.

3 MATERIAL E MÉTODOS

3.1 Dados genômicos – genoma de referência

Para a realização das análises no genoma de *C. arabica*, foi utilizada como referência a anotação disponível no National Center for Biotechnology Information (NCBI), no *link* <https://www.ncbi.nlm.nih.gov/genome/77>. Quando necessário, informações genômicas de referentes a *C. canephora* foram obtidas do banco de dados PLAZA Bioinformatics (VAN BEL *et al.*, 2022), em <https://bioinformatics.psb.ugent.be/plaza/>. Para cada *locus* selecionado, identificaram-se os respectivos transcritos e isoformas proteicas, os quais foram manualmente avaliados para definição daqueles a serem utilizados como referência para cada TPS estudada. Para tanto, utilizou-se como critério a análise de expressão diferencial, detalhada a seguir.

3.2 Análise do perfil transcricional em *Coffea arabica* por RNA-seq

Utilizando dados de RNA-seq disponíveis na literatura, foram realizadas análises da expressão gênica das TPS de interesse frente a diferentes condições visando a identificação de uma possível expressão diferencial dos respectivos genes, considerando seus transcritos. Tais condições compreenderam os estressores ambientais “elevada concentração de CO_2 ”, “elevada temperatura”, e “tratamento com ácido hexanoico”, bem como diferentes estágios de maturação do fruto.

Os dados de RNA-seq utilizados nesta análise foram provenientes de quatro experimentos cujos dados estão disponíveis publicamente. Um dos experimentos analisou o perfil transcricional de folhas de *C. arabica* sob elevada temperatura, depositados por Oliveira *et al.* (2020) sob o código de acesso PRJNA609253; transcritos de folhas de *C. arabica* sob elevada [CO_2], depositados por Marques *et al.* (2021) no NCBI sob o código de acesso BioProject PRJNA630692; transcritos de raízes de *C. arabica* sob tratamento com ácido hexanoico, depositados por Budzinski *et al.* (2021) no repositório European Nucleotide Archive (ENA), sob o código de acesso BioProject PRJEB39901; e transcritos de frutos de *C. arabica* em três estágios de maturação, depositados por Cheng *et al.* (2018) no banco de dados EMBL sob o código de acesso BioProject PRJEB24137. Na Tabela 2, encontram-se sumarizadas as características de cada estudo, bem como as bibliotecas de RNA-seq empregadas nas análises aqui realizadas.

As bibliotecas de RNA-seq referentes aos estudos de elevada temperatura e fases de desenvolvimento fruto encontravam-se disponíveis em sua forma bruta. Uma vez obtidas, foram descompactadas e processadas com o emprego do programa fastp (CHEN *et al.*, 2018) conforme configuração padrão para a eliminação de adaptadores, bem como de bases indeterminadas e de baixa qualidade. Tal programa está disponível na plataforma Galaxy (<https://usegalaxy.org/>), *online* e de acesso gratuito (GIARDINE *et al.*, 2005). Em contrapartida, os dados referentes aos estudos elevada concentração de elevada [CO₂] e tratamento com ácido hexanóico encontravam-se processados e disponíveis em acervo ao qual o grupo de estudos tem acesso, de forma que tal procedimento não foi necessário.

A análise dos respectivos perfis transcricionais dos genes de TPS de interesse em *C. arábica* foi realizada com o emprego da plataforma OmicsBox (<https://www.biobam.com/omicsbox>). Inicialmente, foi utilizada a função *Create Count Table*, opção *Transcript-level quantification*, que se utiliza do pacote de *softwares* RSEM (LI; DEWEY, 2011), associado ao alinhador Bowtie2 (LANGMEAD; SALZBERG, 2012), para realizar a quantificação dos transcritos. Como *input*, foram utilizadas as leituras de sequenciamento de alto rendimento de cada experimento, as quais foram mapeadas contra os transcritos de referência preditos para o genoma de *C. arábica*, obtidos no NCBI (vide item 3.1). Nessa etapa, foi desconsiderada a opção “Estimate RSPD” uma vez que as bibliotecas *Illumina* analisadas foram construídas segundo protocolo padrão, e selecionada a opção “Non-Strand Specific”, ao passo que os dados utilizados não são 5’ ou 3’ específicos. Os demais parâmetros de análise foram mantidos conforme configuração-padrão. Uma vez concluídas as quantificações, foram retornadas tabelas de contagem que exibiram os níveis de transcrição de cada sequência fornecida como *input*.

Em seguida, valendo-se dessas tabelas de contagem, foi empregada a função *Run Differential Expression Analysis*, opção *Pairwise Differential Analysis*, para a detecção de transcrição diferencial dos genes de interesse. Esta baseia-se no pacote do *software* edgeR (ROBINSON *et al.*, 2010), que permite a comparação par-a-par de condições experimentais distintas. Nessa etapa, foram admitidos apenas o cálculo de transcrição diferencial para genes que apresentassem a contagem por milhão (CPM) de transcritos acima de 5 em, ao menos, uma das bibliotecas do experimento (CPM filter = 5; Samples reaching CPM filter = 1), de forma a garantir maior confiabilidade. Os demais parâmetros de análise foram mantidos consoante configurações-padrão. As bibliotecas para cada condição foram comparadas conforme delineado na Tabela 3. Como resultado, foram retornadas tabelas que exibiram aumento,

redução ou manutenção dos níveis de transcrição entre as comparações, de acordo com os parâmetros de filtragem utilizados. Essas foram manualmente analisadas a fim de delimitar os perfis transcricionais dos genes de interesse.

As tabelas de expressão diferencial geradas também foram empregadas na função *Heatmap*, que se utiliza do *software* InChlib (ŠKUTA *et al.*, 2014) para a visualização organizada e exploração interativa dos dados. A partir dessa função, foram geradas representações visuais dos valores de CPM referentes aos transcritos dos genes codificantes das TPS de interesse, frente às condições analisadas.

Em posse desses dados, foram selecionados aqueles genes cuja expressão mostrou-se órgão-específica ou alterada em virtude do experimento para futura validação transcricional por RT-PCR semi-quantitativa.

Tabela 2. Bibliotecas de RNA-seq analisadas
BP = Bioproject; O = Órgão; CV = Variedade; TS = Tipo de sequenciamento; C = Controle; T = Tratamento

BioProject	Condição	Órgão	Cultivar	Tipo de sequenciamento	Controle	Tratamento
PRJNA630692	Elevada [CO ₂]	Folhas	Icatu vermelho	<i>Illumina HighSeq 2000 Single-end (1x50pb)</i>	CA_380ppm_25 C_1A_R1	CA_700ppm_25C _3A_R1
					CA_380ppm_25 C_1B_R1	CA_700ppm_25C _3B_R1
					CA_380ppm_25 C_1C_R1	CA_700ppm_25C _3C_R1
					CA_380ppm_37 C_9A_R1	CA_700ppm_37C _10A_R1
					CA_380ppm_37 C_9B_R1	CA_700ppm_37C _10B_R1
					CA_380ppm_37 C_9C_R1	CA_700ppm_37C _10C_R1
					CA_380ppm_42 C_2A_R1	CA_700ppm_42C _4A_R1
					CA_380ppm_42 C_2B_R1	CA_700ppm_42C _4B_R1

				CA_380ppm_42	CA_700ppm_42C	
				C_2C_R1	_4C_R1	
PRJNA609253	Elevada temperatura	Folhas	Acauã	<i>Illumina HiSeq 2000 Paired-end (2x150pb)</i>	SRR11196525_controle_replicat_a1_R1	SRR11196533_cador_replicata1_R1
					SRR11196525_controle_replicat_a1_R2	SRR11196533_cador_replicata1_R2
					SRR11196526_controle_replicat_a2_R1	SRR11196534_cador_replicata2_R1
					SRR11196526_controle_replicat_a2_R2	SRR11196534_cador_replicata2_R2
					SRR11196530_controle_replicat_a3_R1	SRR11196535_cador_replicata3_R1
					SRR11196530_controle_replicat_a3_R2	SRR11196535_cador_replicata3_R2
					SRR11196538_controle_replicat_a4_R1	SRR11196536_cador_replicata4_R1
					SRR11196538_controle_replicat_a4_R2	SRR11196536_cador_replicata4_R2
					SRR11196539_controle_replicat_a5_R1	SRR11196537_cador_replicata5_R1
					SRR11196539_controle_replicat_a5_R2	SRR11196537_cador_replicata5_R2
PRJEB39901	Ácido hexanóico	Raízes	Catuaí Vermelho	<i>Illumina NovaSeq 6000 Paired-end (2x150pb)</i>	LCS7609_DS_C_C_root1_R1	LCS7609_DS_C_HX_root1_R1
					LCS7609_DS_C_C_root1_R2	LCS7609_DS_C_HX_root1_R2
					LCS7609_DS_C_C_root2_R1	LCS7609_DS_C_HX_root2_R1
					LCS7609_DS_C_C_root2_R2	LCS7609_DS_C_HX_root2_R2
					LCS7609_DS_C_C_root3_R1	LCS7609_DS_C_HX_root3_R1
					LCS7609_DS_C_C_root3_R2	LCS7609_DS_C_HX_root3_R2
					LCS7609_DS_OC_root1_R1	LCS7609_DS_O_HX_root1_R1
					LCS7609_DS_OC_root1_R2	LCS7609_DS_O_HX_root1_R2

				6000	LCS7609_DS_	LCS7609_DS_O
				<i>Paired-end</i>	OC_root2_R1	HX_root2_R1
				(2x150pb)	LCS7609_DS_	LCS7609_DS_O
					OC_root2_R2	HX_root2_R2
					LCS7609_DS_	LCS7609_DS_O
					OC_root3_R1	HX_root3_R1
					LCS7609_DS_	LCS7609_DS_O
					OC_root3_R2	HX_root3_R2
						ERR2231838_am arelo_replicata1_ R1
						ERR2231838_am arelo_replicata1_ R2
					ERR2231835_v erde_replicata1_ R1	ERR2231839_am arelo_replicata2_ R1
					ERR2231835_v erde_replicata1_ R2	ERR2231839_am arelo_replicata2_ R2
				<i>Illumina</i>		ERR2231840_am arelo_replicata3_ R1
				<i>HiSeq</i>		ERR2231840_am arelo_replicata3_ R2
PRJEB24137	Maturação do fruto	Fruto	K7	4000		
				<i>Paired-end</i>		ERR2231841_ver melho_replicata1_ R1
				(2x150pb)		ERR2231841_ver melho_replicata1_ R2
					ERR2231836_v erde_replicata2_ R1	ERR2231842_ver melho_replicata2_ R1
					ERR2231836_v erde_replicata2_ R2	ERR2231842_ver melho_replicata2_ R2
						ERR2231843_ver melho_replicata3_ R1
						ERR2231843_ver melho_replicata3_ R2

Tabela 3. Resumo das análises comparativas de expressão diferencial
C = Condição; CT = Controle; T = Tratamento

C	CT	T
---	----	---

	25 °C 380 ppm CO_2	25 °C 700 ppm CO_2
Elevada [CO_2]	37 °C 380 ppm CO_2	37 °C 700 ppm CO_2
	42 °C 380 ppm CO_2	42 °C 700 ppm CO_2
Temperatura	Ótima	Elevada
	Obatã s/ ác. hexanóico	Obatã c/ ác. hexanóico 0,55mM
Ácido hexanóico	Catuaí s/ ác. hexanóico	Catuaí c/ ác. hexanóico 0,55mM
	Obatã s/ ác. hexanóico	Catuaí s/ ác. hexanóico
	Obatã c/ ác. hexanóico 0,55mM	Catuaí c/ ác. hexanóico 0,55mM
Maturação do fruto	Fruto verde	Fruto amarelo
	Fruto verde	Fruto vermelho

3.3 Material vegetal para análise por RT-PCR semi-quantitativo

O material biológico disponível em laboratório a ser utilizado nas análises aqui propostas consistiam em segmentos de folhas e raízes de *C. arabica* cv. Catuaí Vermelho IAC 144.

As folhas e raízes utilizadas provêm do estudo de Catarino et al. (2021). Neste, mudas de *Coffea arabica* cv. Catuaí Vermelho IAC144 (3–4 meses de idade) obtidas em viveiro certificado e com padrão de crescimento semelhante foram transplantadas para vasos de 12 litros contendo o substrato Floreira (Natus Solos do Brasil), com uma planta por vaso. Após 30 dias de aclimação em casa de vegetação, as plantas foram transferidas para câmaras open-top (OTC) instaladas em casa de vegetação no Centro de Fisiologia e Bioquímica do Instituto de Botânica de São Paulo (atual Instituto de Pesquisas Ambientais) para aclimação do tratamento de CO_2 (ambiente atmosférico – a[CO_2] = 400 ppm ou [CO_2] atmosférico elevado – e[CO_2] = 800 ppm), conforme descrito por Sanches et al. (2017). A injeção de gás foi aplicada através de um cilindro de CO_2 comprimido acoplado ao sistema. Durante este período de aclimação a [CO_2], as plantas foram regadas diariamente com água da torneira e solução nutritiva de Hoagland e Arnon (HOAGLAND; ARNON, 1950) uma vez por semana. Após 30 dias do período de aclimação, metade das plantas cultivadas em a[CO_2] e e[CO_2] foram submetidas a déficit hídrico (WD) com suspensão da irrigação, enquanto a outra metade foi mantida em regime de rega diária (WW) por 40 dias consecutivos, apresentando os seguintes tratamentos: a[CO_2]+WW, a[CO_2]+WD, e[CO_2]+WW e e[CO_2]+WD. A temperatura e umidade relativa do ar (HMP45C-L, Campbell Sci.) e a radiação fotossintética ativa (Li-190R, LiCor) foram monitoradas dentro das OTCs durante o período do experimento. Material foliar e segmentos

de raízes foram obtidos no 20º e 40º dia após a imposição do déficit hídrico, que correspondem aos períodos de “seca moderada” e “seca severa”.

Cabe salientar que todas as amostras disponíveis contavam com, ao menos, três réplicas biológicas (três indivíduos distintos) e técnicas (três repetições de coleta).

3.4 Extração de RNA total

A partir do material vegetal disponível de *C. arabica* cv. Catuaí Vermelho IAC 144 coletadas, o RNA total foi obtido por meio do emprego do *kit* RNeasy Plant (Qiagen, Hilden, North Rhine-Westphalia, Alemanha), conforme instruções do fabricante. Em seguida, determinou-se a concentração do RNA extraído utilizando-se o *kit Qubit RNA BR Assay* (Thermo Fisher Scientific, Wilmington, DE, EUA) e conseguinte quantificação no fluorímetro Qubit (Thermo Fisher Scientific, Wilmington, DE, EUA), conforme instruções do fabricante. A integridade do RNA obtido foi verificada por meio da realização da técnica de eletroforese em gel de agarose a 1,5%. Em seguida, metade do volume obtido foi diluído em água *RNAse-free Water* para padronizar as diferentes concentrações para 35 ng/μl, de maneira a atender demandas específicas da reação de amplificação na etapa seguinte.

3.5 Desenho de iniciadores para TPS de interesse

Primers específicos para os genes selecionados foram desenhados por meio do emprego do programa *Primer 3 Plus* (<https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) para amplificar produtos que variam de 90 a 150 pb. Como *input*, utilizou-se a sequência do respectivo cDNA previsto para a sequência dos transcritos disponíveis no NCBI e, para casos em que o gene apresenta polimorfismo, considerou-se a sequência referente à isoforma X1. Cabe salientar que, dentre os possíveis *primers* gerados pelo programa, foram selecionados aqueles localizados na região 3' do gene visando garantir que os produtos estejam íntegros, e não degradados pela ação putativa de microRNAs.

4 RESULTADOS E DISCUSSÃO

No presente estudo, foram utilizadas 48 bibliotecas de RNA-seq de *C. arabica* oriundos de quatro estudos. Três destes são relacionados à resposta de *C. arabica* a estressores ambientais e ativação de defesas – alta temperatura, altas concentrações de CO_2 , e respostas a um elicitor – e um estudo referente a comparação de perfis transcricionais ao longo da maturação do fruto – PRJEB24137. Essas foram submetidas a análises *in silico* a fim de apurar o perfil transcricional de sete genes que codificam TPS de interesse na espécie em questão, bem como possíveis modulações frente às condições impostas.

4.1 Perfil transcricional das TPS de interesse em *Coffea arabica*

As análises de expressão diferencial ajudaram a revelar perfis transcricionais distintos entre os sete genes analisados no presente estudo (Tabela 4). Ocorreram variações em relação aos órgãos em que esses genes são transcritos, bem como a modulação do nível transcricional em resposta às condições às quais as plantas foram submetidas. Dos sete genes analisados, quatro foram substancialmente transcritos – LOC113705934, LOC113705546, LOC113702909 e LOC113702941 – enquanto os demais foram detectados a níveis mais baixos, e por isso não atingiram os mínimos para análise de expressão diferencial – LOC113705343, LOC113706417 e LOC113703438. Na Figura 3, são ilustrados os níveis de transcrição referente a cada biblioteca analisada em CPM. Em função disso, voltou-se a atenção àqueles que compõem o primeiro grupo.

Os genes LOC113705934 e LOC113702941 mostraram-se órgão-específicos, pois foram detectados transcritos apenas em folhas e raízes de *C. arabica*, respectivamente. Conjuntos de TPS órgão-específicas também foram descritas em diversas espécies vegetais, como *Matricaria recutita* (Asteraceae), *Salvia hispanica* (Lamiaceae) e *Dendrobium officinale* (Orchidaceae). Nesses estudos, foi proposto que tal distribuição desigual manifesta uma regulação estrita da produção de terpenóides com funções variadas, de maneira a conferir características únicas ao órgão em que se encontram, o que pode ocorrer, também, em *C. arabica* (IRMISCH *et al.*, 2012; WIMBERLEY *et al.*, 2020; YU *et al.*, 2020).

Em relação à resposta às condições ambientais avaliadas, a ausência de regulação transcricional observada no gene LOC113705934 na variedade Acauã submetida à condição de elevada temperatura, bem como na transcrição do gene LOC11370294 em raízes das variedades Catuaí e Obatã submetidas à aplicação de ácido hexanóico, indica que estes não participam dos

respectivos mecanismos moleculares e fisiológicos que conferem termotolerância ou que estabelecem um estado de resposta a elicitores na planta (OLIVEIRA *et al.*, 2020; BUDZINSKI *et al.*, 2020), produzindo terpenos em folhas e raízes de maneira constitutiva sob tais condições.

O gene LOC113705546 foi majoritariamente transcrito em folhas e raízes de *C. arabica*. Em folhas, foi constatado o aumento da transcrição do gene em questão em resposta à elevada concentração de $[CO_2]$ na variedade Icatu Vermelho em faixas de temperatura ótima e moderadamente elevada. Tal condição ambiental foi demonstrada favorável à *C. arabica*, uma vez que induz ao aumento da taxa de fotossíntese, bem como à mitigação dos efeitos prejudiciais de outros estresses abióticos, como o calor e o déficit hídrico (SANCHES *et al.*, 2017; MARTINS *et al.*, 2016; RAMALHO *et al.*, 2018; AVILA *et al.*, 2020; CATARINO *et al.*, 2021). Em contrapartida, em raízes, foi constatada redução da transcrição do gene LOC113705546 em resposta à aplicação de ácido hexanóico na variedade Catuaí, agente natural que estabelece na planta um estado de resposta mais rápida e/ou robusta frente a estressores ambientais (BUDZINSKI *et al.*, 2020). Dessa forma, o gene LOC113705546 e as características bioquímicas da TPS por ele codificada em folhas da variedade Icatu Vermelho e raízes de Catuaí constituem elementos interessantes a serem explorados e melhor elucidados para um possível aproveitamento no cultivo da espécie frente às mudanças climáticas previstas. Possivelmente, este gene atua na produção de compostos em respostas a estressores foliares, mas não a estressores em raiz.

Dentre os quatro genes expressos, o LOC113702909 demonstrou o perfil transcricional mais abrangente, sendo encontrado em raízes, folhas e frutos, bem como em todas as variedades analisadas. Nessa perspectiva, é possível que o gene LOC113702909 esteja relacionado à biossíntese de terpenos que atuam em processos metabólicos constitutivos em *C. arabica*, o que justificaria o perfil transcricional ubíquo. Ademais, embora predominantemente estável nos diferentes órgãos e condições, houve redução da expressão em resposta à aplicação de ácido hexanóico em raízes na variedade Obatã, cujo efeito foi previamente explanado. Levando em conta o exposto aqui, o gene LOC113702909 e as características bioquímicas da TPS por ele codificada, em especial em raízes da variedade Obatã, também constituem elementos interessantes a serem melhor elucidados para um possível aproveitamento no cultivo da espécie.

Foram observadas ainda diferenças de níveis transcricionais entre variedades em um mesmo órgão. O gene LOC113705934 foi expresso em folhas da variedade Acauã, porém foi detectado em níveis baixos na variedade Icatu Vermelho; o inverso ocorreu para o gene LOC113705546. Tal variação no padrão de transcrição pode decorrer de características

intrínsecas aos respectivos genótipos das variedades de *C. arabica*, conforme constatado por Marracini et al. (2011) e Oliveira et al. (2020). Variações intraespecíficas na expressão de genes relacionados à biossíntese de terpenos são frequentes em plantas, e foram descritas para outras espécies de interesse, como *Zea mays* e *Cannabis sativa*, de forma que diferentes variedades podem ser mais ou menos vantajosas em contextos específicos (KÖLLNER *et al.*, 2008; BOOTH *et al.*, 2020).

Nota-se, também, a ocorrência de certa variação a nível de transcritos oriundos de um mesmo gene. No que tange ao gene LOC113705546, o transcrito XM_027227439.1 encontrou-se limitado a raízes, enquanto XM_027227438.1 fez-se presente não apenas neste órgão, mas também em folhas. No caso do gene LOC113702909, foi observada certa estabilidade entre os transcritos, ambos expressos em todos os órgãos, ainda que com variações – ausência de XM_027224085.1 em folhas submetidas à condição “elevada [CO_2]” em temperatura ótima e em frutos vermelhos, e ausência de XM_027224086.1 em folhas submetidas à condição “elevada temperatura”. Casos similares foram relatados para a síntese de terpenos nas espécies *Populus trichocarpa*, *Camellia sinensis* e *Isodon rubescens*, nas quais foram identificadas diferenças significativas nas isoformas traduzidas e no produto final, com variações nos padrões de expressão espaço-temporal entre transcritos de um mesmo gene (YUAN *et al.*, 2009; LIU *et al.*, 2018; JIN *et al.*, 2019). Conforme revisado por Lam et al. (2022), a geração de transcritos diversos resultantes de *splicing* alternativo mostra-se relacionada à regulação fina do desenvolvimento ou de resposta a condições de estresse em plantas. Novos estudos podem ser úteis para uma melhor compreensão da regulação dos genes em questão em *C. arabica*.

Os genes analisados podem ser classificados em dois grupos de in-parálogos: LOC113705546, LOC113705934, LOC113703438 e LOC113702909; e LOC113705343, LOC113706417 e LOC113702941. Genes in-parálogos resultam de eventos de duplicação de um mesmo gene em um genoma, que percorreram trajetórias evolutivas distintas frente a mutações e pressão seletiva, em direção à pseudogenização, neofuncionalização ou subfuncionalização (GRIFFITHS *et al.*, 2022a). Tais eventos foram documentados em genes que participam da biossíntese de terpenos em diferentes grupos de plantas terrestres, como a Gimnosperma *Picea abies* e a Angiosperma *Oryza sativa*. Nestas, foi evidenciada a diversificação funcional e plasticidade evolutiva das TPSs do metabolismo secundário das plantas a partir de neofuncionalização baseada em algumas substituições de nucleotídeos não sinônimos (KEELING et al., 2008; WANG *et al.*, 2012). Em *C. arabica*, diferenças marcantes nos padrões de transcrição observadas em elementos de ambos os grupos podem estar

relacionadas à ocorrência desse fenômeno, cuja compreensão demanda maiores investigações moleculares e bioquímicas.

A nível genômico, os genes transcritos dividem-se igualmente entre os subgenomas de *C. arabica*: LOC113705934 e LOC113705546 estão localizados nos cromossomos 8c, oriundo do subgenoma *canephora*; e LOC113702909 e LOC113702941 estão localizados no cromossomo 8e, oriundo do subgenoma *eugenioides* (Figura 1). Ademais, esses mostraram diferenças marcantes no padrão de transcrição, em especial os genes LOC113705546 e LOC113702909. Em oposição a outras espécies aloploplóides, como *Gossypium hirsutum* (ADAMS *et al.*, 2004; FLAGEL *et al.*, 2012), estudos recentes têm demonstrado a ausência de propensão de um subgenoma sobre o outro em *C. arabica*, tanto para padrões do desenvolvimento quanto para estressores abióticos, o que reitera essa co-expressão observada neste estudo. Conforme proposto, tal característica sugere uma regulação mais sofisticada da expressão gênica nessa espécie, cuja elucidação requer maiores investigações (COMBES *et al.*, 2013; LASHERMES *et al.*, 2016; VIEIRA *et al.*, 2019).

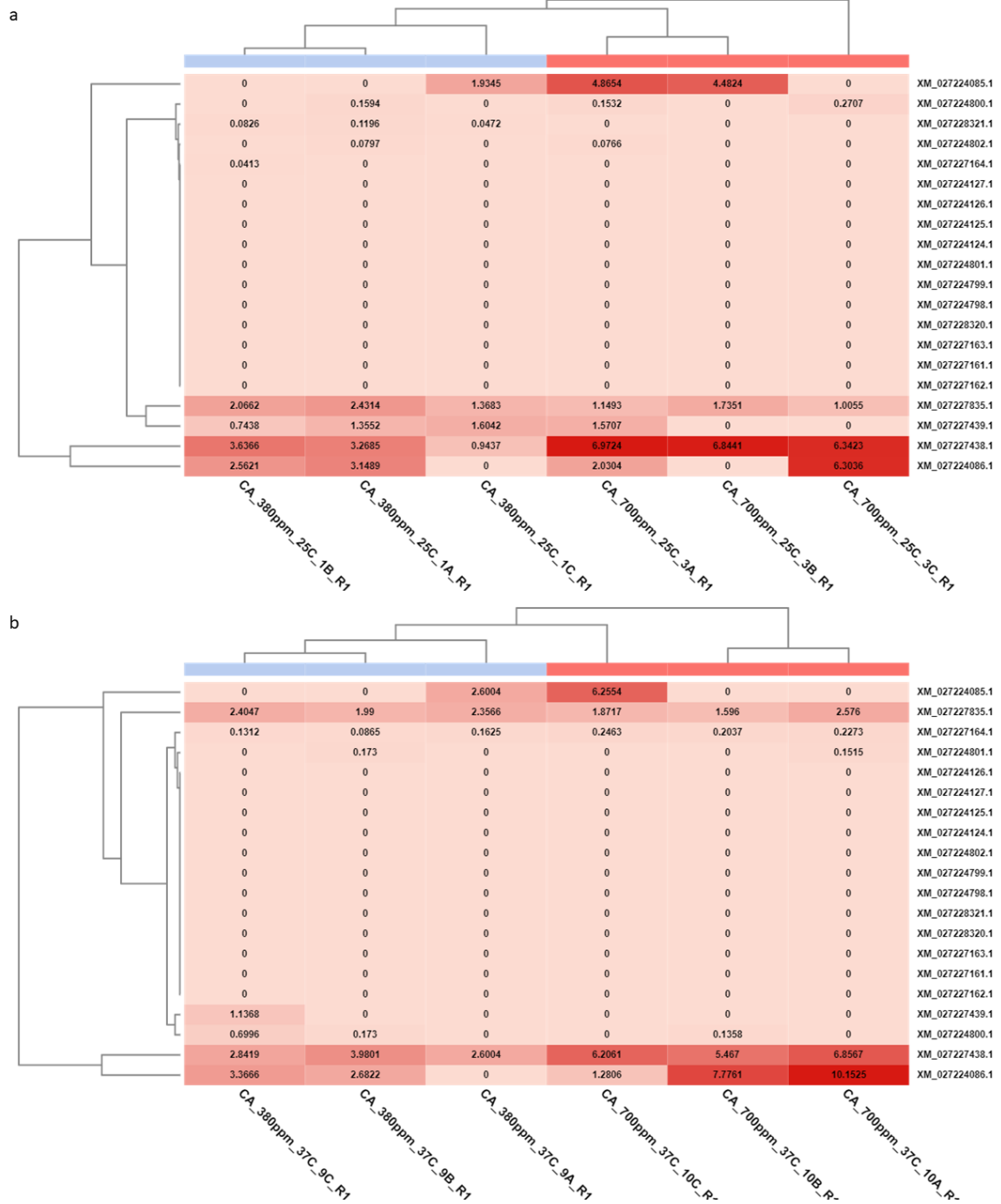
Tabela 4. Perfil transcricional dos genes de interesse

ID = Identificação do gene; IT = identificação do transcrito; T = Temperatura elevada; HX = Ácido hexanóico; CA = Catuaí; O = Obatã; C = Controle; FA = Fruto amarelo; FC = Fruto cereja; * = Transcrição a níveis abaixo de 5CPM; - = Não transcrito

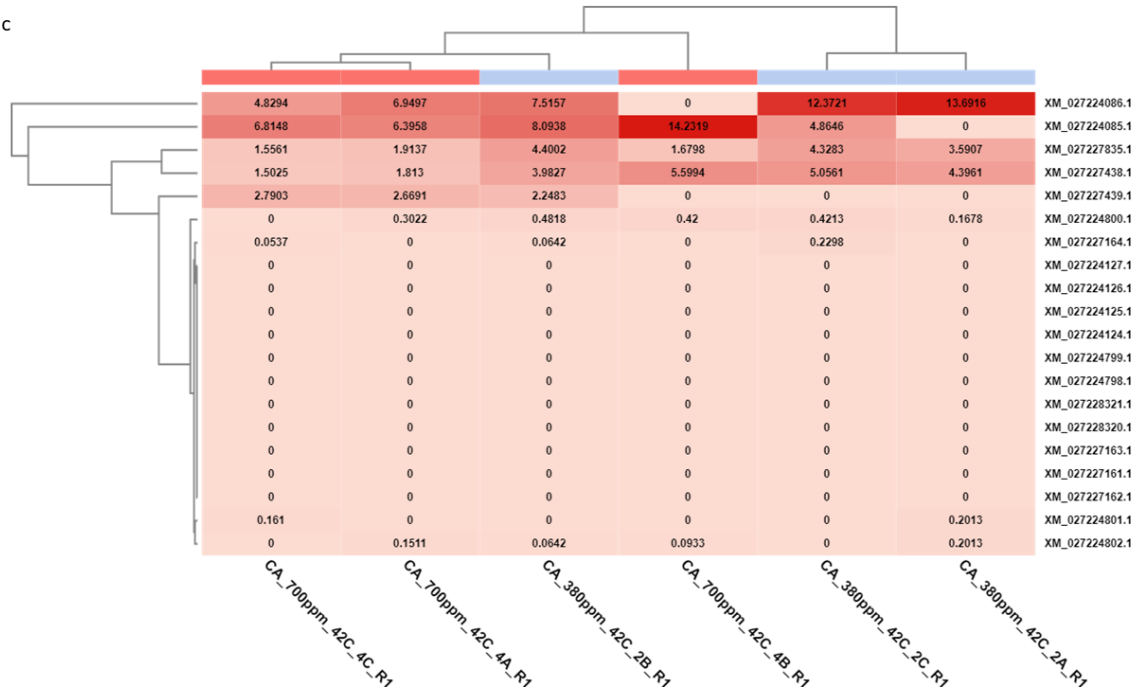
ID	IT	T	HX CA	HX O	C CA x O	HX CA x O	25 °C CO ₂	37 °C CO ₂	42 °C CO ₂	FA	FC
LOC113705343	XM_027227161.1	-	-	-	-	-	-	-	-	-	-
	XM_027227162.1	-	-	-	-	-	-	-	-	-	-
	XM_027227163.1	*	*	*	*	-	-	-	-	*	-
	XM_027227164.1	*	*	*	*	-	-	*	*	*	*
LOC113705934	XM_027227835.1	Estável logFC= 0.05 P- Value= 0.91	-	*	*	-	*	*	*	-	*
LOC113705546	XM_027227438.1	*	Redução logFC= - 10.56 P-Value= 9.17x10 ⁻¹³	Estável logFC= -2.54 P-Value= 0.04	Estável logFC= -0.54 P- Value= 0.04	Estável logFC= 6.53 P- Value= 0.03	Aumento logFC= 1.64 P-Value= 5.06x10 ⁻⁴	Aumento logFC= 1.23 P-Value= 0.001	Estável logFC= 0.11 P- Value= 0.85	*	-
	XM_027227439.1	*	Estável logFC= 0.88 P-Value= 0.70	Estável logFC= -1.55 P-Value= 0.48	Estável logFC= 0.13 P- Value= 0.96	Estável logFC= -2.68 P- Value= 0.15	*	*	*	*	*
LOC113706417	XM_027228320.1	-	*	*	-	-	-	-	-	-	-
	XM_027228321.1	*	*	*	*	-	*	-	-	-	-
LOC113702909	XM_027224085.1	Estável logFC= 7.66 P- Value= 0.001	Estável logFC= -3.09 P-Value= 0.09	Redução logFC= - 12.02 P-Value= 1.06x10 ⁻¹⁶	Estável logFC= -0.14 P- Value= 0.38	Estável logFC= -6.81 P- Value= 0.04	*	Estável logFC= 1.55 P-Value= 0.71	Estável logFC= 1.83 P- Value= 0.23	Estável logFC=1.3 P- Value=0.51	*
	XM_027224086.1	*	Estável	Estável	Estável	Estável	Estável	Estável	Estável	Estável	Estável

		logFC= 3.79 P-Value= 0.19	logFC= 1.45 P-Value= 0.36	logFC= -1.73 P- Value= 0.51	logFC= -0.70 P- Value= 0.67	logFC= 1.01 P-Value= 0.65	logFC= 1.91 P-Value= 0.21	logFC= -0.62 P- Value= 0.67	logFC=0.81 P- Value=0.76	logFC= 1.32 P- Value= 0.44
LOC113703438	XM_027224798.1	-	-	-	-	-	-	-	-	-
	XM_027224799.1	-	-	-	-	-	-	-	-	-
	XM_027224800.1	*	-	-	-	*	*	*	*	*
	XM_027224801.1	*	-	-	-	-	*	*	-	-
	XM_027224802.1	*	-	*	*	-	*	*	*	-
LOC113702941	XM_027224124.1	-	Estável logFC= -2.29 P-Value= 0.13	*	Estável logFC= 1.80 P- Value= 0.01	*	-	-	-	-
	XM_027224125.1	-	-	-	-	-	-	-	-	-
	XM_027224126.1	-	-	-	-	-	-	-	-	-
	XM_027224127.1	-	-	-	-	-	-	-	-	-

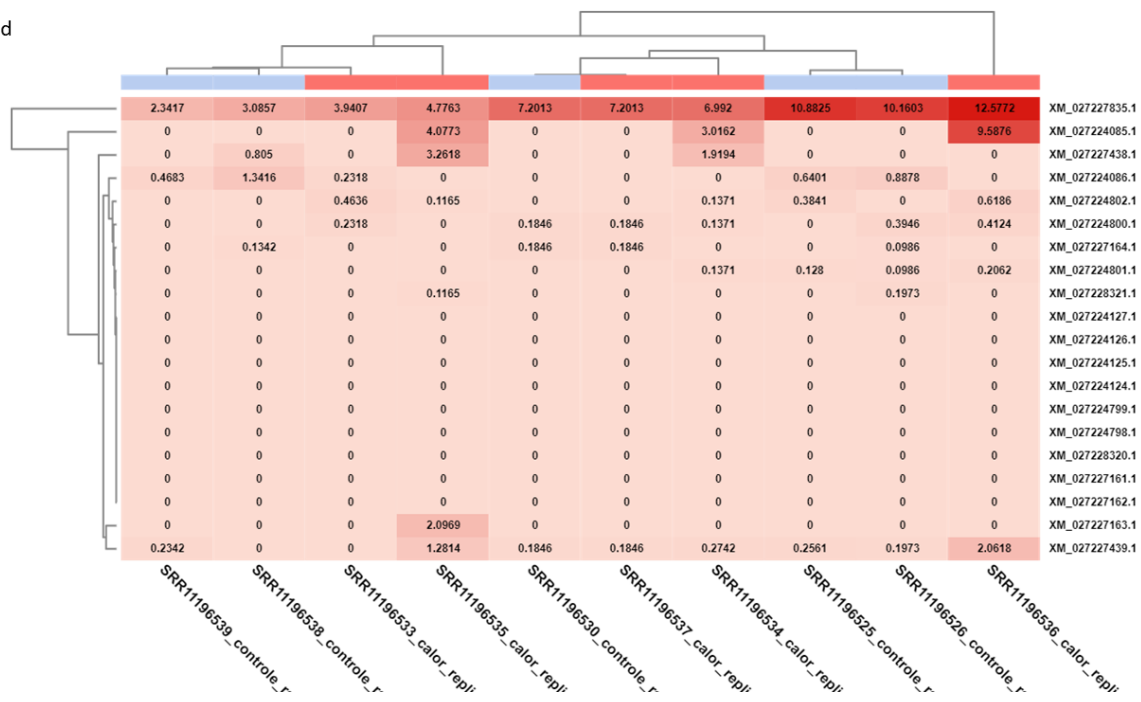
Figura 3. Níveis transcricionais de genes, representados em CPM, por biblioteca. a) 25°C CO₂ 380 ppm vs CO₂ 700 ppm; b) 37°C CO₂ 380 ppm vs CO₂ 380 ppm; c) 42°C CO₂ 380 ppm vs CO₂ 380 ppm; d) elevada temperatura; e) Obatã s/ ác. hexanóico vs Obatã c/ ác. hexanóico 0,55mM; f) Catuaí s/ ác. hexanóico vs Catuaí c/ ác. hexanóico 0,55mM; g) Obatã s/ ác. hexanóico vs Catuaí s/ ác. hexanóico; h) Obatã c/ ác. hexanóico 0,55mM vs Catuaí c/ ác. hexanóico 0,55mM; i) fruto verde vs fruto amarelo; j) fruto verde vs fruto cereja

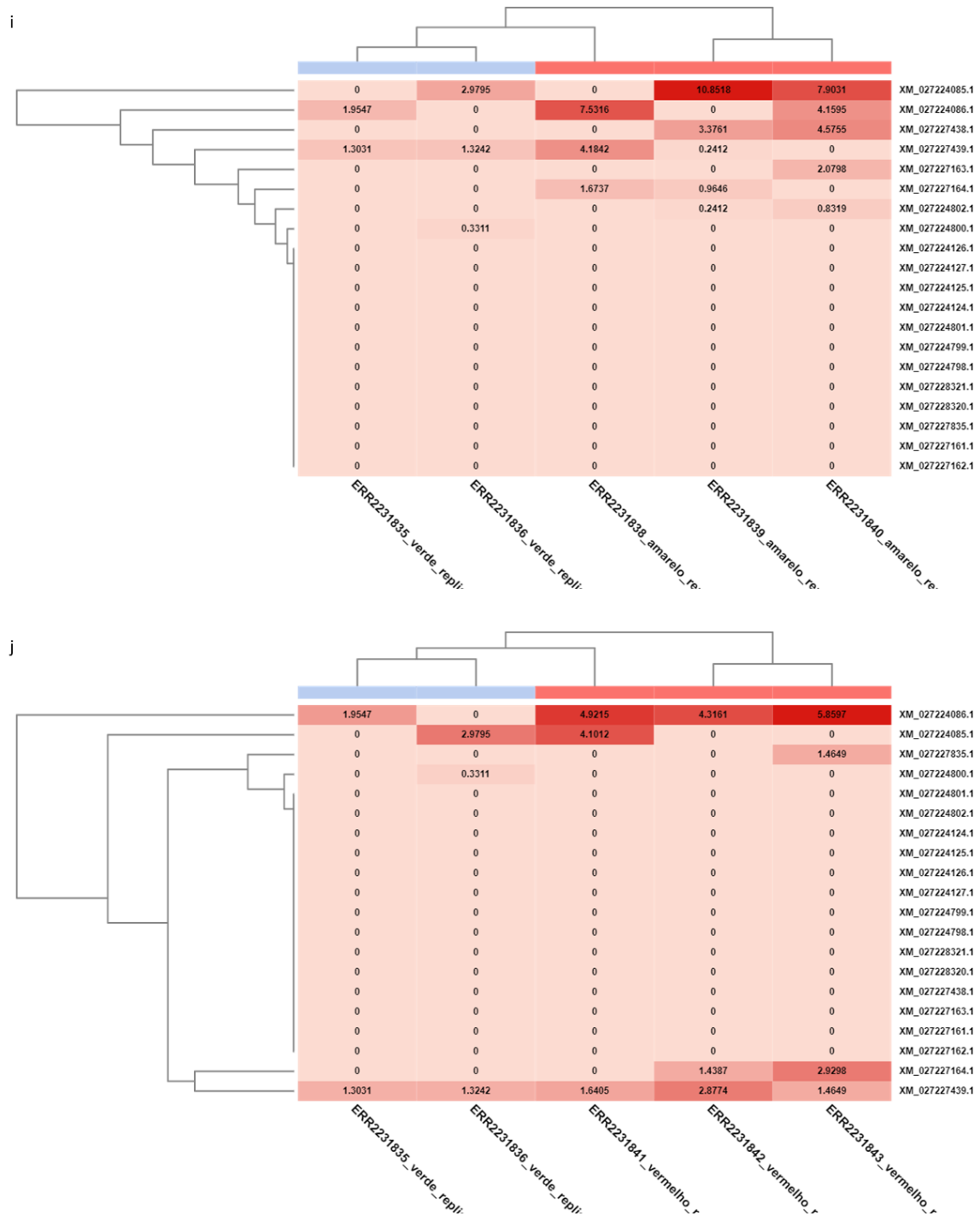


c



d





4.3 – RT-PCR semi-quantitativa: ensaios iniciais

Uma vez delimitados os perfis de expressão dos genes analisados, foram selecionados para validação dos resultados por RT-PCR semi-quantitativa genes que apresentaram perfil de expressão órgão-específicos ou que mostraram modulações no perfil de expressão frente às diferentes condições analisadas. Cabe salientar que, devido a existência de dados anteriores já gerados pelo laboratório, foi priorizada a realização de análises referentes a folhas de *C.*

arabica. Para atender a essa demanda, foram selecionados os genes LOC113705934, LOC113705546 e LOC113702909, cuja expressão foi detectada nesse órgão. *Primers* específicos foram desenhados por meio do emprego do programa *Primer 3 Plus*. As respectivas sequências desenhadas, incluindo-se o normalizador gliceraldeído-3-fosfato desidrogenase (GAPDH) (CRUZ *et al.*, 2009), encontram-se na Tabela 5.

Foi ainda realizada extração de RNA total de folhas e raízes. Para tanto, foi utilizado um protocolo adaptado em um primeiro momento, o que rendeu dificuldades na obtenção de RNA. Em função disso, foi despendido um período de tempo para a realização de alterações e testes de protocolo até o estabelecimento de um protocolo eficiente. Uma vez estabelecido o protocolo, realizou-se a extração de RNA total a partir de segmentos de folhas e, posteriormente, raízes de *C. arabica* em dois momentos do experimento de Catarino *et al.* (2021) – seca moderada e severa. Todas as amostras de RNA obtidas foram quantificadas por meio do fluorímetro Qubit para determinação das respectivas concentrações (Tabela 6). Em seguida, foram, também, submetidas à eletroforese em gel de agarose a 1,5% para verificação da integridade do material (Figura 4). A não visualização de determinadas amostras foi atribuída a baixos níveis de concentração do RNA obtido, a qual estaria atuando como um limitante da técnica. Todavia, foram consideradas íntegras, uma vez que este foi o panorama geral das demais.

Lamentavelmente, não houve tempo hábil para a realização da amplificação e análise semi-quantitativa dos genes de interesse para a validação dos padrões de expressão gênicas em função do período dispendido na otimização de protocolos de extração de RNA. Entretanto, como todas as partes de preparação já foram realizadas, espera-se que este experimento seja realizado em breve.

Tabela 5. Sequências de *primers* desenhados

Gene	Transcrito	Tipo de <i>primer</i>	Sequência
LOC113705934	XM_027227835	<i>Forward</i>	5' - AAAGAGGCCGAAGACAACACTG - 3'
		<i>Reverse</i>	5' - TCCTTGCACACTCTTGG AAC - 3'
LOC113705546	XM_027227438	<i>Forward</i>	5' - AGCGAGGAAGAGGCTATAACG - 3'
		<i>Reverse</i>	5' - AGGAATTGAGCTTCCCTTGG - 3'
LOC113702909	XM_027224085	<i>Forward</i>	5' - CAGAACATGGGCTAAACGAG - 3'

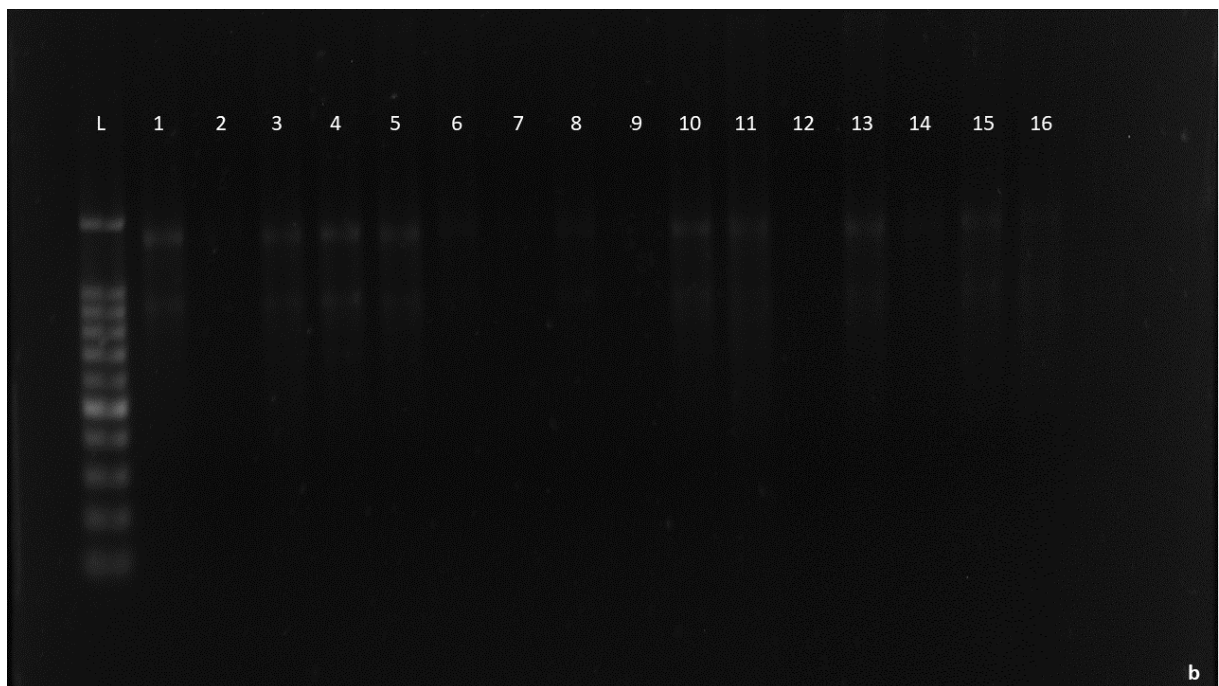
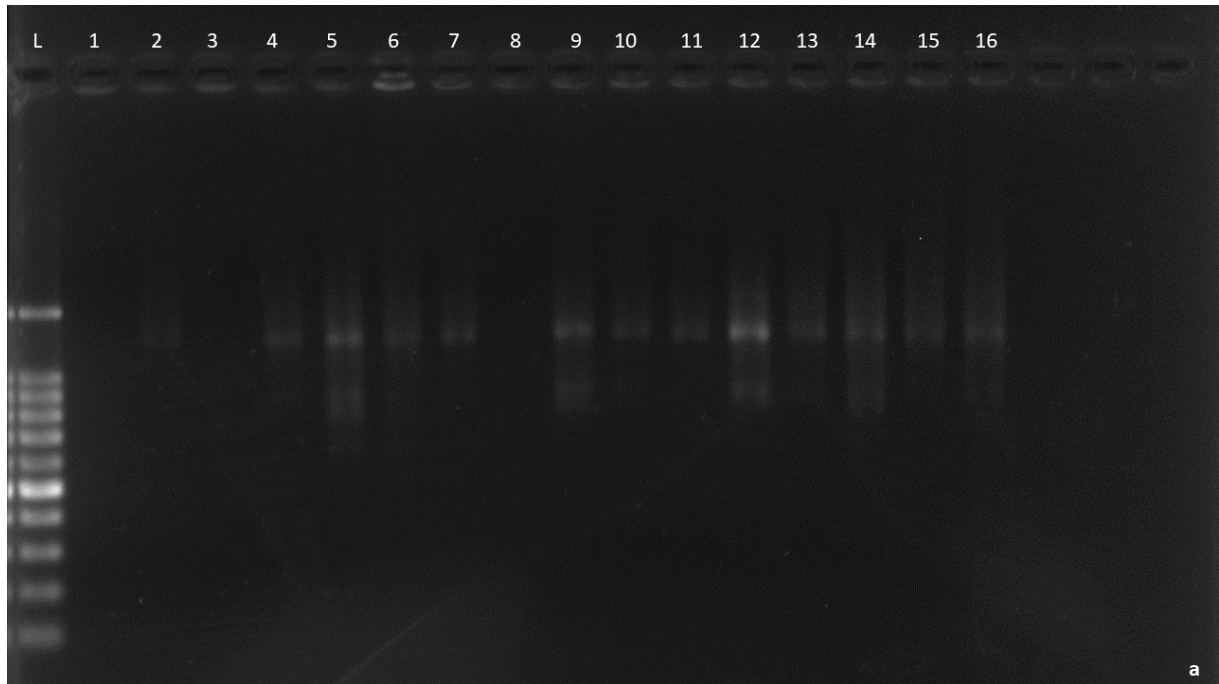
GAPDH	-	<i>Reverse</i>	5' - CGGAAGTTTTATGCAGTCCAG - 3'
		<i>Forward</i>	5' - TTGAAGGGCGGTGCAAA - 3'
		<i>Reverse</i>	5' - AACATGGGTGCATCCTTGCT - 3'

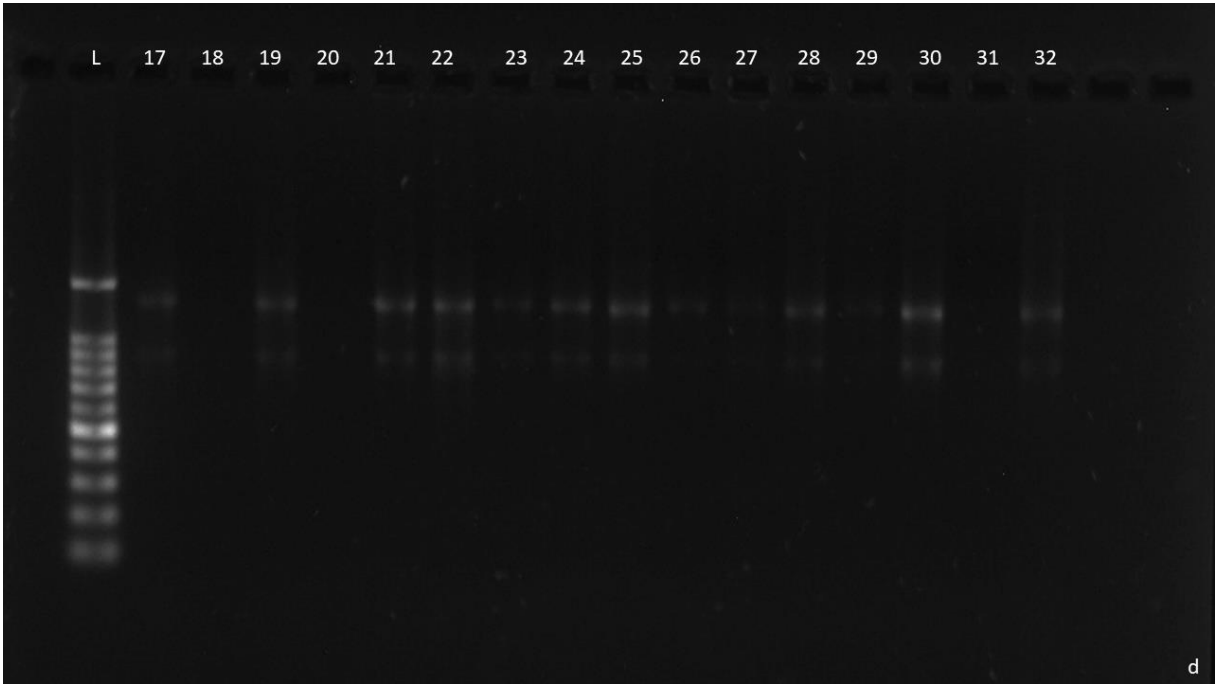
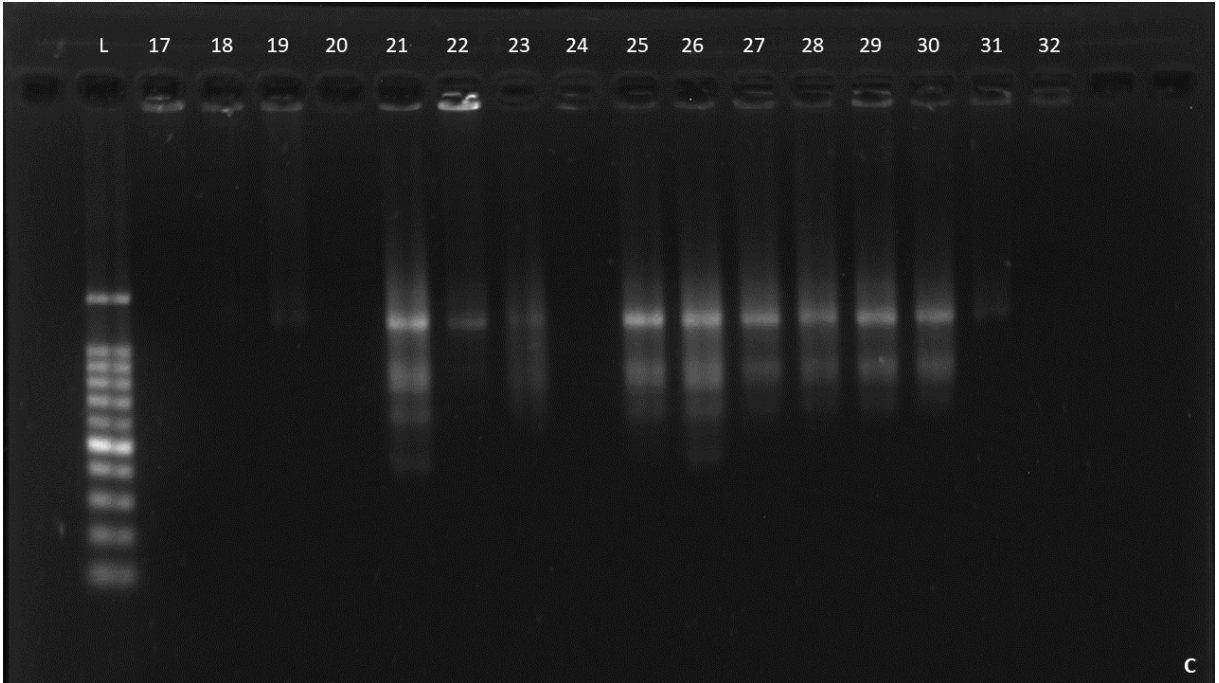
Tabela 6. Concentração das extrações de RNA.

WW = suficiência hídrica; WD = deficiência hídrica; eCO₂ = 700ppm de CO₂; aCO₂ = condição atmosférica normal de CO₂

Amostra	Condição	Tratamento	Folhas (ng/μl)	Raízes (ng/μl)
1	Seca moderada	WW eCO ₂	383	540
2		WW eCO ₂	133	69,6
3		WD eCO ₂	139	369
4		WD eCO ₂	111	150
5		WW eCO ₂	246	321
6		WW eCO ₂	159	470
7		WD eCO ₂	98,3	186
8		WD eCO ₂	166	223
9		WW aCO ₂	170	400
10		WW aCO ₂	151	203
11		WD aCO ₂	169	168
12		WD aCO ₂	189	422
13		WW aCO ₂	261	290
14		WW aCO ₂	240	520
15		WD aCO ₂	171	151
16		WD aCO ₂	259	418
17	Seca severa	WW eCO ₂	313	570
18		WW eCO ₂	273	111
19		WD eCO ₂	225	455
20		WD eCO ₂	78	560
21		WW eCO ₂	277	405
22		WW eCO ₂	90,6	176
23		WD eCO ₂	270	123
24		WD eCO ₂	40,1	168
25		WW aCO ₂	354	115
26		WW aCO ₂	433	427
27		WD aCO ₂	193	520
28		WD aCO ₂	296	135
29		WW aCO ₂	204	413
30		WW aCO ₂	246	140
31		WD aCO ₂	101	350
32		WD aCO ₂	89,5	341

Figura 4. Eletroforeses do RNA total extraído
L = leader; a) folhas sob seca modera; b) raízes sob seca moderada; c) folhas sob seca severa; d) raízes sob seca severa





5 CONCLUSÃO

O presente estudo representa um passo inicial no delineamento do perfil transcricional e entendimento funcional de sete genes organizados em *cluster* que codificam TPS em *C. arabica*, de forma a contribuir para a elucidação da função efetiva desses na produção de metabólitos especializados em café.

Os genes em questão mostraram ampla variação transcricional entre si, perpassando a limitação à detecção a níveis-traço, transcrição órgão-específica ou abrangente, bem como alterações frente a estressores abióticos específicos. Em conjunto, esses dados podem indicar funções distintas desempenhadas pelos terpenos sintetizados, os quais podem atuar como elementos constitutivos de órgãos ou da planta em sua totalidade, ou ainda proporcionar características adaptativas a *C. arabica*. Ademais, parte das variações observadas podem refletir características intrínsecas às variedades analisadas, uma vez também foram detectadas a nível intraespecífico. Os genes LOC113705546 e LOC113702909, em especial, podem deter interesse prático, pois apresentaram a maior faixa de variação transcricional.

Contudo, os dados obtidos até o momento requerem verificação funcional para que sejam efetivamente aplicáveis. Ademais, outros questionamentos emergiram dos resultados obtidos, os quais excedem o escopo deste trabalho. Assim sendo, novos estudos que se utilizem de abordagens moleculares e bioquímicas para melhor caracterizar os genes de interesse e as TPS codificadas devem ser realizados a fim de fornecer pistas para a elucidação desses questionamentos.

Referências bibliográficas

ADAMS, K. L.; PERCIFIELD, R.; WENDEL, J. F. Organ-specific silencing of duplicated genes in a newly synthesized cotton allotetraploid. **Genetics**, v. 168, n. 1, p. 2217–2226, 2004.

AFGAN, E. et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. **Nucleic Acids Research**, v. 46, n. W1, p. W537–W544, 2018.

ANAND, A.; JAUARAMAIAH, R. H.; BEEDKAR, S. D.; et al. Terpene profiling, transcriptome analysis and characterization of cis- β -terpineol synthase from *Ocimum*. **Physiology and Molecular Biology of Plants**, v. 25, n. 1, p. 47–57, 2019.

AVILA, R. T.; CARDOSO, A. A.; DE ALMEIDA, W. L.; et al. Coffee plants respond to drought and elevated [CO₂] through changes in stomatal function, plant hydraulic conductance, and aquaporin expression. **Environmental and Experimental Botany**, v. 177, 2020.

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

BUDZINSKI, I. et al. Transcriptome Analyses of Leaves Reveal That Hexanoic Acid Priming Differentially Regulate Gene Expression in Contrasting *Coffea arabica* Cultivars. **Frontiers In Sustainable Food Systems**, v. 5, p. 2571, 2021.

CATARINO, I. C. A.; MONTEIRO, G. B.; FERREIRA, M. J. P. et al. Elevated [CO₂] Mitigates Drought Effects and Increases Leaf 5-O-Caffeoylquinic Acid and Caffeine Concentrations During the Early Growth of *Coffea Arabica* Plants. **Frontiers in Sustainable Food Systems**, v. 5, p. 1–12, 2021.

CHEN, S.; ZHOU, Y.; CHEN, Y.; GU, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. **Bioinformatics**, 2018.

CHEN, F. et al. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. **The Plant Journal**, v. 66, n.1, p. 212-229, 2011.

- CHENG, B.; AGNELO, F.; HENRY, R. J. The coffee bean transcriptome explains the accumulation of the major bean components through ripening. **Scientific Reports**, v. 8, n. 11414, p. 1-11, 2018.
- CHRISTIANSON, D. W. Structural and chemical biology of terpenoid cyclases. **Chemical Reviews**, v. 117, n. 1, p. 11570-11648, 2017.
- CLARINDO, W. R.; CARVALHO, C. R. (2008). First *Coffea arabica* karyogram showing that this species is a true allotetraploid. **Plant Systematics and Evolution**, v. 274, p. 237-241, 2008.
- COMBES, M. C.; DEREPPER, A.; SEVERAC, D.; BERTRAND, B.; LASHERMES, P. Contribution of subgenomes to the transcriptome and their intertwined regulation in the allopolyploid *Coffea arabica* grown at contrasted temperatures. **New Phytologist**, v. 200, n. 1, p. 251–260, 2013.
- CRUZ, F. et al. Evaluation of coffee reference genes for relative expression studies by quantitative real-time RT-PCR. **Molecular Breeding**, v. 23, n. 4, p. 607–616, 2009.
- SANTOS, C. V.; OLIVEIRA, A. F. FERREIRA FILHO, J. B. D. S. Potential Impacts Of Climate Change On Agriculture And The Economy In Different Regions of Brazil. **Revista de Economia e Sociologia Rural**, v. 60, n. 1, p. 1–24, 2022.
- ESQUIVEL, P.; JIMÉNEZ, V. M. Functional properties of coffee and coffee by-products. **Food Research International**, v. 46, n. 2, p. 488–495, 2012.
- EVERT, R. F.; EICHHORN, S. E. Composição molecular das células vegetais. *In*: EVERT, R. F.; EICHHORN, S. E. **Biologia vegetal**. 8. ed. Rio de Janeiro: Guanabara Koogan, 2018. p. 29-43.
- FLAGEL, L. E.; WENDEL, J. F.; UDALL, J. A. Duplicate gene evolution, homoeologous recombination, and transcriptome characterization in allopolyploid cotton. **BMC Genomics**. v. 13, n. 302, p. 1-13, 2012.
- GIARDINE, B; RIEMER, C; HARDISON, R. C; BURHANS, R.; ELNITSKI, L.; SHAH, P.; ZHANG, Y. BLANKENBERG, D.; ALBERT, I.; TAYLOR, J.; MILLER, W.; KENT, W.J.; NEKRUTENKO, A.. Galaxy: a platform for interactive large-scale genome analysis. **Genome Research**, v. 15, n. 10, p. 1451-1455, 2005.
- GRIFFITHS, A. J. F.; DOEBLEY, J.; PEICHEL, C.; WASSARMAN, D. A. Evolução de genes, traços e espécies. *In*: GRIFFITHS, A. J. F.; DOEBLEY, J.; PEICHEL, C.; WASSARMAN, D. A., **Introdução à genética**, 12. ed. Rio de Janeiro: Guanabara Koogan, 2022. p. 664-694a.
- HESKES, A. M. et al. Biosynthesis of bioactive diterpenoids in the medicinal plant *Vitex agnus-castus*. **Plant Journal**, v. 93, n. 5, p. 943–958, 2018.
- HOAGLAND, D. R.; ARNON, D. I. The water-culture method for growing plants without soil. **California Agricultural Experiment Station**, v. 347, p 1–32, 1950.

- IVAMOTO, S. T.; REIS, O.; DOMINGUES, D. S. et al. Transcriptome analysis of leaves, flowers and fruits perisperm of *Coffea arabica* L. reveals the differential expression of genes involved in raffinose biosynthesis. **PLoS ONE**, v. 12, n. 1, p. 1–17, 2017a.
- IVAMOTO, S. T.; SAKURAY, L. M.; FERREIRA, L. P.; et al. Diterpenes biochemical profile and transcriptional analysis of cytochrome P450s genes in leaves, roots, flowers, and during *Coffea arabica* L. fruit development. **Plant Physiology and Biochemistry**, 2017b.
- IRMISCH, S.; KRAUSE, S. T.; KUNERT, G.; GERSHENZON, J.; DEGENHARDT, J.; KÖLLNER, T. G. The organ-specific expression of terpene synthase genes contributes to the terpene hydrocarbon composition of chamomile essential oils. **BMC Plant Biology**, v. 12, n. 84, p. 1-13, 2012.
- JIN, B.; GUO, J.; TANG, J.; et al. An alternative splicing alters the product outcome of a class I terpene synthase in *Isodon rubescens*. **Biochemical and Biophysical Research Communications**, v. 512, n. 2, p. 310–313, 2019.
- KEELING, C. I.; WEISSHAAR, S.; LIN, R. P. C.; BOHLMANN, J. Functional plasticity of paralogous diterpene synthases involved in conifer defense. **Proceedings of the National Academy of Sciences of the United States of America**, v. 105, n. 3, p. 1085–1089, 2008.
- KÖLLNER, T. G.; HELD, M.; LENK, C.; HILTPOLD, I; TURLINGS, T. C. J.; GERSHENZON, J.; DEGENHARDT, J. A maize (E)-b-caryophyllene synthase implicated in indirect defense responses against herbivores is not expressed in most American maize varieties. **Plant Cell**, v. 20, n. 1, p. 482–494, 2008.
- LAM, P. Y.; WANG, L.; LO, C.; ZHU, F. Y. Alternative Splicing and Its Roles in Plant Metabolism. **International Journal of Molecular Sciences**, v. 23, n. 13, p. 1–20, 2022.
- LANGMEAD, B.; SALZBERG, S. L. Fast gapped-read alignment with Bowtie 2. **Nature Methods**, v. 9, n. 4, p. 357–359, 2012.
- LASHERMES, P.; ANDRADE, A. C.; ETIENNE, H. Genomics of coffee, one of the world's largest traded commodities. In: MOORE, P. H.; MING, R. (Eds.), **Genomics of tropical crop plants**. New York: Springer, 2008. p. 203–225.
- LASHERMES, P.; HUEBER, Y.; COMBES, M. C., SEVERAC, D.; DEREPPER, A. Inter-genomic DNA exchanges and homeologous gene silencing shaped the nascent allopolyploid coffee genome (*Coffea arabica* L.). **G3: Genes, Genomes, Genetics**, v. 6, n. 9, p. 2937–2948, 2016.
- LI, B.; DEWEY, C. N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. **BMC Bioinformatics**, v. 12, n. 323, p. 1-16, 2011.
- LIU, G. F.; LIU, J. J.; HE, Z. R.; et al. Implementation of CsLIS/NES in linalool biosynthesis involves transcript splicing regulation in *Camellia sinensis*. **Plant Cell and Environment**, v. 41. n. 1, p. 176–186, 2018.

- MARRACINI, P.; FREIRE, L. P.; ALVES, G. S. C.; et al. RBCS1 expression in coffee: *Coffea* orthologs, *Coffea arabica* homeologs, and expression variability between genotypes and under drought stress. **BMC Plant Biology**, v. 11, n. 1, p. 1-23, 2011.
- MARQUES, I.; FERNANDES, I.; PAULO, O. S.; et al. A transcriptomic approach to understanding the combined impacts of supra-optimal temperatures and CO₂ revealed different responses in the polyploid *Coffea arabica* and its diploid progenitor *C. canephora*. **International Journal of Molecular Sciences**, v. 22, n. 6, p. 1–21, 2021.
- MARTINS, M. Q.; RODRIGUES, W. P.; FORTUNATO, A. S.; et al. Protective response mechanisms to heat stress in interaction with high [CO₂] conditions in coffee spp. **Frontiers in Plant Science**, v. 7, p. 1-18, 2016.
- MATARESE, F.; CUZZOLA, A.; SCALABRELLI, G.; D'ONOFRIO, C. Expression of terpene synthase genes associated with the formation of volatiles in different organs of *Vitis vinifera*. **Phytochemistry**, v. 105, n. 1, p. 12–24, 2014.
- MATSUBA, Y. et al. Evolution of a complex locus for terpene biosynthesis in Solanum. **Plant Cell**, v. 25, n. 6, p. 2022–2036, 2013.
- MATSUBA, Y. et al. Biosynthesis of the diterpenoid lycosantalonal via neryleryl diphosphate in solanum lycopersicum. **PLoS ONE**, v. 10, n. 3, p. 1–16, 2015.
- MOEENFARD, M.; ALVES, A. (2020). New trends in coffee diterpenes research from technological to health aspects. **Food Research International**, v. 134, p. 109207, 2020.
- MARTIANSYAH, I.; AMANAH, D. M.; PUTRANTO, R. A. Semi-quantitative RT-PCR analysis of transcripts encoding protease inhibitor in *Hevea brasiliensis* Muell. Arg latex. **IOP Conference Series: Earth and Environmental Science**, v. 183, n. 1, p. 1-10, 2018.
- OLIVEIRA, R et al. Elevated Temperatures Impose Transcriptional Constraints and Elicit Intraspecific Differences Between Coffee Genotypes. **Frontiers In Plant Science**, v. 11, n. 1, p. 1-16, 2020.
- PICHERSKY, E.; RAGUSO, R. A. Why do plants produce so many terpenoid compounds? **New Phytologist**, v. 220, n. 1, p. 692-702, 2018.
- RAMAKERS, C. et al. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. **Neuroscience Letters**, v. 339, n. 1, p. 62–66, 2003.
- RAMALHO, J. C.; PAIS, I. P.; LEITÃO, A. E.; et al. Can elevated air [CO₂] conditions mitigate the predicted warming impact on the quality of coffee bean? **Frontiers in Plant Science**, v. 9, p. 1–14, 2018.
- ROBINSON, M. D.; MCCARTHY, D. J.; SMYTH, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. **Bioinformatics**, v. 26, n. 1, p. 139–140, 2010.
- SALMONA, J. et al. Deciphering transcriptional networks that govern *Coffea arabica*

seed development using combined cDNA array and real-time RT-PCR approaches. **Plant Molecular Biology**, v. 66, n. 1, p. 105-124, 2008).

SALLAUD, C. et al. Characterization of two genes for the biosynthesis of the labdane diterpene Z-abienol in tobacco (*Nicotiana tabacum*) glandular trichomes. **Plant Journal**, v. 72, n. 1, p. 1–17, 2012.

SANCHES, R. F. E.; CATARINO, I. C. A.; BRAGA, M. R.; SILVA, E. A. da. Influência da alta concentração atmosférica de CO₂(↑[CO₂]atm) × disponibilidade hídrica nas relações hídricas, trocas gasosas e acúmulo de carboidratos em *Coffea arabica* L. **Hoehnea**, v. 44, n. 4, p. 635–643, 2017.

SCHILMILLER, A. L. et al. Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. **Proceedings of the National Academy of Sciences of the United States of America**, n. 106, v. 26, p. 10865–10870, 2009.

TAYLOR, S. et al. A practical approach to RT-qPCR-Publishing data that conform to the MIQE guidelines. **Methods**, v. 50, n. 4, p. S1-S5, 2010.

TETALI, S. D. Terpenes and isoprenoids: a wealth of compounds for global use. **Planta**, v. 249, n. 1, p. 1–8, 2019.

THOLL, D. Biosynthesis and Biological Functions of Terpenoids in Plants. **Advances in biochemical engineering/biotechnology**, v. 123, n. 1, p. 127-141, 2015.

VAN BEL, M. et al. PLAZA 5.0: extending the scope and power of comparative and functional genomics in plants. **Nucleic Acids Research**, v. 50, n. 1, p. D1468–D1474, 2022.

VIEIRA, H. D. (2008). Coffee: The plant and its cultivation. In: SOUZA, M. (Ed.), **Plant parasitic nematodes of coffee**. Dordrecht: Springer, 2008. p. 3-18.

VIEIRA, N. G.; FERRARI, I. F.; REZENDE, J. C.; MAYER, J. L. S.; MONDEGO, J. M. C. Homeologous regulation of Frigida-like genes provides insights on reproductive development and somatic embryogenesis in the allotetraploid *Coffea arabica*. **Scientific Reports**, v. 9, n. 1, p. 1–15, 2019.

WANG, Q.; HILLWIG, M. L.; WU, Y.; PETERS, R. J. CYP701A8: A rice ent-kaurene oxidase paralog diverted to more specialized diterpenoid metabolism. **Plant Physiology**, v. 158, n. 3, p. 1418–1425, 2012.

WIMBERLEY, J.; CAHILL, J.; ATAMIAN, H. S. De novo Sequencing and Analysis of *Salvia hispanica* Tissue-Specific Transcriptome and Identification of Genes Involved in Terpenoid Biosynthesis. **Plants**, v. 9, n. 405, p. 1–15, 2020.

XLSTAT Version 2014.5.03 Copyright Addinsoft 1995–2014, 2014 XLSTAT and Addinsoft are Registered Trademarks of Addinsoft. Disponível em: <<https://www.xlstat.com>>. Acesso em 14 abr. 2022.

YANG, M. et al. Divergent Evolution of the Diterpene Biosynthesis Pathway in Tea Plants (*Camellia sinensis*) Caused by Single Amino Acid Variation of ent-Kaurene Synthase. **Journal of Agricultural and Food Chemistry**, v. 68, n. 37, p. 9930–9939, 2020.

YUAN, Y.; CHUNG, J. D.; FU, X.; et al. Alternative splicing and gene duplication differentially shaped the regulation of isochorismate synthase in *Populus* and *Arabidopsis*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, n. 51, p. 22020–22025, 2009.

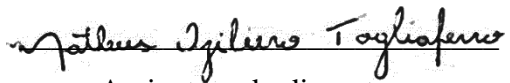
YU, Z.; ZHAO, C.; ZHANG, G.; TEIXEIRA DA SILVA, J. A.; DUAN, J. Genome-wide identification and expression profile of tps gene family in *dendrobium officinale* and the role of dotps10 in linalool biosynthesis. **International Journal of Molecular Sciences**, v. 21, n. 15, p. 1–22, 2020.

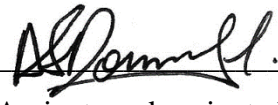
ZERBE, P.; BOHLMANN, J. Bioproducts, Biofuels, and Perfumes: Conifer Terpene Synthases and their Potential for Metabolic Engineering. *In*: Jetter, R. (eds) **Phytochemicals – Biosynthesis, Function and Application**. Recent Advances in Phytochemistry, v. 44. Springer, Cham, 2014. p. 58-108.

ZERBE, P., & BOHLMANN, J. Plant diterpene synthases: Exploring modularity and metabolic diversity for bioengineering. **Trends in Biotechnology**, v. 33, n. 7, p. 419–428, 2015.

ZERBE, P. et al. Exploring diterpene metabolism in non-model species: Transcriptome-enabled discovery and functional characterization of labda-7,13E-dienyl diphosphate synthase from *Grindelia robusta*. **Plant Journal**, v. 83, n. 5, p. 783–793, 2015.

ZHOU, F.; PICHERSKY, E. More is better: the diversity of terpene metabolism in plants. **Current Opinion in Plant Biology**, v. 55, n. 1, p. 1-10, 2020.


Assinatura do discente


Assinatura do orientador