UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
INSTITUTO DE BIOCIÊNCIAS – RIO CLARO

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS
(BIOLOGIA VEGETAL)

GENÔMICA COMPARATIVA DE LIPOXIGENASES EM PLANTAS E PERFIL TRANSCRICIONAL EM *Coffea arabica*

PAULA OLIVEIRA CAMARGO

Rio Claro – SP
2023

UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
INSTITUTO DE BIOCIÊNCIAS – RIO CLARO

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS (BIOLOGIA VEGETAL)

GENÔMICA COMPARATIVA DE LIPOXIGENASES EM PLANTAS E PERFIL TRANSCRICIONAL EM *Coffea arabica*

PAULA OLIVEIRA CAMARGO

Tese apresentada ao Instituto de Biociências do Câmpus de Rio Claro, Universidade Estadual Paulista, como parte dos requisitos para obtenção do título de Doutor em Ciências Biológicas (Biologia Vegetal)
Orientador: Dr. Douglas Silva Domingues

Rio Claro – SP
2023

C172g

Camargo, Paula Oliveira
Genômica comparativa de lipoxigenases em plantas e perfil transcricional em Coffea arabica / Paula Oliveira Camargo. -- Rio Claro, 2023
98 p.

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Instituto de Biociências, Rio Claro
Orientador: Douglas Silva Domingues

1. Famílias multigênicas. 2. Diversidade gênica em Angiospermas. 3. LOX em Coffea. 4. Elicitor ácido hexanoico. 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.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: **Genômica comparativa de lipoxigenases em plantas e perfil transcricional em Coffea arabica**

AUTORA: PAULA OLIVEIRA CAMARGO

ORIENTADOR: DOUGLAS SILVA DOMINGUES

Aprovada como parte das exigências para obtenção do Título de Doutora em Ciências Biológicas (Biologia Vegetal), área: Biologia Vegetal pela Comissão Examinadora:

Pesquisador Dr. DOUGLAS SILVA DOMINGUES (Participaçao Virtual)
Departamento de Botanica / IB UNESP Rio Claro

gov.br Documento assinado digitalmente
DOUGLAS SILVA DOMINGUES
Data: 05/06/2023 12:09:13-0300
Verifique em https://validar.iti.gov.br

Profa. Dra. MARINA ALVES GAVASSI (Participaçao Virtual)
Instituto de Biociências de Rio Claro-IB-UNESP / Rio Claro/SP

Profª. Drª. MAYRA COSTA DA CRUZ GALLO DE CARVALHO (Participaçao Virtual)
Centro ce Ciências Biológicas / Universidade Estadual do Norte do Paraná

Profª. Drª ILARA GABRIELA FRASSON BUDZINSKI (Participaçao Virtual)
Genética / Escola Superior de Agricultura Luiz de Queiroz - ESALQ/USP

Profª. Drª. MIRIAN PEREZ MALUF (Participaçao Virtual)
Embrapa Café / Instituto Agronômico de Campinas

Rio Claro, 25 de maio de 2023

AGRADECIMENTOS

Eu quero agradecer primeiramente a Deus por ter me dado força, saúde, perseverança e energia para concluir mais uma etapa da minha formação acadêmica, a qual considero uma grande conquista.

Aos meus pais Vilmar Rota Camargo e Surli Oliveira Camargo, por serem a minha base, o meu apoio e por sempre acreditarem em mim.

Ao meu marido Adriano Polican Ciena, por todo apoio, ensinamentos e companheirismo. Ele foi uma peça fundamental para a conclusão da minha tese. E sempre será a minha inspiração, por ser um estimado professor e pesquisador.

Aos meus familiares (tios, tias, primos e primas), que mesmo de longe, sempre torceram pelo meu sucesso.

Aos meus amigos de Londrina, que sempre me encentivaram a persistir na minha carreira acadêmica.

Aos meus queridos ex e atuais alunos, que de alguma forma, sempre me mandam energia positiva.

Aos meus ex-alunos e agora amigos do coração, Mathilde, Nathália e Lauro, que sempre me fizeram acreditar que eu sou capaz de exercer uma boa atividade acadêmica.

Aos meus amigos de laboratório, Natacha, Gian, Samara, Thalita, Daniel, Raíssa, Ilara, Ingrid, Ana, Matheus e em especial a Natália Calzado, que acompanharam mais de perto o andamento da pesquisa, colaborararam com palavras de incentivo e trocas de conhecimentos e experiências acadêmicas.

Ao meu Orientador Dr. Douglas Silva Domingues, por todo ensinamento do conhecimento científico. Por toda ajuda durante os anos de doutorado. Por ter aceitado a orientação e por ter deixado eu continuar a profissão em que eu mais amo (professora) durante o doutorado.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de

Financiamento 001. Também através do processo **nº2016/10896-0, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).** E juntamente ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (Processo nº 170155/2018-8) .

Dedicatória

Ao meu maior amor e ao meu maior orgulho...

Meu filho, Davi Camargo Ciena.

Resumo

As lipoxigenases (LOXs) são enzimas que desempenham diversas funções fisiológicas nos vegetais, incluindo crescimento e desenvolvimento, e estão relacionadas com vias de defesa vegetal. O ácido jasmônico é um hormônio sinalizador para ativação das LOXs em resposta a ataques de herbívoros ou patógenos. Elicitores, como o ácido hexanoico, também podem sinalizar a ativação de vias de jasmonatos e, consequentemente, a ativação das LOXs. As LOXs são codificadas por uma família gênica, geralmente estudada individualmente em uma espécie ou em poucas espécies relacionadas. No entanto, a família gênica LOX nunca foi estudada em detalhe no gênero *Coffea*. Diante deste cenário, o objetivo deste estudo foi analisar a diversidade, evolução e expressão dos genes LOX em espécies de angiospermas. Além disso, buscamos identificar genes codificadores de enzimas lipoxigenases em três espécies do gênero *Coffea* (*Coffea arabica*, *Coffea canephora* e *Coffea eugenioides*) e avaliar se o elicitor ácido hexanoico pode modular o perfil transcricional da família gênica LOX em *C. arabica*. Identificamos 247 genes LOX entre 23 espécies de angiospermas e plantas basais e em análises filogenéticas identificamos um novo subclado de LOX. Foram identificados 18 genes de lipoxigenases em *Coffea arabica*, enquanto em *Coffea canephora* e *Coffea eugenioides* foram encontrados 9 genes LOX. Nosso trabalho observou que os genes LOX no genoma tetraploide de *Coffea arabica* tem distribuição em seus subgenomas similar aos diploides parentais, *Coffea canephora* e *Coffea eugenioides*. Verificamos que a aplicação exógena de ácido hexanoico pode modular o perfil transcricional de genes de lipoxigenases em folhas e raízes de *C. arabica* cv. Catuaí Vermelho (CV) e *C. arabica* cv. Obatã (OB). Três genes apresentam alta correlação entre a atividade da enzima lipoxigenase e a atividade transcricional de genes LOX. Com isso, esperamos contribuir para o entendimento da diversidade e evolução de LOX em Angiospermas bem como para o entendimento da regulação da defesa vegetal mediada pelas LOXs em plantas de *Coffea*.

Palavras-chave: Famílias multigênicas; diversidade gênica em Angiospermas; LOX em Coffea; elicitor ácido hexanoico.

Abstract

Lipoxygenases (LOXs) are enzymes that perform several physiological functions in plants, including growth and development, and are related to plant defense pathways. Jasmonic acid is a signaling hormone for activation of LOXs in response to herbivore or pathogen attacks. Elicitors, such as hexanoic acid, can also signal the activation of jasmonate pathways and, consequently, the activation of LOXs. LOXs are encoded by a gene family, usually studied individually in a species or a few related species. However, the LOX gene family has never been studied in detail in the genus *Coffea*. Given this scenario, the aim of this study was to analyze the diversity, evolution and expression of LOX genes in angiosperm species. In addition, we sought to identify genes encoding lipoxygenase enzymes in three species of the *Coffea* genus (*Coffea arabica*, *Coffea canephora* and *Coffea eugenioides*) and evaluate whether the hexanoic acid elicitor can modulate the transcriptional profile of the LOX gene family in *C. arabica*. We identified 247 LOX genes among 23 species of angiosperms and basal plants and in phylogenetic analyzes we identified a new subclade of LOX. 18 lipoxygenase genes were identified in Coffea arabica, while in Coffea canephora and Coffea eugenioides 9 LOX genes were found. Our work observed that the LOX genes in the tetraploid genome of *Coffea arabica* have distribution in their subgenomes similar to the parental diploids, *Coffea canephora* and *Coffea eugenioides*. We verified that the exogenous application of hexanoic acid can modulate the transcriptional profile of 12 lipoxygenase genes in leaves and roots of *C. arabica* cv. Catuaí Vermelho (CV) and C. arabica cv. Obatã (OB). Three genes show high correlation between lipoxygenase enzyme activity and transcriptional activity of LOX genes. With this, we hope to contribute to the understanding of the diversity and evolution of LOX in Angiosperms as well as to the understanding of the regulation of plant defense mediated by LOXs in *Coffea* plants.

Keywords: Multigene families; gene diversity in Angiosperms; LOX in Coffea; hexanoic acid elicitor.

Sumário

1. INTRODUÇÃO 10
REFERÊNCIAS 13
2. CAPÍTULO 1 - Genome-Wide Analysis of Lipoxygenase (LOX) Genes in Angiosperms 15
2.1 Introduction 16
2.2 Results 17
2.3 Discussion 25
2.4 Materials and Methods 31
2.4.1 Identification and Annotation of LOX Family Genes 31
2.4.2 Multiple Sequence Alignment and Phylogenetic Analysis 33
2.4.3 Determination of Gene Structures 33
2.4.4 Selection Pressure and Evolutionary Analysis 33
2.4.5 Analysis of LOX Gene Expression Profiles in Angiosperms 34
2.4.6 Investigation of Motif Sequences and Cellular Localization of LOX Genes 34
2.5 Conclusions 34
References 35
3. CAPÍTULO 2 - Genome-wide identification and characterization of the lipoxygenase gene family in the tetraploid *Coffea arabica* L. and its diploid parental genomes 39
3.1 Introdução 41
3.2 Resultados 42
3.2.1 Identificação e análise filogenética de genes LOX em Coffea 42
3.2.2 Mapa cromossômico dos genes de *Coffea arabica, Coffea canephora e Coffea eugenioides* 45
3.2.3 Estrutura de éxons e íntrons dos genes LOX em Coffea 47
3.2.4 Atividade enzimática de lipoxigenases em *C. arabica* cv Catuaí Vermelho e Obatã 48
3.2.7 Análises de correlação entre os valores de RNAseq, qPCR e atividade enzimática 54
3.3 Discussão 57
3.4 Material e Métodos 62
3.4.1 Identificação de Genes da Família LOX inclusos na árvore evolutiva 62

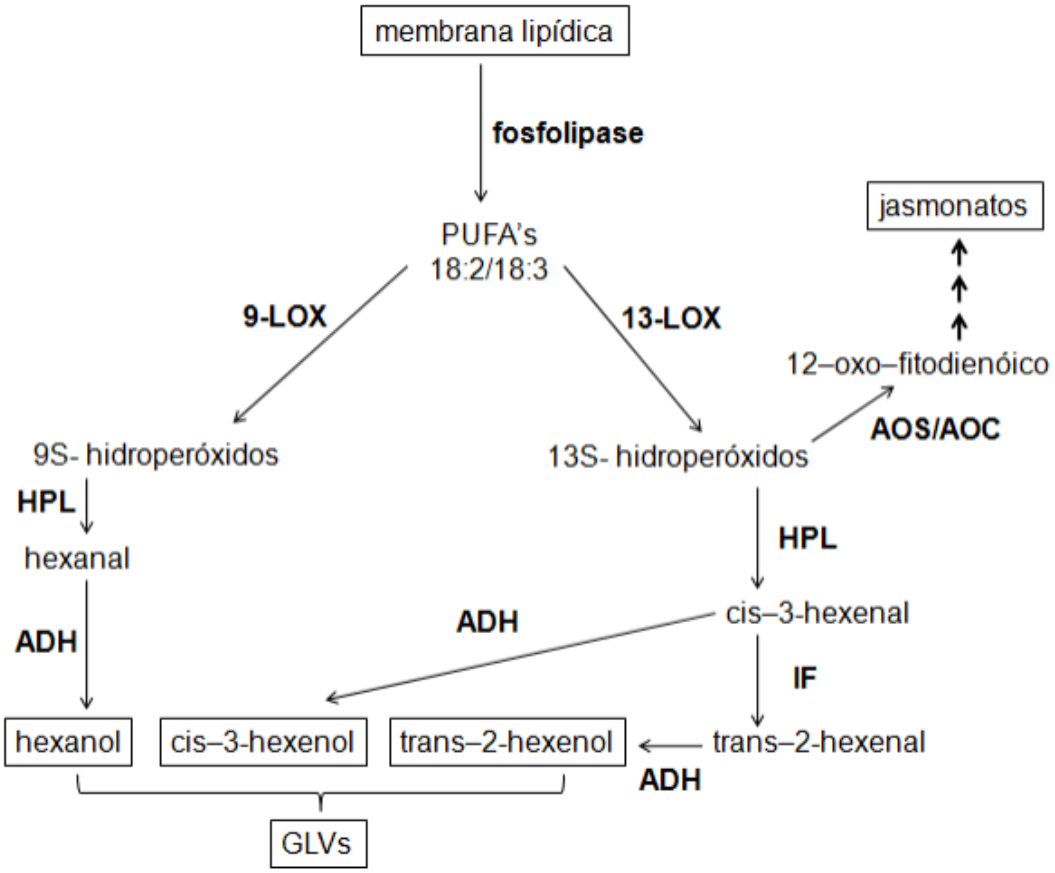
3.4.2 Alinhamento de Múltiplas Sequências e Análise Filogenética 62
3.4.3 Determinação das estruturas genéticas em *Coffea arabica, Coffea canephora* e *Coffea eugenioides* 63
3.4.4 Localização Celular de Genes LOX em *Coffea arabica, Coffea canephora* e *Coffea eugenioides* 63
3.4.5 Mapeamento dos genes LOX 63
3.4.6 Experimento para análise transcricional de genes LOX em *Coffea arabica*: efeito do ácido hexanoico 64
3.4.7 Extração de RNA total para preparo da biblioteca - RNA-Seq e para as análises de RT- qPCR 64
3.4.8 Preparo da biblioteca de RNAseq e sequenciamento..... 65
3.4.9 Desenho dos primers para sequências de *Coffea arábica* 65
3.4.10 Análises do perfil transcricional de genes LOX por RT-qPCR 66
3.4.11 Determinação da atividade enzimática de lipoxigenase – extrato foliar 67
3.4.12 Determinação da atividade de lipoxigenases..... 67
3.4.13 Análises de correlação para proteínas Lipoxigenases..... 68
3.5 Conclusões 68
4. CONSIDERAÇÕES FINAIS 69
Apêndice I 75
Apêndice II 83
Apêndice III 98

1. INTRODUÇÃO

A lipoxigenase (LOX) é uma enzima oxidoredutase amplamente distribuída em plantas. Essas enzimas já foram identificadas em diversas espécies vegetais (UMATE et al., 2011; FENG et al., 2010; ZHU et al., 2018; SHABAN et al., 2018; SARDE et al., 2018). O peso molecular de LOX nas espécies vegetais varia entre 90 – 110 kDa e as reações mediadas por essas enzimas incorporaram oxigênio molecular no carbono 9 (9-LOX) ou 13 (13-

LOX) de ácidos graxos poliinsaturados (PUFAs) contendo 18 carbonos (SINGH et al., 2022). Na Figura 1, apresentamos as reações mediadas por estas enzimas de maneira resumida.

Figura 1. Diagrama das reações básicas da via da lipoxigenase. Caminho para a biossíntese do ácido jasmônico (JA) e GLVs (*Green Leaf Volatiles*).



Legenda: LOX – lipoxigenase, AOS – aleno óxido sintase, AOC – aleno óxido ciclase, HPL – hidroperóxidoliase, ADH – álcool desidrogenase, IF – fator de isomerização. Adaptado de BATE; ROSTHSTEIN, 1998 e GAO; KOLOMIETS, 2009.

As enzimas LOX, presentes nas plantas, são proteínas estáveis geralmente encontradas nas folhas. Essas enzimas produzem hidroperóxidos de ácidos graxos nos vegetais, que podem ser metabolizados em três tipos de derivados. Um deles é a mistura de ácidos graxos epóxi e hidroxi, formada através da co-oxidação com peroxidase, que tem importância na constituição da cutina.

Outra reação envolve as enzimas hidroperóxidos liases, que quebram os hidroperóxidos, resultando na formação de um aldeído e um ácido graxo oxo-insaturado. Esses compostos são formados nas plantas como resposta à ferimentos, proteção contra patógenos e também participam das respostas ao estresse abiótico, estimulando a expressão de genes relacionados ao estresse.

Portanto, diversas funções fisiológicas relacionadas ao desenvolvimento vegetal, como a maturação dos estames em *Arabidopsis thaliana* (Acosta & Przybyl, 2019), proteção contra patógenos, como visto em *Oryza sativa* (Liao et al., 2022), e resposta a estresses abióticos, como relatado por Liu et al. (2021) em bananas, dependem da associação das enzimas lipoxigenases com os hormônios jasmonatos.

Foi observado que lipoxigenases desempenham um papel importante nos processos de defesa vegetal em folhas de *Coffea arabica* submetidas à herbivoria (Meriño-Cabrera et al., 2018). Da mesma forma, essas enzimas também atuam na defesa bioquímica natural em frutos de café (*Coffea arabica*) cultivados organicamente, que são mais suscetíveis a infecções por patógenos (Patui et al., 2007).

Além dessas funções, Ding et al. (2019) descreveram que algumas isoenzimas de lipoxigenases podem co-oxidar carotenoides. Estes carotenoides captam energia luminosa para a fotossíntese e ajudam a proteger as plantas de espécies reativas de oxigênio. Até o momento, os estudos evolutivos sobre os genes LOX em plantas têm se limitado a uma única espécie ou a grupos de espécies relacionadas. Por isso, análises filogenéticas mais detalhadas, como as realizadas neste estudo, são necessárias para melhor compreender a relação entre isoformas de lipoxigenases em plantas.

O gênero *Coffea* (Rubiaceae) tem 124 espécies, das quais duas apresentam um maior interesse econômico: *Coffea arábica* (cafeeiro arábica) e *Coffea canephora* (Robusta ou Conilon). *C. arabica* é uma espécie alotetraploide (2n = 4x = 44), derivado de uma fusão genômica entre as espécies *C. canephora* (Robusta ou Conilon) e *C. eugenioides* (espécie selvagem) (SATTLER et al., 2022). Em *Coffea canephora*, genes de lipoxigenases envolvidos em rotas metabólicas de produção do ácido jasmônico já foram relatados por BHARATHI et al. (2017).

Há indícios de que as vias de biossíntese de jasmonatos podem ser reguladas em resposta de defesa em plantas submetidas ao ácido hexanoico (C6H12O2) (SCALSCHI et al., 2013; FINITI et al., 2014). Esse ácido - também conhecido como ácido caproico, por conta do seu odor característico (http://gestis-en.itrust.de/nxt/gateway.dll/gestis en/028160.xml) é um composto de seis carbonos, derivado do hexano. Respostas de defesas induzidas por ácido hexanoico já foram relatadas em diferentes sistemas vegetais, como em tomate e batata (CAMAÑES et al., 2015; LÓPEZ-GALIANO et al, 2019). ARANEGA-BOU et al. (2014), postularam que a aplicação de baixas concentrações (faixa de 1 a 5mM) de ácido hexanoico teria um efeito indutor de resistência em vegetais, e que concentrações mais altas (de 6 a 20mM) teriam um efeito bactericida e fungicida, mas sem efeitos de toxicidade em plantas.

Diante do exposto, nosso trabalho teve como objetivo a análise filogenética e a identificação de genes LOX em 23 espécies de angiospermas (eudicotiledôneas, monocotiledôneas e plantas basais), assim como a identificação de genes de lipoxigenases em três espécies de *Coffea*, e a avaliação se o ácido hexanoico modula genes LOX em *C. arabica*.

A apresentação dos resultados desta tese está dividida em dois capítulos. No primeiro capítulo apresentamos o manuscrito publicado na revista *Plants*, que versa sobre as análises evolutivas de LOX em Angiospermas. No segundo capítulo, apresentamos os resultados referentes à identificação de LOX em *Coffea* e análises transcricionais em *C. arabica*.

Suplementarmente, são ainda colocados como apêndice outros trabalhos desenvolvidos ao longo desta tese, relacionados às respostas transcricionais do metabolismo em *C. arabica*, como no metabolismo de terpenoides (SILVA et al., 2020). Entre eles, destacamos a geração do transcriptoma em larga escala de *C. arabica* em resposta ao elicitor ácido hexanoico (BUDZINSKI et al., 2021; BUDZINSKI 2022), que foram fontes de dados para o capítulo 2. Estes materiais estão anexados (Apêndice I).

REFERÊNCIAS

Acosta, I. F.; Przybyl, M. Jasmonate Signaling during Arabidopsis Stamen Maturation. *Plant Cell Physiol* **2019**, 12, 2648-2659. doi: 10.1093/pcp/pcz201

Aranega-Bou, P.; de la O Leyva, M.; Finiti, I.; García-Agustín, P.; González-Bosch, C. Priming of plant resistance by natural compounds. Hexanoic acid as a model. *Front Plant Sci* **2014**, 5 488. doi: 10.3389/fpls.2014.00488

Bate, N. J.; rothstein, S. C6 – volatiles derived from the lipoxygenase pathway induce a subset of defense – related genes. The Plant Journal, **1998**, 16, 5, 561-569

Bharathi, K.; Sreenath, H. L. Identification and Analysis of Jasmonate Pathway Genes in *Coffea canephora* (Robusta Coffee) by *In Silico* Approach. *Pharmacogn Mag* **2017**, 13, S196-S200. doi: 10.4103/pm.pm_518_16

Camañes, G.; Scalschi, L.; Vicedo, B.; González-Bosch, C.; García-Agustín, P. An untargeted global metabolomic analysis reveals the biochemical changes underlying basal resistance and priming in Solanum lycopersicum, and identifies 1-methyltryptophan as a metabolite involved in plant responses to Botrytis cinerea and Pseudomonas syringae. *Plant J* **2015** 1,125-39. doi: 10.1111/tpj.12964

Christie, W. W. Harwood, J. L. Oxidation of polyunsaturated fatty acids to produce lipid mediators. *Essays Biochem* **2020**, 3, 401-421. doi: 10.1042/EBC20190082

Ding, Y.; Yang, W.; Su, C.; Ma, H.; Pan, Y.; Zhang, X.; Li, J. Tandem 13-Lipoxygenase Genes in a Cluster Confers Yellow-Green Leaf in Cucumber. *Int. J. Mol. Sci* **2019**, *20*, 3102. doi.org/10.3390/ijms20123102

Feng, B.; Dong, Z.; Xu, Z.; An, X.; Quin, H.; Wu, N.; Wang, D.; Wang, T. Molecular analysis of lipoxygenase (LOX) genes in common wheat and phylogenetic investigation of LOX proteins from model and crop plants. *J. Cereal Sci* **2010**, *52*, 387–394. doi.org/10.1016/j.jcs.2010.06.019

Finiti, I.; de la O Leyva, M.; Vicedo, B.; Gómez-Pastor R.; López-Cruz, J.; García-Agustín, P.; Real, M. D.; González-Bosch, C. Hexanoic acid protects tomato plants against Botrytis cinerea by priming defence responses and

reducing oxidative stress. *Mol Plant Pathol* **2014**, 6, 550-62. doi: 10.1111/mpp.12112

Gao, X.; kolomiets, M. V. Host – derived lipids and oxylipins are crucial signals in modulating mycotoxin production by fungi. *Toxin Reviews* **2009**, v. 28, p. 79 – 88. doi.org/10.1080/15569540802420584

Liao, Z.; Wang, L.; Li, C.; Cao, M.; Wang, J.; Yao, Z.; Zhou, S.; Zhou, G.; Zhang, D.; Lou, Y. The lipoxygenase gene OsRCI-1 is involved in the biosynthesis of herbivore-induced JAs and regulates plant defense and growth in rice. *Plant Cell Environ* **2022**, 9, 2827-2840. doi: 10.1111/pce.14341

Liu, F.; Li, H.; Wu, J.; Wang, B.; Tian, N.; Liu, J.; Sun, X.; Wu, H.; Huang, Y.; Lü, P.; Cheng, C. Genome-wide identification and expression pattern analysis of lipoxygenase gene family in banana. *Sci Rep* **2021**, 1, 9948. doi: 10.1038/s41598-021-89211-6

López-Galiano, M. J.; García-Robles, I.; Ruiz-Arroyo, V. M.; Oltra, S. S.; Marko Petek, M.; Rausell, C.; Real, M. D. Colorado potato beetle chymotrypsin genes are differentially regulated in larval midgut in response to the plant defense inducer hexanoic acid or the Bacillus thuringiensis Cry3Aa toxin. *Journal of Invertebrate Pathology* **2019**, 166, 107224. doi.org/10.1016/j.jip.2019.107224

Patui, S.; Peresson, C.; Braidot, E.; Tubaro, F.; Colussi, A.; Bonnlander, B.; Macri, F.; Vianello, a. Lipoxygenase Distribution in Coffee (*Coffea arabica* L.) Berries**.** *Journal of Agricultural and Food Chemistry* **2007**, 55, 20, 8223-8230. doi: 10.1021/jf070982s

Sarde, S.J.; Kumar, A.; Remme, R.N.; Dicke, M. Genome-wide identification, classification and expression of lipoxygenase gene family in pepper. *Plant Mol. Biol.* **2018**, *98*, 375–387. doi.org/10.1007/s11103-018-0785-y

Sattler, M. C., de Oliveira, S. C.; Mendonça, M. A. C.; Clarindo, W. R. Coffea cytogenetics: from the first karyotypes to the meeting with genomics. *Plants* **2022**, 6, 112. doi: 10.1007/s00425-022-03898-z.

Scalschi, L.; Vicedo, B.; Camañes, G.; Fernandez-Crespo, E.; Lapeña, L.; González-Bosch, C.; García-Agustín, P. Hexanoic acid is a resistance inducer that protects tomato plants against Pseudomonas syringae by priming the jasmonic acid and salicylic acid pathways. *Mol Plant Pathol*, **2013**, 4, 342-55. doi: 10.1111/mpp.12010

Shaban, M.; Ahmed, M.M.; Sun, H.; Ullah, A.; Zhu, L. Genome-wide identification of lipoxygenase gene family in cotton and functional characterization in response to abiotic stresses. *BMC Genom.* **2018**, *19*, 599. doi.org/10.1186/s12864-018-4985-2

Silva, N.; Ivamoto-Suzuki, S.T.; Camargo, P.O.; Rosa, R.S.; Pereira, L.F.P.; Domingues, D.S. Low-Copy Genes in Terpenoid Metabolism: The Evolution and Expression of *MVK* and *DXR* Genes in Angiosperms. *Plants* **2020**, *9*, 525. https://doi.org/10.3390/plants9040525

Singh, P.; Arif, Y.; Miszczuk, E.; Bajguz, A.; Hayat, S. Specific Roles of Lipoxygenases in Development and Responses to Stress in Plants. *Plants* (Basel). **2022**, 7, 979. doi: 10.3390/plants11070979

Umate, P. Genome-wide analysis of lipoxygenase gene family in Arabidopsis and rice. *Plant Signal Behav.* **2011**, *6*, 335–338. doi.org/10.4161/psb.6.3.13546

Vogt, J.; Schiller, D.; Ulrich, D.; Schwab, W.; Dunemann, F. Identification of lipoxygenase (LOX) genes putatively involved in fruit flavour formation in apple (*Malus* × *domestica*). *Tree Genet. Genomes* **2013**, *9*, 1493–1511. doi.org/10.1007/s11295-013-0653-5

Zhu, J.; Wang, X.; Guo, L.; Xu, Q.; Zhao, S.; Li, F.; Yan, X.; Liu, S.; Wei, C. Characterization and alternative splicing profiles of the lipoxygenase. Gene Family in Tea Plant (*Camellia sinensis*). *Plant Cell Physiol.* **2018**, *59*, 1765–1781 https://doi.org/10.1093/pcp/pcy091

2. CAPÍTULO 1 - Genome-Wide Analysis of Lipoxygenase (LOX) Genes in Angiosperms

Abstract

Lipoxygenases (LOXs) are enzymes that catalyze the addition of an oxygen molecule to unsaturated fatty acids, thus forming hydroperoxides. In plants, these enzymes are encoded by a multigene family found in several organs with varying activity patterns, by which they are classified as LOX9 or LOX13. They are involved in several physiological functions, such as growth, fruit development, and plant defense. Despite several studies on genes of the LOX family in plants, most studies are restricted to a single species or a few closely related species. This study aimed to analyze the diversity, evolution, and expression of LOX genes in angiosperm species. We identified 247 LOX genes among 23 species of angiosperms and basal plants. Phylogenetic analyses identified clades supporting LOX13 and two main clades for LOX9: LOX9_A and LOX9_B. Eudicot species such as *Tarenaya hassleriana*, *Capsella rubella*, and *Arabidopsis thaliana* did not present LOX9_B genes; however, LOX9_B was present in all monocots used in this study. We identified that there were potential new subcellular localization patterns and conserved residues of oxidation for LOX9 and LOX13 yet unexplored. In summary, our study provides a basis for the further functional and evolutionary study of lipoxygenases in angiosperms.

Keywords: lipoxygenase gene family; angiosperms; purifying selection

2.1 Introduction

Lipoxygenases (LOXs; EC 1.13.11.12) are enzymes belonging to the class of oxidoreductases that catalyze the addition of an oxygen molecule to unsaturated fatty acids, thus forming hydroperoxides that decompose into short-chain acids, aldehydes, and ketones. The most common plant fatty acids broken down by LOXs are linoleic and linolenic acids. LOXs are widely present in living organisms, occurring in bacteria, fungi, animals, and plants [1].

In plants, LOXs are found in several organs in varying concentrations; they are involved in several physiological functions including growth and development, vegetative reserve, senescence, resistance to insects and pathogens, seed germination, and as precursors of hormones and volatile substances [2]. It is known that lipoxygenase proteins effectively participate in the biosynthesis of the plant hormone jasmonate. Therefore, several physiological functions in plants depend on the association of these enzymes with this hormone. In tobacco plants, lipoxygenases are associated with responses involved in plant defense and resistance to stress through their regulatory elements such as methyl jasmonate (MeJA). Lipoxygenases are also involved, through MeJA biosynthesis, in metabolic pathways that regulate the transcription of the leaf senescence process, a fact observed in experiments carried out with the model species *Arabidopsis thaliana*. In *Cucurbita pepo*, the hormone jasmonate, synthesized by lox3a, controls petal elongation and flowering opening as well as fruit abortion in the absence of fertilization [3,4,5].

In higher plants, LOX enzymes can produce fatty acid hydroperoxides through two pathways known as the LOX pathways. The hydroperoxides formed are reactive molecules that can be mobilized in higher plants by enzymatic complexes involving enzymes such as hydroperoxide cyclase and hydroperoxide lyase. The latter, in turn, produces six-carbon compounds such as trans-2-hexenal, which is a characteristic component of fruit flavor and odor. Twelve-carbon compounds can also be produced by this enzyme, such as thaumatin, which is involved in signaling and cell division processes in

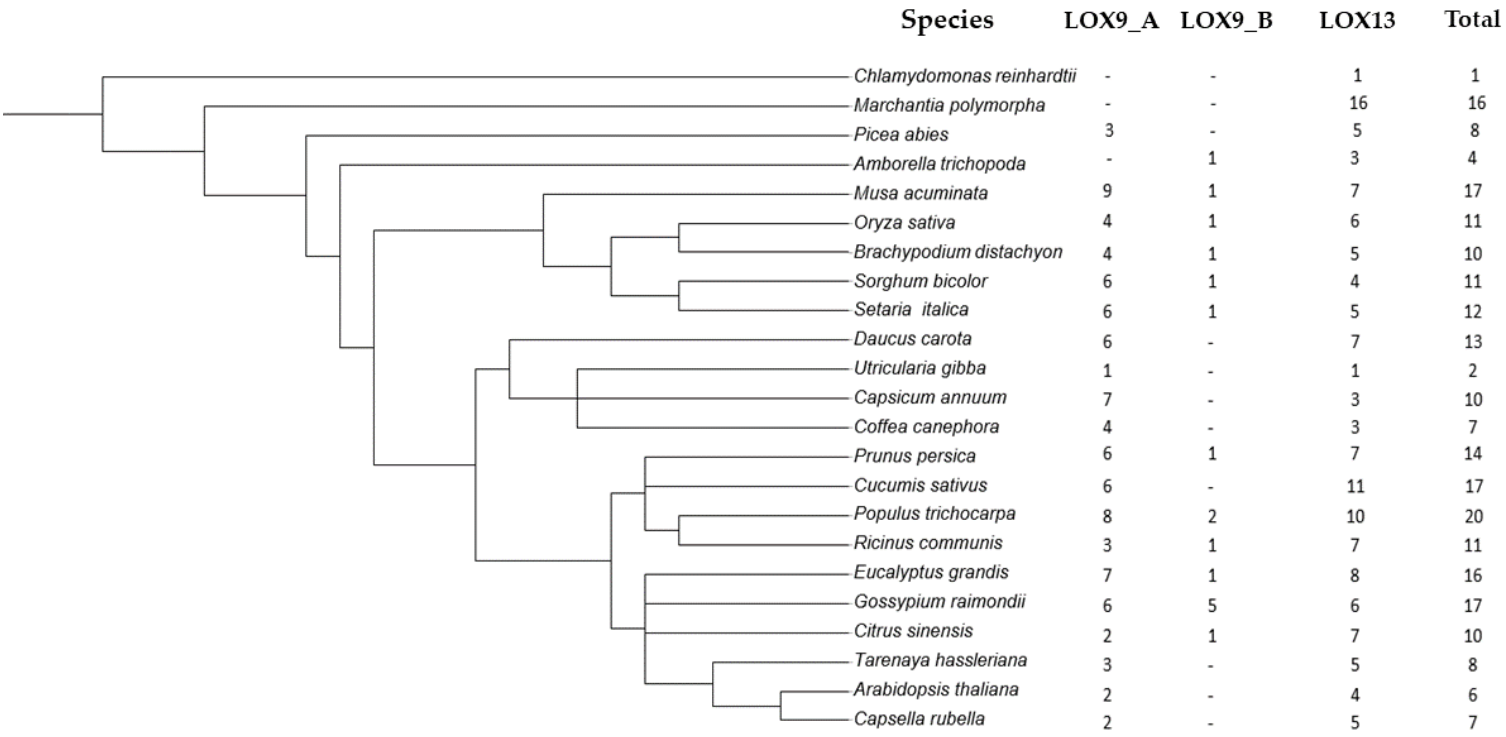
response to plant injuries [6]. To date, LOXs have been classified according to their oxidation position of polyunsaturated fatty acids — LOX9 and LOX13 are responsible for the oxygenation of linoleic acid at carbons 9 and 13, respectively — or based on their cellular location — LOX type I was found in the cytoplasm, and LOX type II in the organelle-targeting signal peptides [7]. *Arabidopsis thaliana*, a reference plant for the evolutionary analysis presented in this study, contains six LOX genes, of which two are of the LOX9 type and four are of the LOX13 type [8].

So far, evolutionary studies on LOX genes in plants are restricted to a single species or a few closely related species [7,9,10]. Given this context, more detailed phylogenetic analyses were performed in this study using LOX members from 23 angiosperm plant species to comprehensively assess the relationships between plants and LOX enzymes.

2.2 Results

A total of 247 LOX genes were found among 23 plant species: *Arabidopsis thaliana*, *Citrus sinensis*, *Capsella rubella*, *Gossypium raimondii*, *Tarenaya hassleriana*, *Prunus persica*, *Eucalyptus grandis*, *Ricinus communis*, *Cucumis sativus*, *Capsicum annuum*, *Utricularia gibba*, *Daucus carota*, *Coffea canephora*, *Brachypodium distachyon*, *Setaria italica*, *Populus trichocarpa*, *Oryza sativa* ssp. *japonica*, *Musa acuminata*, *Sorghum bicolor*, *Picea abies*, *Marchantia polymorpha*, *Amborella trichopoda*, and *Chlamydomonas reinhardtii* (Table S1).

In eudicots, the number of LOX genes varied between two (*Utricularia gibba*) and twenty (*Populus trichocarpa*), with an average number of genes of 11.29. In monocots, the number of genes varied between 10 (*Brachypodium distachyon*) and 16 (*Musa acuminata*); the average was 12 genes. In basal plants, the number of LOX genes varied between one (*Chlamydomonas reinhardtii*) and sixteen (*Marchantia polymorpha*), and the mean number of genes was 7.25 (Figure 1).



Species	LOX9_A	LOX9_B	LOX13	Total
Chlamydomonas reinhardtii	-	-	1	1
Marchantia polymorpha	-	-	16	16
Picea abies	3	-	5	8
Amborella trichopoda	-	1	3	4
Musa acuminata	9	1	7	17
Oryza sativa	4	1	6	11
Brachypodium distachyon	4	1	5	10
Sorghum bicolor	6	1	4	11
Setaria italica	6	1	5	12
Daucus carota	6	-	7	13
Utricularia gibba	1	-	1	2
Capsicum annuum	7	-	3	10
Coffea canephora	4	-	3	7
Prunus persica	6	1	7	14
Cucumis sativus	6	-	11	17
Populus trichocarpa	8	2	10	20
Ricinus communis	3	1	7	11
Eucalyptus grandis	7	1	8	16
Gossypium raimondii	6	5	6	17
Citrus sinensis	2	1	7	10
Tarenaya hassleriana	3	-	5	8
Arabidopsis thaliana	2	-	4	6
Capsella rubella	2	-	5	7

Figure 1. Number of LOX genes distributed among angiosperm groups. Fourteen species of eudicots, five species of monocots, and four basal species were analyzed.

The evolutionary tree was constructed based on amino acid sequence alignments. The LOX genes were divided into three groups with bootstrap support above 90%. Therefore, according to our data, we proposed a new nomenclature of the clades as follows: LOX13 group, LOX9_A (previously called LOX9), and LOX9_B (Figure 2).

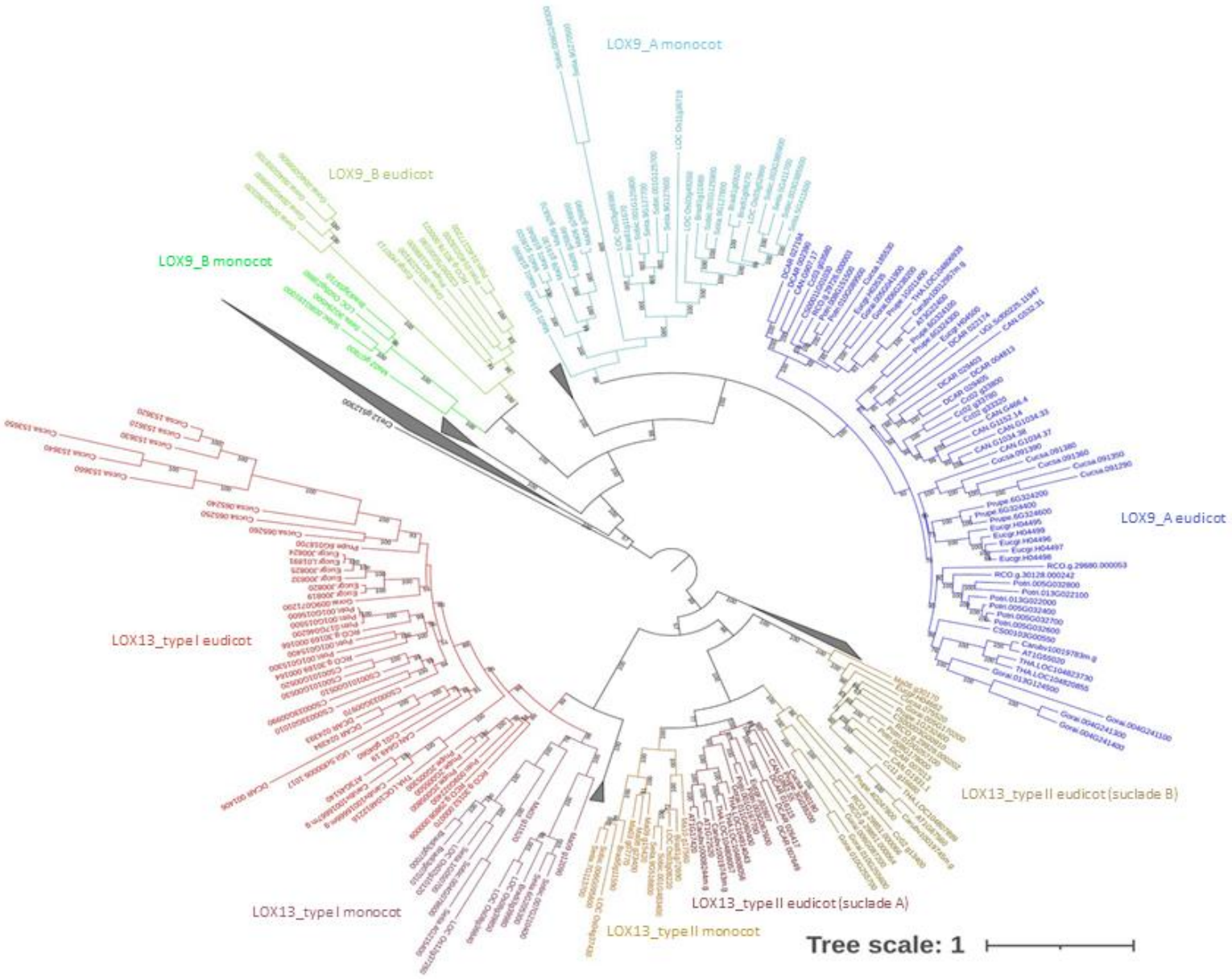


Figure 2. Evolutionary tree (maximum likelihood method, 1000 replicates per bootstrap) of lipoxygenases from 23 plant species. To facilitate visualization, we included the name of each clade with its corresponding color. Legend: eudicot LOX13_type I is in light-red, monocot LOX13_type I is in light-purple, monocot LOX13_type II is in light-brown, eudicot LOX13_type II (subclade A) is in dark-purple, and eudicot LOX13_type II (subclade B) is represented in dark-brown color. The LOX9_A group's respective division among angiosperms is represented with the following colors: eudicot LOX9_A—dark blue and monocot LOX9_A—light blue. The LOX9_B group is represented in two shades of green: eudicot LOX9_B—dark green and monocot LOX9_B—light green.

We found inconsistencies within the two groups of LOX13, which until now were classified as LOX13 type I and LOX 13 type II. Although LOX13 type II presented a signal peptide for targeting organelles, LOX13 type I could also present a signal peptide for signaling in organelles. Thus, it was possible to

infer that cellular localization was not the main and only mode used to classify LOX13 proteins. The LOX13 type II genes showed two distinct subclades for division into monocots and eudicots. However, the LOX13 type I genes showed two subdivisions for eudicots and one for monocots, indicating a more complex evolution of the type I LOX13s. We also observed that the LOX9 clade had two distinct main sub-clades. There was a highly supported phylogenetic sub-clade, this being a sub-clade supported by external groups (*Amborella trichopoda* and *Picea abies*) with a distinct division between monocots and dicots; however, we found a regular distribution among the species of angiosperms, with an expansion of LOX genes in *Gossypium raimondii*. All the monocots used in this study had at least one representative of LOX9 in this sub-clade, but the same was not observed for the eudicots since not all the species had at least one representative of this sub-clade. The only angiosperm species that presented at least one copy of the LOX gene for this sub-clade were *Prunus persica*, *Populus trichocarpa*, *Ricinus communis*, *Eucalyptus grandis*, *Citrus sinensis*, and *Gossypium raimondii*. In the Brassicaceae species, including the model *Arabidopsis thaliana*, representative LOX genes for this putative new clade were detected.

We identified that representatives of the LOX9_B subgroup had, in one of their oxidation domains, a specific site for the conservation of the amino acids leucine (L) or methionine (M) (Figure 3), rather than valine (V), observed in all the members of the LOX9_A group.

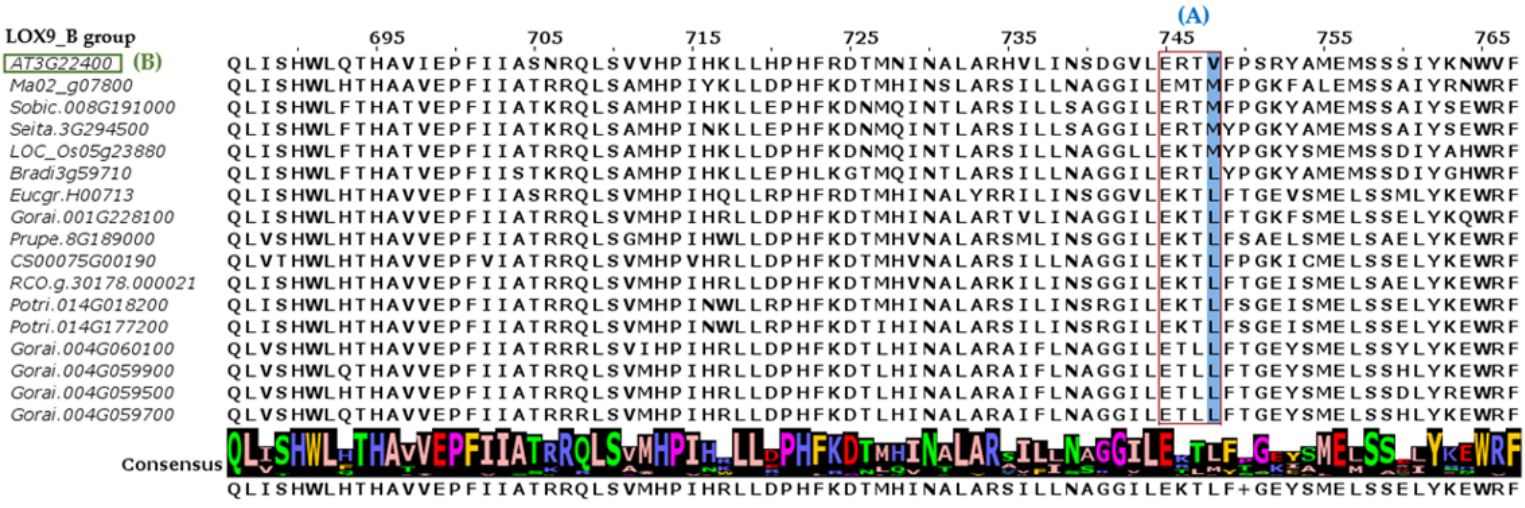


Figure 3. LOX9_B group and its pattern of amino acid conservation in motif sequences (**A**). This subclade had a specific conservation site for amino acids leucine (L) or methionine (M) (blue column). Diverging from the conservation pattern of LOX9

proteins in angiosperm species, including *Arabidopsis thaliana*, represented by AT3G22400 (**B**), a LOX9 gene, which had, in the same place (blue column), the amino acid valine (V), was conserved. We used MAFFT software version 7 for alignment and JALview for visualization.

LOX proteins were also analyzed based on their cellular location. LOX9_A and LOX9_B were 100% identified as cytoplasmic proteins. LOX13 type I (previous nomenclature) proteins were identified as cytoplasmic proteins (42%), chloroplasts (53%), or proteins present in another cellular compartment (5%). This same classification was observed for the LOX13 type II (previous nomenclature) proteins, such as cytoplasmic proteins (58%), chloroplasts (40%), or proteins present in other cellular compartments (2%) (Figure 4). Details are also available at https://doi.org/10.5281/zenodo.7374887 (accessed on 3 January 2022).

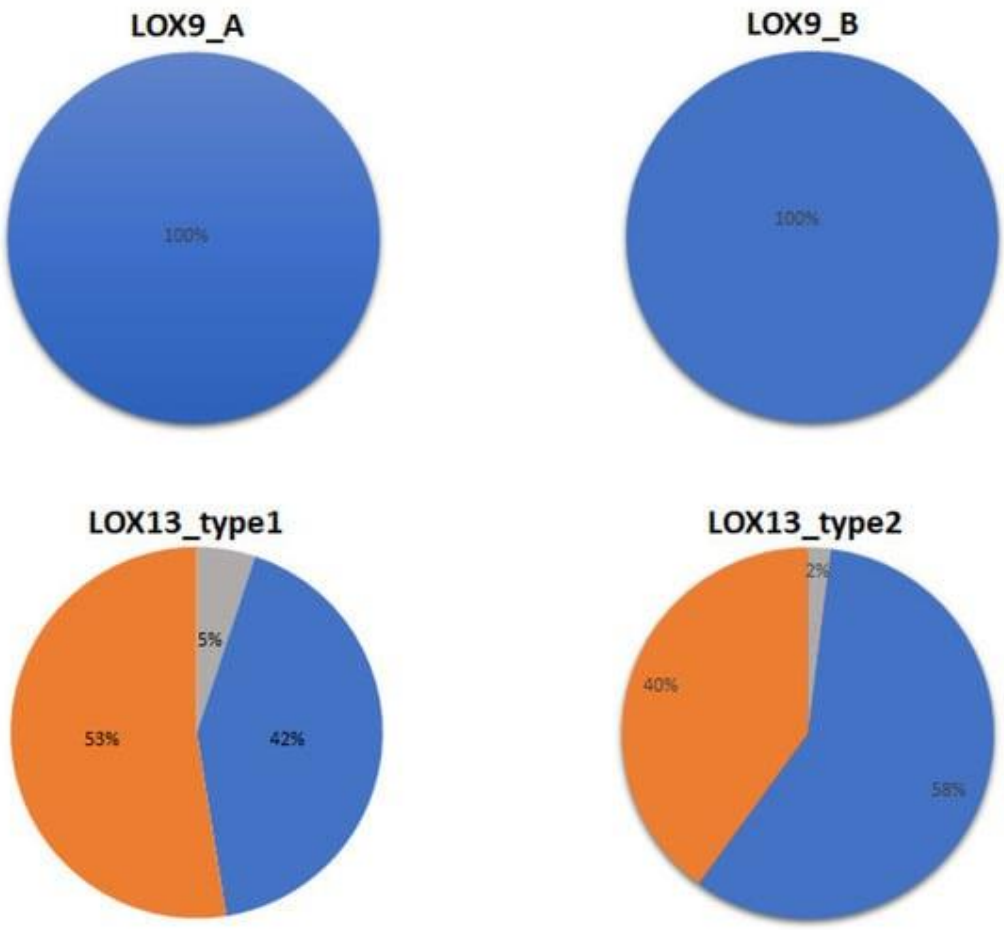


Figure 4. Subcellular localization of LOX proteins (in %). Proteins with cytoplasmic localization are represented in blue, proteins with signal peptide targeting chloroplasts are represented in orange, and proteins directed to other compartments are represented in gray.

An analysis of the individual selection profile of each amino acid, as well as dN and dS substitution of the LOX genes in the eudicot and monocot species, was performed to verify the possibility of different evolutionary pressures. The clades LOX9_A (eudicotyledon and monocotyledon), LOX13 (A) (monocotyledon), and LOX13 (B) (eudicotyledon and monocotyledon) showed diversifying selection with a dN/dS value greater than 1. The clades LOX9_B (eudicotyledon and monocotyledon) and LOX13 (B) (eudicotyledonous) showed a purifying selection with a dN/dS value less than 1 (Figure 5).

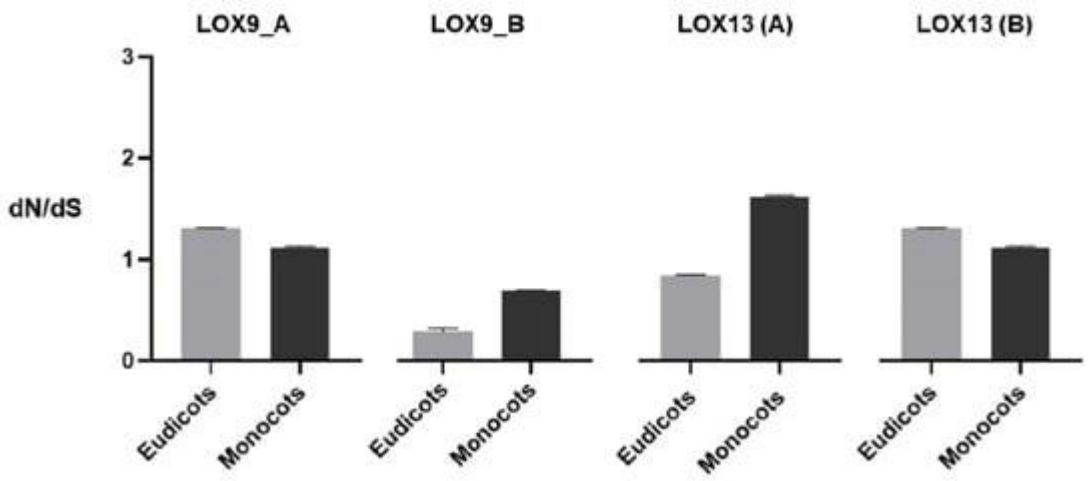


Figure 5. Comparison between the ratios of non-synonymous (dN) and synonymous (dS) substitutions between the LOX groups. Values obtained through the MEGAX program. Calculations performed on the Datamonkey platform ($p < 0.05$) using models of positive selections and dN/dS replacement ratios of LOX genes in eudicots and monocots.

Therefore, the selection models FEL, FUBAR, MEME, and SLAC were grouped in Venn diagrams (Figure 6 and Figure 7).

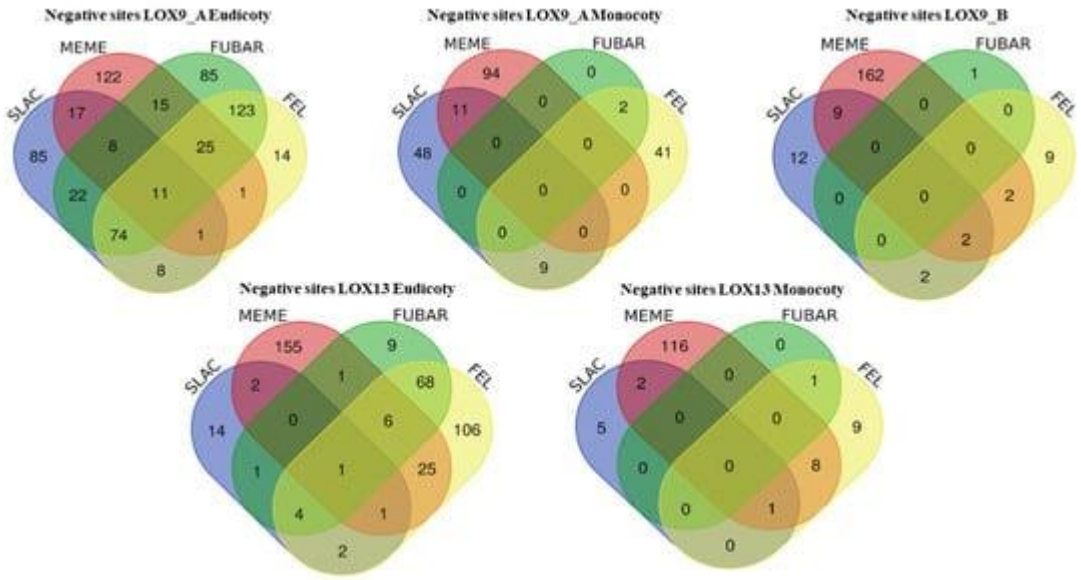


Figure 6. Venn diagrams. Comparison of negative sites between groups of lipoxygenases (LOX9_A eudicot, LOX9_A monocot, LOX9_B, LOX13 eudicot, and LOX13 monocot).

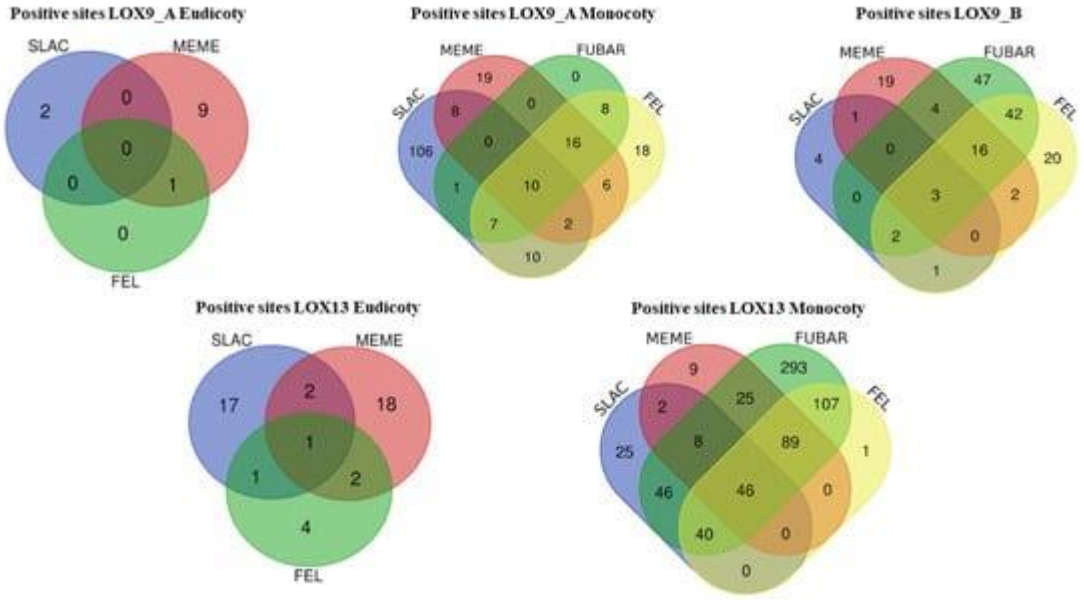


Figure 7. Venn diagram. Comparison of positive sites between groups of lipoxygenases (LOX9_A eudicot, LOX9_A monocot, LOX9_B, LOX13 eudicot, and LOX13 monocot).

For the LOX9_A eudicots, eleven negative positions were common for the four models analyzed (FEL, FUBAR, MEME, and SLAC), and one positive position was common for MEME and FEL. In the LOX9_A monocots, the

largest number of negative positions (11) was shared between the MEME and SLAC models, and the largest number of positive positions (16) was shared between the FEL, FUBAR, and MEME models. For LOX9_B, nine negative sites were shared between the SLAC and FEL models and forty-two positive sites were shared between the FUBAR and FEL models. In the LOX13 eudicots, the largest number of negative positions (68) was shared between the FUBAR and FEL models, and the largest number of positive positions (2) was shared between the SLAC and MEME models. For the LOX13 monocots, the FEL and MEME models shared the largest number of negative positions (8), and the FUBAR and FEL models shared 106 positive positions. Finally, positive selections detected with at least two different methods and moderately supported positive selections with only one method were categorized as strongly supported.

Public RNA-seq data were used to understand the LOX gene expression profiles in the angiosperms. Five plant species, including three eudicots (*Gossypium raimondii*, *Prunus persica*, and *Ricinus communis*) and two monocots (*Brachypodium distachyon* and *Sorghum bicolor*) were chosen based on the available literature. LOX genes were grouped into heatmaps according to their function: #LOX9_A, +LOX9_B, and *LOX13 (Figure 8).

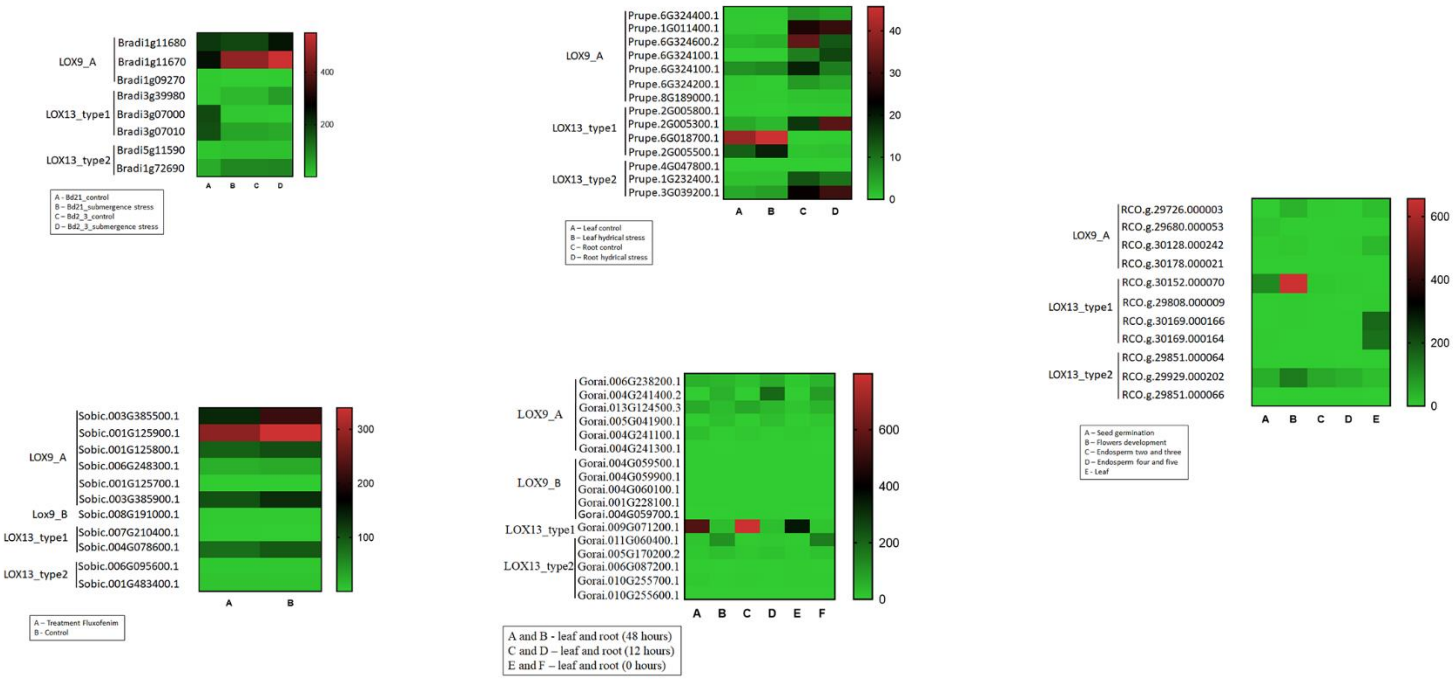


Figure 8. Transcription profile of the LOX gene family members in five angiosperm species (values in TPM—transcription per million). Symbols: LOX9_A, LOX9_B, LOX13_type 1 e LOX13_type 2. (**A**) *Brachypodim distachyon* (monocot.), control, and submersion stress. (**B**) *Prunus persica* (eudicot.), leaf (control and water stress) and root (control and water stress). (**C**) *Sorghum bicolor* (monocot.) control and treatment of Fluxophenim. (**D**) *Gossypium raimondii* (eudicot.), leaf and root (48, 12, and 0 h). (**E**) *Ricinus Communis* (eudicot.), seed germination, flower development, endosperm development II/III, endosperm development IV/V, and leaves. Data were obtained by the CLC Genomics Workbench program).

To better understand the similarities between the LOX gene sequences generated in this study, the structural positions—exons and introns—were obtained for each identified clade (Figure S1). The maximum numbers of exons and introns found were ten and nine, respectively, as shown in DCAR_027194. Moreover, the minimum numbers of exons and introns identified were four and three, respectively, as shown in LOC_Os03g49380. The smallest LOX gene occurred in Cc02_g33800 and Cc02_g33320 (≅5 Kb), and the largest occurred in THA.LOC104807899 (≅29 Kb).

2.3 Discussion

The present study aimed to determine the number of LOX genes in several plant species, expanding the analyses to species that had never been studied, and to confirm the number of LOX genes, thus updating information

regarding previously studied angiosperm species. The number of LOX genes in *Arabidopsis thaliana*, *Brachypodium distachyon*, and *Populus trichocarpa* identified in this study corroborated the number of LOX genes identified in previous studies [8,11,12]. However, we identified annotation errors for the LOX genes for some species. Shaban [**7**] identified fourteen LOX genes in *Gossypium raimondii*; in addition to these, we identified the presence of four additional LOX genes (Gorai.004G059500, Gorai.004G059900, Gorai.004G060100, and Gorai.004G059700); Gorai.004G092100, initially assigned as a LOX gene, was not included considering our parameters, as it did not present the domain IPR001024 (PLAT/LH2) and because its protein did not have a molecular weight of 90–110 kDa [13], thus having, in this sense, a high probability of being a pseudogene.

We identified 14 LOX genes in *Prunus persica*; in previous studies, the LOX copy number for this species ranged between 16 [14] and 12 [2]. The LOX genes ppa002308, ppa001112, and ppa001082 identified by Li [14] were grouped in the same branch in the present evolutionary tree, which led to the hypothesis that these genes are the result of alternative splicing. Using search tools (Blastn), we identified these three sequences as a single gene, coded as Prupe.047800 (PLAZAv.4 code) [15], so the sequences grouped in the previous study [14] may have resulted from alternative processing in *Prunus persica*. Studies on the evolution and regulation of the genes of the LOX family from alternative splicing processes have shown that alternative transcripts are regulated according to the stress variation to which a particular plant is subjected. This way, competitive or compensatory regulation mechanisms between isoforms arise [12].

A total of fourteen LOX genes have been described in *Oryza sativa* ssp. *japonica* [8]; however, in our study, we identified 11 LOX genes. LOC_Os12g37320 (55.29 kDa), LOC_Os02g19790 (50.74 kDa), and LOC_Os06g04420 (14.16 kDa) were not considered genes belonging to the LOX family as they did not present the domains IPR001024 (PLAT/LH2), IPR013819 (LOX, C-terminal), IPR001246 (LOX, plant), and IPR000907 (LOX), and these proteins were not of the average molecular weight (90–110 kDa) for the family. In *Capsicum annuum*, our study identified ten LOX genes, whereas Sarde [10] identified eight LOX genes for this species. The

Capana03g003 sequence (59.16 kDa) was not included in our data because it did not present the average molecular weight for the LOX proteins. We used the strategies of repredicting the exon–intron structures of this gene in order to check if this was a problem of gene prediction. However, even so, the Capana03g003 gene did not meet the pre-established criteria in our study to be considered a gene of the LOX family. Through comparison analysis (Blastp), Capana01g001574 and Capana01g001578 were considered the same gene, as they were encoded as CAN.G649.19 (PLAZAv.4 code) [15]. In our first analyses, the CAN.G532.31 gene was not included as a LOX gene, as it did not present the IPR001024 (PLAT/LH2) domain; however, it had the molecular weight of average LOX genes. Therefore, CAN.G532.31 (with 2463 base pairs, 820 amino acids, and a molecular weight protein of 92.01 kDa) was included in the LOX gene family for *Capsicum annuum* in order to follow the nomenclature of Sarde [10].

Although we identified 17 LOX genes for *Cucumis sativus*, the presence of 23 LOX genes for this same species has been described in a previous study [16]. Csa013924 (57.97 kDa), Csa010340 (65.49 kDa), Csa009893 (82.93 kDa), and Csa019335 (49.93 kDa) were not included as LOX genes since they lacked IPR001024 (PLAT/LH2) and because they did not have the average molecular weight of LOX proteins. The Csa022479 gene, with a molecular protein weight of 29.94 kdA, did not present the domains IPR001024 (PLAT/LH2) and IPR001246 (LOX, plant). Finally, through the comparison analysis (Blastp), Csa006735 and Csa006736 were considered to be the same gene, which, in our analyses, was encoded by Cucsa.091350 (code PLAZAv.4) [15].

We found that, for the 247 sequences used in the construction of the evolutionary tree, the number of LOX9_B genes was much smaller when compared to the numbers of LOX9_A and LOX13 genes [10,17]. The LOX9_B group was restricted to *Amborella trichopoda* (basal, one gene), *Musa acuminata* (monocot, one gene), *Setaria italica* (monocot, one gene), *Oryza sativa* ssp. *japonica* (monocot, one gene), *Sorghum bicolor* (monocot, one gene), *Brachypodium distachyon* (monocot, one gene), *Populus trichocarpa* (dicot, two genes), *Prunus persica* (eudicot, one gene), *Ricinus communis*, (eudicot, one gene), *Eucalyptus grandis* (eudicot, one

gene), *Gossypium raimondii* (eudicot, five genes), and *Citrus sinensis* (eudicot, one gene). Therefore, the LOX9_B genes were distributed among the species, mainly in only one copy, except for *Gossypium raimondii*, which presented five copies of LOX genes.

The LOX9_B subclade has already been reported in *Glycine max* and was considered to be exclusive to soybeans [18]. However, according to our results, LOX9_B had a wider distribution in the angiosperms. Using *Glycine max* LOX9_B as a query in https://shoot.bio/ (accessed on 4 April 2022), we confirmed that this clade was widespread in angiosperms, despite its patchy distribution (Figure S2). One hypothesis raised was that the LOX9_B genes may have been lost in eudicots over time. Eudicot species such as *Tarenaya hassleriana*, *Capsella rubella*, and *Arabidopsis thaliana* did not present any LOX9_B genes. We suggested that this loss in eudicots may have resulted from duplication events that occurred during the diversification of the Brassicaceae family, as in the present study we found *Tarenaya hassleriana* to be the representative species of this family. It is estimated that around 31.8 to 42.8 million years ago, close to the emergence of Brassicaceae, there was a duplication event where new classes of glucosinolates (compounds related to plant chemical defense) emerged [19]. Thus, both the duplication of glucosinolate genes and the loss of LOX_B genes in the Brassicaceae may have been favored during this evolution.

Another factor that reinforced the idea that the representatives of the LOX9_B subclade constituted a new group, when compared to other species of angiosperms, was the differential presence of conserved amino acids in a specific domain of lipoxygenase. We observed in our study that the species representing the LOX9_B group had a specific site with the conservation of the amino acids leucine (L) or methionine (M) in one of their domains. Vogt [9] identified, in this same position and in some plant species, including *Arabidopsis thaliana*, the amino acid valine (V) as conserved for the LOX9 group and the amino acid phenylalanine (F) as conserved for the LOX13 group (Figure 3). So, the LOX9_B clade is new to the literature.

Another point highlighted in our work was the way of classifying LOX proteins. In plants, most of the LOXs reported so far belong to LOX13, which plays a crucial role in the synthesis of jasmonates [1]. The LOX13 pathway

catalyzes the conversion of unsaturated fatty acids (PUFAs) such as linolenic acid and arachidonic acid to hydroperoxide octadecatrienoic acid (HPOT13), which is metabolized in the plant as the signaling compounds jasmonates and green-leaf volatile compounds (GLVs). In *Physcomitrella patens*, a moss species, we demonstrated that the LOX13 type II (LOX13) protein acted on a linolenic acid substrate, whereas another LOX (LOX9_B) protein acts on an arachidonic acid substrate [20].

Up until now, the classification of LOX proteins was based on their oxidation position or cellular location [7,10,17]. LOX9 and LOX13 have been reported to be responsible for the oxygenation of linoleic acid at carbons 9 and 13, respectively. Furthermore, LOX9 enzymes have highly similar sequences, and the sequences of LOX13 type II (LOX13) enzymes are only moderately similar and contain an N-terminal chloroplast signal peptide, whereas LOX13 type I (LOX_B) enzymes have highly similar sequences and lack a chloroplast signal peptide [13,21]. However, this form of classification (LOX9, LOX13 type I, and LOX13 type II) is not the most adequate for grouping LOX proteins, as it is known that some LOX enzymes can perform both carbon-9 and carbon-13 oxidation [22,23]. Lipoxygenase proteins can also be classified based on their cellular location—LOX type I was found in the cytoplasm and LOX type II in the organelle-targeting signal peptides [7]. However, given our results, it was identified that LOX13 proteins, both type I and type II, were cytoplasmic proteins, proteins present in chloroplasts, or proteins present in another cell compartment. Thus, subcellular localization is not the best way to classify LOX proteins.

Finally, according to the literature, all type I LOXs are also necessarily type 13. However, the type II LOX group has a mix of type 9 LOXs and type 13 LOXs, which can cause classification errors [7]. Therefore, the re-annotation of LOX genes in angiosperm families in non-model species—as was carried out in our study—was necessary to improve the phylogenetic resolution.

After an analysis of the individual selection profile of amino acids in LOX proteins, we observed that the clades LOX9_A (eudicotyledonous and monocotyledonous), LOX13 (A) (monocotyledonous), and LOX13 (B) (eudicotyledonous and monocotyledonous) showed diversifying selection, that is, a dN/dS value greater than 1, suggesting that genetic modifications in the

LOX genes for these clades were positively fixed throughout their evolution. The clades LOX9_B (eudicotyledonous and monocotyledonous) and LOX13 (B) (eudicotyledonous) presented a purifying selection, that is, a dN/dS value less than 1, suggesting a conservation of the function of the LOX genes for these clades. Thus, differences in selection pressure between the eudicotyledonous and monocotyledonous groups were observed only among the LOX13 clade (A). A ratio of dN/dS > 1 indicates acceleration, with evolution based on positive gene selection, while a ratio of dN/dS = 1 indicates that the genes are under the influence of a neutral selection action, and when the ratio of dN/dS is less than 1, the selection is indicated as purifying [24,25].

To understand the LOX gene expression profiles in angiosperms, we used public RNA-seq data from five plant species: three eudicots—*Gossypium raimondii* [26], *Prunus persica* [27], and *Ricinus communis* [28], and two monocots—*Brachypodium distachyon* [29] and *Sorghum bicolor* [30] (Figure 8). In *Brachypodium distachyon,* it was possible to notice that LOX9_A (Bradi1g11680 and Bradi1g11670) a greater expression value followed by LOX13 (Bradi3g07000 and Bradi3g07010), and LOX9_B presented the lowest expression value when comparing the LOX groups. LOX13 (Bradi3g07000 and Bradi3g07010) showed a higher expression in the control plants when compared to the plants submitted to immersion. In *Gossypium raimondii*, when studying the data obtained for leaves and roots (48, 12, and 0 h), the gene LOX13 Gorai.006G087200 had the highest expression value, and this same gene presented a differential expression between leaves (highest expression value) and roots (smallest expression value). For *Prunus persica*, when studying the leaves (control and water stress) and roots (control and water stress), the genes LOX9_B (Prupe.8G189000.1) and LOX13 (Prupe.2G005800.1 and Prupe.4g047800.1) showed the lowest values of expression. The Prupe.1G011400.1, Prupe.6g324600.2, Prupe.6G324100.1, Prupe.6G324300.1, Prupe.3G039200.1, Prupe.2G005300.1, and Prupe.1G232400.1 genes showed higher expression values in the control plants for root and water stress. Furthermore, Prupe.2G005500.1 and Prupe.6G018700.1 showed higher expression values for leaf control and water stress. These results showed that, regardless of the plant condition — control or water stress — these genes were related to the control of specific tissues.

In *Ricinus communis*, LOX13 (RCO.g.30152.000070) showed a higher expression value in terms of flower development, which was followed by seed germination. The LOX13 gene (RCO.g.29929.000202) showed a higher expression value for flower development, and the LOX13 genes RCO.g.30169.000166 and RCO.g.30169.000164 showed higher expression values for development of leaf. In *Sorghum bicolor* (when considering control and fluphenim treatment), we observed higher expression values not for the specific groups of LOX but for both conditions, i.e., control and treatment. Thus, Sobic. 003G385500.1, Sobic. 001G125900.1, Sobic. 001G125800.1, Sobic. 003G385900.1, and Sobic. 006G095600.1 had the highest expression values.

The analysis of the structure and organization of the LOX genes revealed that the number of introns and exons varied little within each identified clade. That is, the function of the LOX genes within these clades was probably the same, corroborating the groups formed in the evolutionary tree.

2.4 Materials and Methods

2.4.1 Identification and Annotation of LOX Family Genes

Genomic sequences of LOX genes were obtained in twenty-three representative angiosperm species (Table 1) with a total of thirteen dicots, six monocots, the basal angiosperm *Amborella trichopoda*, and three species as outgroups: a gymnosperm (*Picea abies*), a bryophyte (*Marchantia polymorpha*), and a green alga (*Chlamydomonas reinhardtii*).

Table 1. Species used for LOX analysis. Fourteen species of eudicots, five species of monocots, and four basal species were analyzed.

Species	LOX9-A	LOX9-B	LOX13	Total
Chlamydomonas reinhardtii	.	.	1	1
Marchantia polymorpha	.	.	16	16
Picea abies	3	.	5	8
Amborella trichopoda	.	1	3	4
Musa acuminata	9	1	7	17
Setaria italica	6	1	5	12
Sorghum bicolor	6	1	4	11
Oryza sativa	4	1	6	11
Brachypodium distachyon	4	1	5	10
Daucus carota	6	.	7	13
Coffea canephora	4	.	3	7
Capsicum annuum	7	.	3	10
Utricularia gibba	1	.	1	2
Cucumis sativus	6	.	11	17
Prunus persica	6	1	7	14
Ricinus communis	3	1	7	11
Populus trichocarpa	8	2	10	20
Eucalyptus grandis	7	1	8	16
Gossypium raimondii	6	5	6	17
Citrus sinensis	2	1	7	10
Tarenaya hassleriana	3	.	5	8
Capsella rubella	2	.	5	7
Arabidopsis thaliana	2	.	4	6

Genes were searched by BLAST using LOX proteins from *Arabidopsis thaliana* as queries in PLAZA 4.0 [15], in which sequences that obtained a score greater than 200 and an e-value less than e-50 were recovered. All genes obtained were later manually analyzed to confirm the presence of typical LOX domains. We considered as LOX genes those that simultaneously presented the following InterPro domains in their respective proteins: IPR001024 (PLAT/LH2), IPR013819 (lipoxygenase, C-terminal), IPR001246 (lipoxygenase, plant), and IPR000907 (lipoxygenase). We also obtained in PLAZA [15], using an InterPro domain search, all genes that satisfied these criteria and were not found by a BLAST search. Besides domain composition, we selected the genes whose encoded proteins had a molecular weight between 90 to 110 kilodaltons for further analysis [13]. In the cases of gene prediction errors, gene prediction was confirmed using the FGNESH tool implemented on the Softberry website (http://www.softberry.com/) (accessed on 23 February 2022).

2.4.2 Multiple Sequence Alignment and Phylogenetic Analysis

The coding sequences (CDS) in nucleotides were aligned with MUSCLE [31] and translated into an amino acid alignment in the translatorX tool (http://translatorx.co.uk/) (accessed on 25 April 2019). Amino acid alignments were used to trace the phylogenetic profile of LOX family members using the maximum likelihood method in MEGAX [25], 1000 bootstrap replicates [32], Poisson's model and uniform rates for the option 'rates among sites', and gaps in the alignment were treated as 'pairwise deletion'. After running the protein model tests implemented in MEGA, we chose LG + G + I + F [33] as the best matrix model of amino acid substitution for the phylogenetic analysis. To retrack the evolutionary relationships among the 23 plant species, an evolutionary tree was constructed using PhyloT [34] and was visualized and annotated with iTOL [35].

2.4.3 Determination of Gene Structures

Gene Structure Display Server v2.0 [36] was used with standard parameters to analyze the exon–intron structure of the LOX genes. Genomic and CDS sequences in FASTA format corresponding to the genes of all the 23 plant species were inserted to generate the gene structures.

2.4.4 Selection Pressure and Evolutionary Analysis

Non-synonymous (dN) and synonymous (dS) nucleotide substitutions of the LOX gene sequences were classified and used for the dN/dS ratio. The indices dN/dS = 1, dN/dS < 1, and dN/dS > 1 represented Darwinian neutral evolution, purifying selection, or positive selection, respectively. Individual dN/dS indices for each amino acid of the predicted proteins for each gene were determined using the statistical test suite available in MEGAX [25]. Four sets of paralogous LOX genes, LOX9_A dicots and monocots, LOX9_B dicots and monocots, and LOX13, which was subdivided into (A) dicots and (B) monocots, were analyzed to detect positive and negative selection signatures. The position of sites subjected to positive selection was predicted with FUBAR, SLAC, FEL, and MEME based on a threshold p-value < 0.05 (or a posterior probability > 0.95). All these tools were implemented using the

Datamonkey 2.0 online platform (https://datamonkey.org)(accessed on 26 November 2019) [37]. Positive and negative positions in each model were compared and grouped in Venn diagrams using the Bioinformatics & Evolutionary Genomics platform (http://bioinformatics.psb.ugent.be/webtools/Venn/) (accessed on 5 December 2019). Sites that evolved under positive selection were categorized as strongly supported (i.e., detected with at least two different methods) or moderately supported (i.e., detected with only one method). The files used for this analysis are available at https://doi.org/10.5281/zenodo.7374887.) (Accessed on 7 December 2019).

2.4.5 Analysis of LOX Gene Expression Profiles in Angiosperms

To understand the LOX gene expression profiles in the angiosperms, we used public RNA-seq data from five plant species: three eudicots, i.e., *Gossypium raimondii* [26], *Prunus persica* [27], and *Ricinus communis* [28], and two monocots, i.e., *Brachypodium distachyon* [29] and *Sorghum bicolor* [30]. Heatmaps were constructed with RPKM values obtained using CLC Genomics Workbench (CLC Bio–http://www.clcbio.com) (accessed on 3 March 2020).

2.4.6 Investigation of Motif Sequences and Cellular Localization of LOX Genes

LOX motif sequences were aligned using MAFFT version 7 [38] with the default parameters. The LOX recognition motifs were identified based on previously known domains [9]. The subcellular locations of all the LOX protein identified were also predicted. For this, two websites were used: CELLO v.2.5: subCELular localization predictor [39] and targetP-2.0.

2.5 Conclusions

In summary, we performed a comprehensive analysis of the LOX genes in 23 species of angiosperms and basal plants. We suggested that the 247 LOX members found in this study should receive a new nomenclature: LOX9_A, LOX9_B, and LOX13. The cell locations and oxidation positions of LOX9 and LOX13 should not be the most significant factors for classifying LOX genes.

The distribution of these genes in the eudicots may indicate the loss of LOX9_B genes during the diversification process of the Brassicaceae family. The rhythm for LOX gene duplication and deletion events over time was not the same between the eudicot, monocot, and basal species. The pattern of synonymous substitution in the eudicots was higher than in the monocots; however, this was not observed in the groups LOX9_B and LOX13. Finally, the LOX expression profiles showed differential expression responses in tissues such as leaves and roots and in developing endosperms and seeds as well as a differential expression of LOX genes in the species subjected to water stress.

References

1. Viswanath, K.K.; Varakumar, P.; Pamuru, R.R.; Basha, S.J.; Rao, A.D. Plant Lipoxygenases and their role in plant physiology. J. Plant Biol. 2020, 63, 83–95. [Google Scholar] [CrossRef]
2. Guo, S.; Song, Z.; Ma, R.; Yang, Y.; Yu, M. Genome-wide identification and expression analysis of the lipoxygenase gene family during peach fruit ripening under different postharvest treatments. Acta Physiol. Plant 2017, 39, 111. [Google Scholar] [CrossRef]
3. Zheng, K.; Wang, Z.; Pang, L.; Song, Z.; Zhao, H.; Wang, Y.; Wang, B.; Han, S. Systematic identification of methyl jasmonate-responsive long noncoding rnas and their nearby coding genes unveils their potential defence roles in tobacco BY-2 cells. Int. J. Mol. Sci. 2022, 23, 15568. [Google Scholar] [CrossRef]
4. Cebrián, G.; Segura, M.; Martínez, J.; Iglesias-Moya, J.; Martínez, C.; Garrido, D.; Jamilena, M. Jasmonate-deficient mutant lox3a reveals crosstalk between JA and ET in the differential regulation of male and female flower opening and early fruit development in Cucurbita pepo. J. Exp. Bot. 2022, erac468. [Google Scholar] [CrossRef] [PubMed]
5. He, S.; Zhi, F.; Min, Y.; Ma, R.; Ge, A.; Wang, S.; Wang, J.; Liu, Z.; Guo, Y.; Chen, M. The MYB59 Transcription factor negatively regulates salicylic acid- and jasmonic acid-mediated leaf senescence. Plant Physiol. 2022, kiac589. [Google Scholar] [CrossRef] [PubMed]
6. Silva, M.D.; Oliveira, M.G.A.; Lanna, A.C.; Pires, C.V.; Piovesan, N.D.; José, I.C.; Bárbara, R.; Batista, E.G.B.; Moreira, M.A. Characterization of lipoxygenase pathway of soybean plants resistant and susceptible to diaphorte phaseolorum f.sp. meridionalis, pathogen responsible for stem canker. Rev. Bras. Fisiol. Vegl. 2001, 13, 316–328. [Google Scholar] [CrossRef]
7. Shaban, M.; Ahmed, M.M.; Sun, H.; Ullah, A.; Zhu, L. Genome-wide identification of lipoxygenase gene family in cotton and functional characterization in response to abiotic stresses. BMC Genom. 2018, 19, 599. [Google Scholar] [CrossRef]

8. Umate, P. Genome-wide analysis of lipoxygenase gene family in Arabidopsis and rice. Plant Signal Behav. 2011, 6, 335–338. [Google Scholar] [CrossRef][Green Version]
9. Vogt, J.; Schiller, D.; Ulrich, D.; Schwab, W.; Dunemann, F. Identification of lipoxygenase (LOX) genes putatively involved in fruit flavour formation in apple (Malus × domestica). Tree Genet. Genomes 2013, 9, 1493–1511. [Google Scholar] [CrossRef]
10. Sarde, S.J.; Kumar, A.; Remme, R.N.; Dicke, M. Genome-wide identification, classification and expression of lipoxygenase gene family in pepper. Plant Mol. Biol. 2018, 98, 375–387. [Google Scholar] [CrossRef]
11. Feng, B.; Dong, Z.; Xu, Z.; An, X.; Quin, H.; Wu, N.; Wang, D.; Wang, T. Molecular analysis of lipoxygenase (LOX) genes in common wheat and phylogenetic investigation of LOX proteins from model and crop plants. J. Cereal Sci. 2010, 52, 387–394. [Google Scholar] [CrossRef]
12. Zhu, J.; Wang, X.; Guo, L.; Xu, Q.; Zhao, S.; Li, F.; Yan, X.; Liu, S.; Wei, C. Characterization and alternative splicing profiles of the lipoxygenase. Gene Family in Tea Plant (Camellia sinensis). Plant Cell Physiol. 2018, 59, 1765–1781. [Google Scholar] [CrossRef][Green Version]
13. Brash, A.R. Lipoxygenases: Occurrence, functions, catalysis, and acquisition of substrate. J. Biol. Chem. 1999, 274, 23679–23682. [Google Scholar] [CrossRef] [PubMed][Green Version]
14. Li, M.; Li, L.; Dunwell, J.M.; Qiao, X.; Liu, X.; Zhang, S. Characterization of the lipoxygenase (LOX) gene family in the Chinese white pear (Pyrusbretschneideri) and comparison with other members of the Rosaceae. BMC Genom. 2014, 15, 444. [Google Scholar] [CrossRef][Green Version]
15. Van Bel, M.; Diels, T.; Vancaester, E.; Kreft, L.; Botzki, A.; Van de Peer, Y.; Coppens, F.; Vandepoele, K. PLAZA 4.0: An integrative resource for functional, evolutionary and comparative plant genomics. Nucleic Acids Res 2018, 46, D1190–D1196. [Google Scholar] [CrossRef]
16. Liu, S.Q.; Liu, X.H.; Jiang, L.W. Genome-wide identification, phylogeny and expression analysis of the lipoxygenase gene family in cucumber. Genet. Mol. Res. 2011, 10, 2613–2636. [Google Scholar] [CrossRef]
17. Wang, J.; Hu, T.; Wang, W.; Hu, H.; Wei, Q.; Wei, X.; Bao, C. Bioinformatics analysis of the lipoxygenase gene family in radish (Raphanussativus) and functional characterization in response to abiotic and biotic stresses. Int. J. Mol. Sci. 2019, 20, 6095. [Google Scholar] [CrossRef][Green Version]
18. Song, H.; Wang, P.; Li, C.; Han, S.; Lopez-Baltazar, J.; Zhang, X.; Wang, X. Identification of lipoxygenase (LOX) genes from legumes and their responses in wild type and cultivated peanut upon Aspergillus flavus infection. Sci. Rep. 2016, 12, 35245. [Google Scholar] [CrossRef][Green Version]
19. Edger, P.P.; Hall, J.C.; Harkess, A.; Tang, M.; Coombs, J.; Mohammadin, S.; Schranz, M.E.; Xiong, Z.; Leebens-Mack, J.; Meyers, B.C.; et al. Brassicales phylogeny inferred from 72 plastid genes: A reanalysis of the phylogenetic localization of two paleopolyploid events and origin of novel chemical defenses. Am. J. Bot. 2018, 105, 463–469. [Google Scholar] [CrossRef] [PubMed][Green Version]

20. Anterola, A.; Göbel, C.; Hornung, E.; Sellhorn, G.; Feussner, I.; Grimes, H. Physcomitrella patens has lipoxygenases for both eicosanoid and octadecanoid pathways. Phytochemistry 2009, 70, 40–52. [Google Scholar] [CrossRef]
21. Feussner, I.; Wasternack, C. The lipoxygenase pathway. Annu. Rev. Plant Biol. 2002, 53, 275–297. [Google Scholar] [CrossRef] [PubMed]
22. Palmieri-Thiers, C.; Canaan, S.; Brunini, V.; Lorenzi, V.; Tomi, F.; Desseyn, J.L.; Garscha, U.; Oliw, E.H.; Berti, L.; Maury, J. A lipoxygenase with dual positional specificity is expressed in olives (Olea europaea L.) during ripening. Biochim. Biophys. Acta Mol. Cell Res. 2009, 1791, 339–346. [Google Scholar] [CrossRef] [PubMed][Green Version]
23. Wang, R.; Shen, W.; Liu, L.; Jiang, L.; Liu, Y.; Su, N.; Wan, J. A novel lipoxygenase gene from developing rice seeds confers dual position specificity and responds to wounding and insect attack. Plant Mol. Biol. 2008, 66, 401–414. [Google Scholar] [CrossRef]
24. Lynch, M.; Conery, J.S. The evolutionary fate and consequences of duplicate genes. Science 2000, 290, 1151–1155. [Google Scholar] [CrossRef][Green Version]
25. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol. Bio. Evo. 2018, 35, 1547–1549. [Google Scholar] [CrossRef]
26. Dong, Q.; Magwanga, R.O.; Cai, X.; Lu, P.; Nyangasi Kirungu, J.; Zhou, Z.; Wang, X.; Wang, X.; Xu, Y.; Hou, Y.; et al. RNA-Sequencing, Physiological and RNAi Analyses Provide Insights into the Response Mechanism of the ABC-Mediated Resistance to Verticilliumdahliae Infection in Cotton. Genes 2019, 10, 110. [Google Scholar] [CrossRef] [PubMed][Green Version]
27. Ksouri, N.; Jiménez, S.; Wells, C.E.; Contreras-Moreira, B.; Gogorcena, Y. Transcriptional Responses in Root and Leaf of Prunus persica under Drought Stress Using RNA Sequencing. Front. Plant Sci. 2016, 7, 1715. [Google Scholar] [CrossRef][Green Version]
28. Tan, M.; Xue, J.; Wang, L.; Huang, J.; Fu, C.; Yan, X. Transcriptomic analysis for different sex type of ricinus communis L. during development from apical buds to inflorescences by digital gene expression profiling. Front. Plant Sci. 2016, 12, 1208. [Google Scholar] [CrossRef][Green Version]
29. Rivera-Contreras, I.K.; Zamora-Hernández, T.; Huerta-Heredia, A.A.; Capataz-Tafur, J.; Barrera-Figueroa, B.E.; Juntawong, P.; Peña-Castro, J.M. Transcriptomic analysis of submergence-tolerant and sensitive Brachypodium distachyon ecotypes reveals oxidative stress as a major tolerance factor. Sci. Rep. 2016, 6, 27686. [Google Scholar] [CrossRef][Green Version]
30. Baek, Y.S.; Goodrich, L.V.; Brown, P.J.; James, B.T.; Moose, S.P.; Lambert, K.N.; Riechers, D.E. Transcriptome profiling and genome-wide association studies reveal gsts and other defense genes involved in multiple signaling pathways induced by herbicide safener in grain sorghum. Front. Plant Sci. 2019, 10, 192. [Google Scholar] [CrossRef]
31. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 2004, 32, 1792–1797. [Google Scholar] [CrossRef][Green Version]
32. Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 1985, 39, 783–791. [Google Scholar] [CrossRef]

33. Le, S.Q.; Gascuel, O. An Improved General Amino Acid Replacement Matrix. Mol. Biol. Evol. 2008, 25, 1307–1320. [Google Scholar] [CrossRef] [PubMed][Green Version]
34. Letunic, I.; Bork, P. Interactive tree of life v2: Online annotation and display of phylogenetic trees made easy. Nucleic Acids Res 2011, 39, 475–478. [Google Scholar] [CrossRef] [PubMed]
35. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v3: An online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res 2016, 44, 242–245. [Google Scholar] [CrossRef]
36. Hu, B.; Jin, J.; Guo, A.Y.; Zhang, H.; Luo, J.; Gao, G. GSDS 2.0: An upgraded gene feature visualization server. Bioinformatics 2015, 31, 1296–1297. [Google Scholar] [CrossRef] [PubMed][Green Version]
37. Weaver, S.; Shank, S.D.; Spielman, S.J.; Li, M.; Muse, S.V.; Pond, S.L.K. Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes. Mol. Biol. Evol. 2018, 35, 773–777. [Google Scholar] [CrossRef] [PubMed][Green Version]
38. Katoh, K.; Rozewicki, J.; Yamada, K.D. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform. 2019, 20, 1160–1166. [Google Scholar] [CrossRef][Green Version]
39. Yu, C.; Chen, Y.; Lu, C.; Hwang, J. Prediction of protein subcellular localization. Proteins: Struct. Funct. Bioinform. 2006, 64, 643–651. [Google Scholar] [CrossRef]
40. Armenteros, J.J.A.; Salvatore, M.; Emanuelsson, O.; Winther, O.; Heijne von, G.; Elofsson, A.; Henrik, N. Detecting sequence signals in targeting peptides using deep learning. Life Sci. Alliance 2019, 2, e201900429. [Google Scholar] [CrossRef]

3. CAPÍTULO 2 - Genome-wide identification and characterization of the lipoxygenase gene family in the tetraploid *Coffea arabica* L. and its diploid parental genomes

Resumo

As enzimas lipoxigenases (LOXs) têm um papel crucial no crescimento, desenvolvimento e defesa das plantas, e estão envolvidas na produção do hormônio ácido jasmônico (JA). Estudos indicam que o ácido hexanoico (C6H12O2) é um elicitor pode estimular a defesa das plantas por meio da biossíntese de jasmonatos. Até agora, não há informações detalhadas sobre a expressão diferencial de lipoxigenases em plantas de cafeeiro sob a ação de elicitores.

Nosso estudo buscou identificar genes codificantes de enzimas lipoxigenases em três espécies de cafeeiro: *Coffea arabica*, *Coffea canephora* e *Coffea eugenioides*, bem como avaliar se genes de lipoxigenases são expressos diferencialmente em função da aplicação de ácido hexanoico em *C. arabica*.

Encontramos 18 genes de lipoxigenases em *Coffea arabica* e 9 genes em *Coffea eugenioides* e *Coffea canephora*. A análise dos posicionamentos cromossômicos dos genes mostrou uma alta correspondência entre os genes do tetraploide *Coffea arabica* e os genomas de seus possíveis parentais, *Coffea eugenioides* e *Coffea canephora*. Observamos ainda que a aplicação de ácido hexanoico pode alterar a expressão de alguns genes de lipoxigenases em folhas e raízes de C. arabica cv. Catuaí Vermelho e C. arabica cv. Obatã. A aplicação de ácido hexaoico também altera a atividade da enzima lipoxigenase e três genes apresentaram alta correlação entre a atividade da enzima lipoxigenase e a modulação da atividade transcricional de genes LOX.

Com base nos dados, concluímos que alguns genes de lipoxigenases em *Coffea arabica* podem ser candidatos para análises mais específicas em estudos relacionados ao gênero *Coffea*, por serem possíveis contribuintes majoritários para a atividade da enzima lipoxigenase.

Palavras-chave: Ácido hexanoico; elicitação de cafeeiro; Expressão gênica diferencial; Lipoxigenases.

Abstract

Lipoxygenase enzymes (LOXs) play a crucial role in plant growth, development, and defense and are involved in the production of the hormone jasmonic acid (JA). Studies indicate that hexanoic acid (C6H12O2) is an elicitor that can stimulate plant defense through the biosynthesis of jasmonates. So far, there is no detailed information on the differential expression of lipoxygenases in coffee plants under the action of elicitors. Our study aimed to identify genes encoding lypoxigenase in three coffee species: *Coffea arabica*, *Coffea canephora*, and *Coffea eugenioides*, as well as to evaluate whether lipoxygenase genes are differentially expressed as a result of hexanoic acid application in *C. arabica*. We found 18 lipoxygenase genes in Coffea arabica and 9 genes in *Coffea eugenioides* and *Coffea canephora*. The analysis of chromosomal positions of the genes showed a high correspondence between the genes of the tetraploid *Coffea arabica* and the genomes of its possible parental species, *Coffea eugenioides* and *Coffea canephora*. We also observed that the application of hexanoic acid can alter the expression of some lipoxygenase genes in leaves and roots of *C. arabica* cv. Catuaí Vermelho and C. arabica cv. Obatã. Hexanoic acid application also alters the activity of the lipoxygenase enzyme, and three genes show a high correlation between lipoxygenase enzyme activity and the modulation of LOX gene transcriptional activity. Based on the data, we conclude that some lipoxygenase genes in Coffea arabica may be candidates for more specific analyses in studies related to the *Coffea* genus, as they are possible major contributors to lipoxygenase enzyme activity.

Keywords: Hexanoic acid; Coffee plant elicitation; Differential gene expression; Lipoxygenases.

3.1 Introdução

O café é popular em todo mundo, sendo o Brasil o maior produtor dessa cultura. A produção de *Coffea arabica* corresponde a 64% e a produção de *Coffea canephora* (conhecida como conilon), corresponde a 36% da safra total dos cafés do Brasil (CONAB, 2022). *C. arabica* é uma espécie alotetraploide (2n = 4x = 44), derivado de uma fusão genômica entre as espécies *C. canephora* (Robusta ou Conilon) e *C. eugenioides* (espécie selvagem) (SATTLER et al., 2022).

Diferentes pragas afetam a produção e a qualidade do café e ácido jasmônico (JA) é um hormônio que desempenha um papel importante no crescimento, desenvolvimento e mecanismos de defesa, principalmente contra insetos-praga. Por exemplo, a presença do bicho-mineiro aumenta a concentração de lipoxigense em folhas de cafeeiro (MERIÑO-CABRERA et al., 2018). As principais enzimas envolvidas na produção de JA são as lipoxigenases (LOX), essas enzimas, oxidorredutases, estão distribuídas naturalmente nos vegetais. A oxidação por LOX resulta na produção de aroma e também participa na regulação da biossíntese de compostos voláteis. As LOX podem atuar como um agente aromatizante natural para a produção de alimentos (SINGH et al., 2022). A “via das lipoxigenases” é iniciada pela lipoxigenase, 9-LOX ou 13-LOX. As enzimas LOX nas plantas são proteínas estáveis, muitas vezes presentes em altas quantidades, especialmente nas folhas (CHRISTIE; HARWOOD, 2020).

Há relatos que plantas submetidas ao ácido hexanoico (C6H12O2) podem apresentar resposta de defesa, através da via de biossíntese de jasmonatos (SCALSCHI et al., 2013; FINITI et al., 2014). O ácido hexanoico é um composto de seis carbonos, derivado do hexano. É sabido que esse composto pode gerar respostas de defesa em alguns sistemas vegetais, nesses casos, o ácido hexanoico pode apresentar um efeito bactericida e fungicida (CAMAÑES et al., 2015; LÓPEZ-GALIANO et al, 2019). ARANEGA-BOU et al. (2014).

Alguns perfis de expressão de lipoxigenases em plantas submetidas às diferentes condições já foram relatados. Em *Brachypodium distachyon*, as lipoxigenases Bradi1g11680 e Bradi1g11670 tem sua expressão reprimida em

plantas submetidas ao estresse de submersão (CAMARGO et al., 2022). Em *Prunus persica* submetidas ao estresse hídrico, foi possível observar que a expressão diferencial de lipoxigenases estava relacionada ao controle de tecidos específicos, como folha e raiz (KSOURI et al., 2016).

Até o presente há pouca informação sobre a caracterização de lipoxigenases em cafeeiro. PATUI et al. (2007) avaliaram a atividade de lipoxigenases em frutos de cafeeiro arábica, encontrando maior atividade na polpa e MERIÑO-CABRERA et al. (2018) observaram a modulação da atividade enzimática de lipoxigenase em função do ataque de bicho-mineiro em folhas de cafeeiro. No entanto, não há informações sobre análises especificas de expressão diferencial da família gênica de lipoxigenases em plantas de cafeeiro submetidas a elicitores. Diante do exposto, nosso trabalho teve como objetivo a identificação de enzimas lipoxigenases em três espécies de Café: *Coffea arabica*, *Coffea canephora* e *Coffea eugenioides*, assim como, a identificação de genes de lipoxigenases diferencialmente expressos pela aplicação de ácido hexanoico em *C. arabica,* para as cultivares Catuaí e Obatã.

As cultivares foram selecionadas com base em suas histórias distintas de reprodução e resistência. A cultivar Catuaí é uma das cultivares mais plantadas no Brasil, porém é suscetível à ferrugem, a principal doença do café arábica em todo mundo (TALHINHAS et al., 2017). Já a cultivar Obatã apresenta-se como moderadamente resistente à ferrugem (DEL GROSSI et al., 2013).

3.2 Resultados

3.2.1 Identificação e análise filogenética de genes LOX em Coffea

Foram identificados em *Coffea arabica* (Ca) 18 genes de lipoxigenases. Em *Coffea eugenioides* (Ce) e *Coffea canephora* (Cc) foram identificados 9 genes LOX (Tabela 1). Esse resultado está de acordo com o observado em *Coffea arabica*, espécie alotetraploide (2n = 4x = 44), derivada de uma fusão genômica entre as espécies *C. canephora* e *C. eugenioides*. Os números de pares de bases (pb) dos genes LOX variou entre 2820 pb a 2091 pb, já o

tamanho médio dos aminoácidos ficou entre 939 pb e 696 aa. Por fim, o peso molecular, medido em Kilodalton (kD), das proteínas LOX variou entre 76.78 kD e 106.53 kD. Estes dados estão detalhados na Tabela 1.

Tabela 1. Genes de Lipoxigenases identificados em *Coffea arabica, Coffea eugenioides* e *Coffea canephora*.

Espécies	ID gene	ID mRNA	Nome gene	Lozalização cromossômica	CDS/pb	Tamanho/aa	Peso Molecular/kD
Coffea arabica	LOC113727628	XM_027251892.1	CaLOX1	2c: NC_039900.1 (59528810..59533257)	2571	856	88.91
	LOC113732840	XM_027258829.1	CaLOX2	2e: NC_039901.1 (65479485..65484033)	2565	854	88.69
	LOC113732839	XM_027258827.1	CaLOX3	2e: NC_039901.1 (65426956..65433029)	2586	861	90.54
	LOC113727627	XM_027251891.1	CaLOX4	2c: NC_039900.1 (59504040..59510098)	2586	861	90.5
	LOC113727625	XM_027251889.1	CaLOX5	2c: NC_039900.1 (59477945..59487346)	2580	859	90.3
	LOC113732838	XM_027258826.1	CaLOX6	2e: NC_039901.1 (65405503..65414130)	2580	859	90.08
	LOC113738240	XM_027265485.1	CaLOX7	3e: NC_039903.1 (2929755..2934192)	2649	882	91.98
	LOC113734115	XM_027260465.1	CaLOX8	3c: NC_039902.1 (2801597..2806034)	2649	882	91.98
	LOC113727626	XM_027251890.1	CaLOX9	2c: NC_039900.1 (59488470..59496591)	2388	795	83.02
	LOC113741416	XM_027268943.1	CaLOX10	4e: NC_039905.1 (719683..724527, complement)	2757	918	95.25
	LOC113741614	XM_027269183.1	CaLOX11	4e: NC_039905.1 (748134..752980, complement)	2757	918	95.25
	LOC113741495	XM_027269026.1	CaLOX12	4e: NC_039905.1 (1537669..1542545)	2757	918	95.32
	LOC113714586	XM_027238534.1	CaLOX13	1c: NC_039898.1 (12190762..12202237, complement)	2700	899	93.45
	LOC113716533	XM_027240915.1	CaLOX14	11c: NC_039918.1 (35207313..35211921, complement)	2781	926	96.81
	LOC113700875	XM_027221306.1	CaLOX15	1e: NC_039899.1 (10792524..10804053, complement)	2700	899	93.67
	LOC113730976	XM_027255986.1	CaLOX16	2e: NC_039901.1 (11190355..11196852)	2781	926	89.65
	LOC113726022	XM_027249481.1	CaLOX17	2c: NC_039900.1 (11420434..11427279)	2601	866	90.15
	LOC113719047	XM_027244036.1	CaLOX18.1	11e: NC_039919.1 (41296241..41301156, complement)	2601	866	96.77
	LOC113719047	XM_027244035.1	CaLOX18.2	11e: NC_039919.1 (41296241..41301156, complement)	2514	837	86.69
Coffea eugenioides	LOC113763778	XM_027307697.1	CeLOX1	2: NC_040036.1 (73407070..73411494)	2565	854	88.69
	LOC113761107	XM_027303951.1	CeLOX2	2: NC_040036.1 (73360479..73366982)	2586	861	90.5
	LOC113760605	XM_027303251.1	CeLOX3	2 NC_040036.1 (73343060..73351673)	2580	859	90.25
	LOC113765110	XM_027309170.1	CeLOX4	3: NC_040037.1 (3272051..3276705)	2649	882	92.13
	LOC113753037	XM_027297112.1	CeLOX5	11: NC_040045.1 (40363926..40368580)	2649	882	92.13
	LOC113767484	XM_027311595.1	CeLOX6	4: NC_040038.1 (1040424..1045395)	2757	918	95.35
	LOC113759572	XM_027302152.1	CeLOX7	1: NC_040035.1 (11817844..11829437, complement)	2700	899	93.58
	LOC113751729	XM_027295841.1	CeLOX8	11: NC_040045.1 (46388687..46393584, complement)	2781	926	96.77
	LOC113762616	XM_027306149.1	CeLOX9	2: NC_040036.1 (13029466..13036021)	2601	866	89.73
Coffea canephora	GSCOCT00013092001	Cc02_g33790.1	CcLOX1	chr2:46692788..46697429	2091	696	79.61
	GSCOCG00029162001	Cc02t13400.1	CcLOX2	2:11512899..11519353	2817	938	106.53
	GSCOCG00038212001	Cc11t16680.1	CcLOX3	11:32734546..32738830	2820	939	95.15
	GSCOCG00004674001	Cc00_g30760	CcLOX4	chr0:185343573..185347521	2397	798	89.57
	GSCOCG00026050001	Cc03t03580.1	CcLOX5	3:2736559..2740860	2649	798	88.97
	GSCOCG00015359001	Cc02t33320.1	CcLOX6	2:45665773..45670461	2571	856	85.56
	GSCOCG00013091001	Cc02t33780.1	CcLOX7	2:46680121..46686068	2574	857	86.63
	GSCOCG00039270001	Cc01t04060.1	CcLOX8	1:10752168..10764896	2514	837	83.4
	GSCOCG00013093001	Cc02t33800.1	CcLOX9	2:46698748..46702482	2400	799	76.78

Legenda: Foram identificados 18 genes em *Coffea arabica*, 9 genes em *Coffea eugenioides* e 9 genes em *Coffea canephora*, com seus respectivos posicionamentos no genoma, tamanho da sequência de nucleotídeos (CDS/pb), número de aminoácidos presentes na sequência proteica e tamanho da proteína LOX, em Kilodaltons (kD).

Foi desenhada uma árvore filogenética para analisar a relação dos genes LOX em *Coffea arabica*, *Coffea eugenioides* e *Coffea canephora*,

utilizando o alinhamento das sequências de aminoácidos. Para aprimorar a compreensão das relações, os genes LOX das eudicotiledôneas *Arabidopsis thaliana* e *Daucus carota* também foram incluídos na construção da árvore. Os genes LOX foram divididos em 2 grupos, LOX9 e LOX13, nos quais a maioria dos clados apresentaram bootstrap acima de 90%. Com base nas informações filogenéticas, 10 genes de *Coffea arabica*, 4 genes de *Coffea eugenioides* e 4 genes de *Coffea canephora* puderam ser classificados como LOX13. Também foram inclusos como genes do grupo LOX9, 9 genes para *Coffea arabica* (incluindo os genes CaLOX18.1 e CaLO18.2 – provenientes, possivelmente, de processos de *splicing* alternativo), 5 genes para *Coffea eugenioides* e 5 genes para *Coffea canephora* (Figura 2 e Apêndice III).

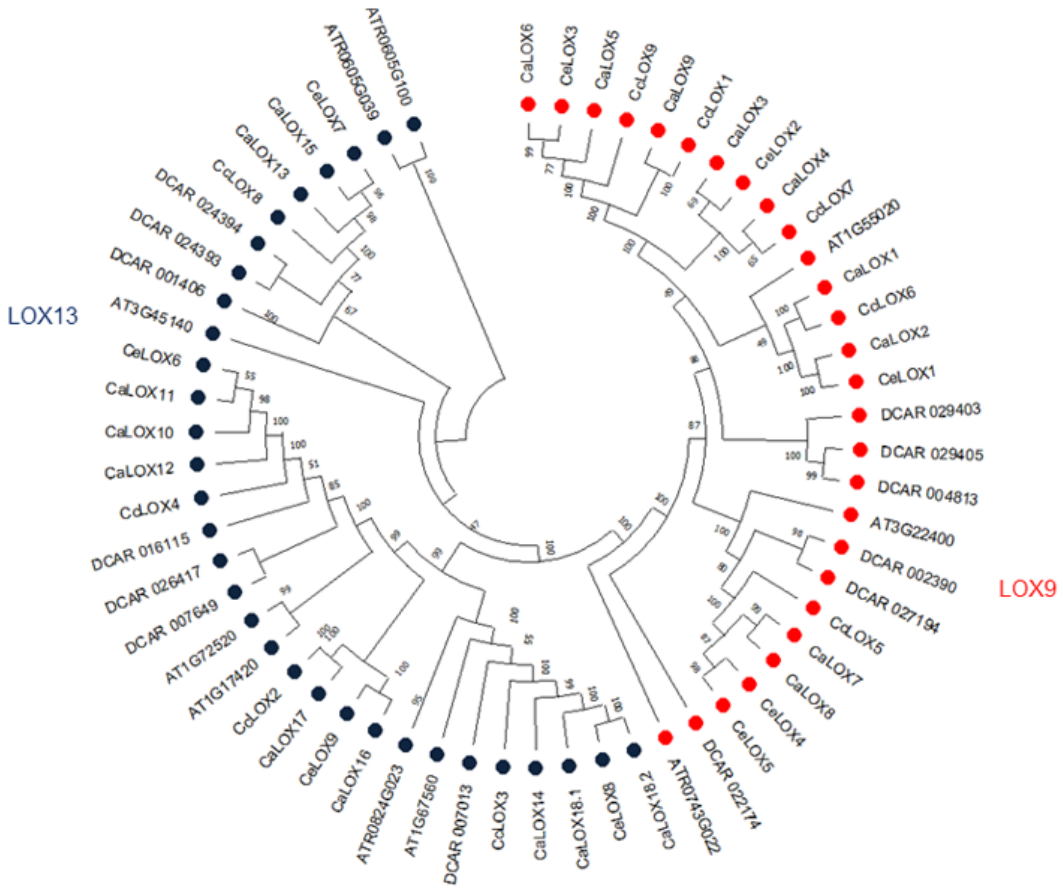


Figura 2. Árvore filogenética (método de máxima verossimilhança, bootstrap de 1000 réplicas) de lipoxigenases. Legenda: Lipoxigenasas de *Coffea arabica* (CaLOX), *Coffea eugenioides* (CeLOX), *Coffea canephora* (CcLOX), *Arabidopsis thaliana* (AT) e *Daucus carota* (DCAR). O grupo LOX9 está representado pela cor vermelha, o grupo LOX13 está representado pela cor azul.

3.2.2 Mapa cromossômico dos genes de *Coffea arabica, Coffea canephora e Coffea eugenioides*

As localizações cromossômicas dos genes de lipoxigenases foram identificadas em *Coffea arabica* (Ca, *Coffea eugenioides* (Ce) e *Coffea canephora* (Cc) (Figura 3). Foi observado que houve uma alta correspondência no posicionamento dos genes ao comparar o genoma tetraploide de *Coffea arabica* com seus respectivos genomas parentais, *Coffea eugenioides* e *Coffea canephora*. As únicas diferenças de posicionamento foram identificadas no cromossomo Ca4e de *Coffea arabica*, que possuía três genes LOX, enquanto o cromossomo Ce4 de *Coffea eugenioides* apresentava apenas um gene LOX. Além disso, no cromossomo Ca11e de *Coffea arabica* foi identificado um gene LOX, enquanto o cromossomo Ce11 de *Coffea eugenioides* apresentava dois genes LOX.

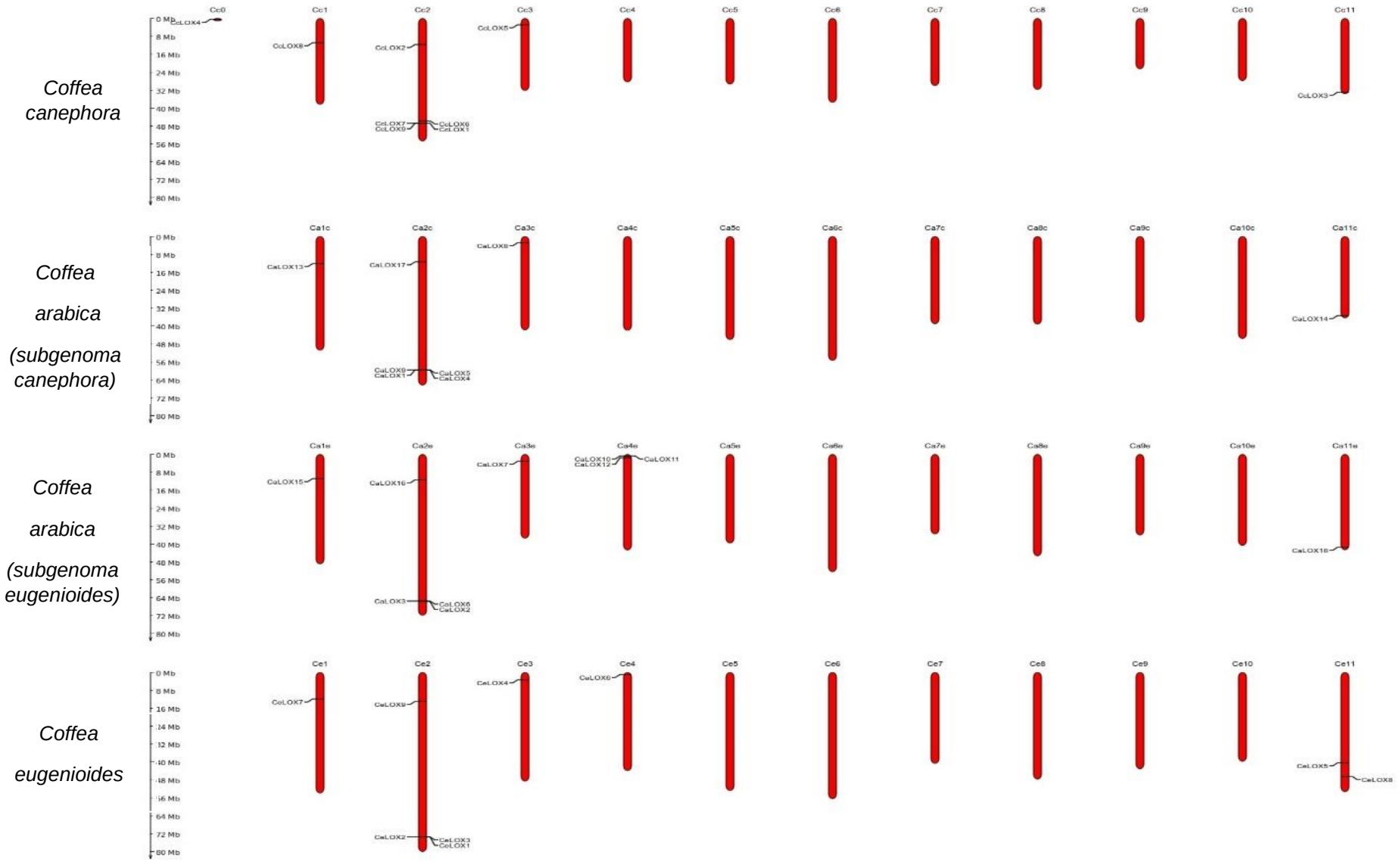

Coffea canephora
Cc0
CcLOX4
Cc1
CcLOX8
Cc2
CcLOX2
CcLOX7
CcLOX9
CcLOX6
CcLOX1
Cc3
CcLOX5
Cc4
Cc5
Cc6
Cc7
Cc8
Cc9
Cc10
Cc11
CcLOX3
Coffea arabica (subgenoma canephora)
Ca1c
CaLOX13
Ca2c
CaLOX17
CaLOX9
CaLOX1
CaLOX5
CaLOX4
Ca3c
CaLOX8
Ca4c
Ca5c
Ca6c
Ca7c
Ca8c
Ca9c
Ca10c
Ca11c
CaLOX14
Coffea arabica (subgenoma eugenioides)
Ca1e
CaLOX15
Ca2e
CaLOX16
CaLOX3
CaLOX6
CaLOX2
Ca3e
CaLOX7
Ca4e
CaLOX10
CaLOX12
CaLOX11
Ca5e
Ca6e
Ca7e
Ca8e
Ca9e
Ca10e
Ca11e
CaLOX18
Coffea eugenioides
Ce1
CeLOX7
Ce2
CeLOX9
CeLOX2
CeLOX3
CeLOX1
Ce3
CeLOX4
Ce4
CeLOX6
Ce5
Ce6
Ce7
Ce8
Ce9
Ce10
Ce11
CeLOX5
CeLOX8

Figura 3. Localização cromossômica de genes LOX em Coffea. *Coffea arabica* (Ca), *Coffea eugenioides* (Ce) e *Coffea canephora* (Cc).

3.2.3 Estrutura de éxons e íntrons dos genes LOX em Coffea

Foram identificados os padrões de estrutura éxon-íntron dos genes LOX em *Coffea* (Figura 4). Em *Coffea arabica* o número de éxons foi de 8 a 9 e o número de íntrons variou de 7 a 8. Em *Coffea eugenioides* o menor número de éxons foi de 7 e o maior número foi de 11, seguidos por 6 íntrons e 10 íntrons, respectivamente. Em *Coffea canephora*, foi identificada a menor variação na estrutura de éxons e íntrons, sendo que a maioria dos genes apresentou 9 éxons e 8 íntrons, exceto pelo gene CcLOX6, que apresentou 8 éxons e 7 íntrons (Figura 4). Também foram preditas as localizações subcelulares das proteínas LOX em *Coffea arabica, Coffea canephora* e *Coffea eugenioides* (Tabela 2).

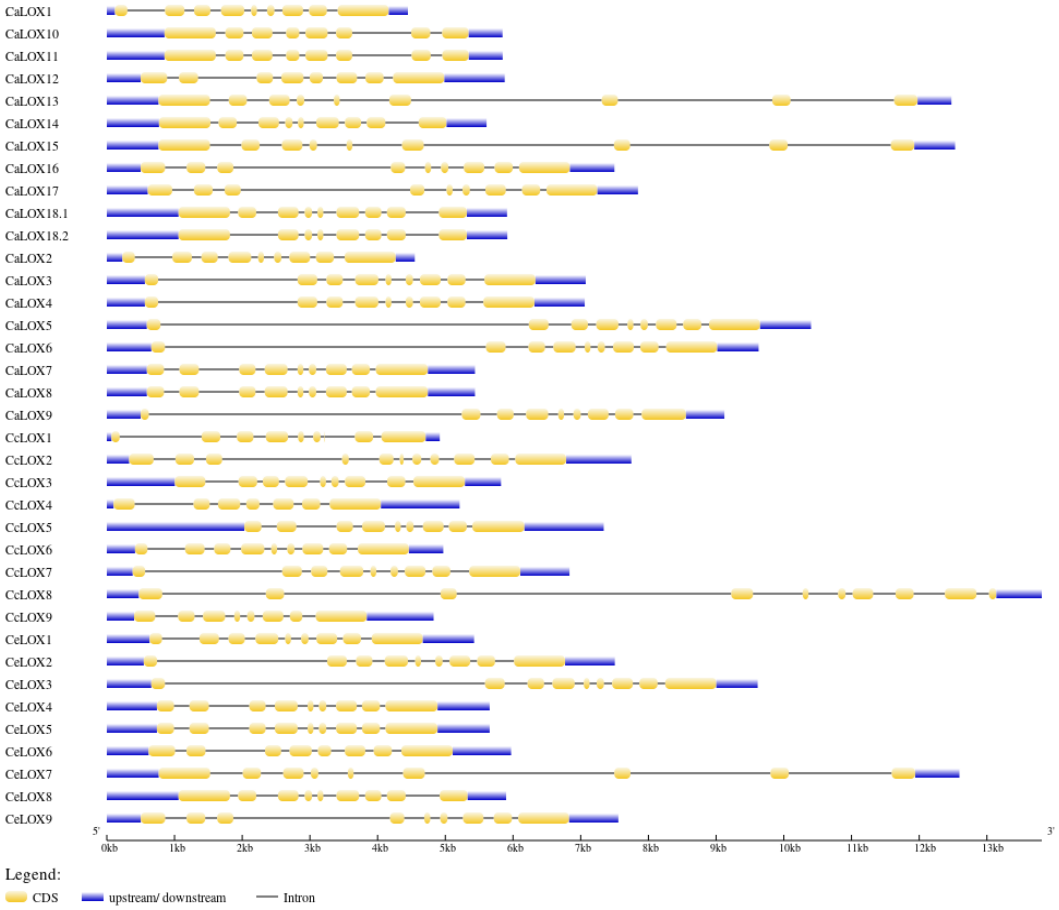


Figura 4. Estruturas de éxons e íntrons em *Coffea arabica* (CaLOX), *Coffea eugenioides* (CeLOX) e *Coffea canephora* (CcLOX). Os éxons são representados pelas barras amarelas e os íntrons são representados pelas linhas cinzas.

Tabela 2. Localização subcelular em *Coffea arabica, Coffea eugenioides* e *Coffea canephora*.

Espécies	Gene ID	mRNA ID	Nome do Gene	LocalizaçãoSubcelular	Plant-mSubP Dipep	DeepLoc - 2.0	Euk-mPLoc 2.0
Coffea arabica	LOC113727628	XM_027251892.1	CaLOX1	Outro	Citoplasma	Citoplasma	Citoplasma
	LOC113732840	XM_027258829.1	CaLOX2	Outro	Citoplasma	Citoplasma	Citoplasma
	LOC113732839	XM_027258827.1	CaLOX3	Outro	Citoplasma	Citoplasma	Citoplasma-Cloroplasto
	LOC113727627	XM_027251891.1	CaLOX4	Outro	Citoplasma	Citoplasma	Cloroplasto
	LOC113727625	XM_027251889.1	CaLOX5	Outro	Citoplasma	Citoplasma	Citoplasma
	LOC113732838	XM_027258826.1	CaLOX6	Outro	Citoplasma	Citoplasma	Citoplasma
	LOC113738240	XM_027265485.1	CaLOX7	Outro	Citoplasma	Citoplasma	Citoplasma-Cloroplasto
	LOC113734115	XM_027260465.1	CaLOX8	Outro	Citoplasma	Citoplasma	Citoplasma-Cloroplasto
	LOC113727626	XM_027251890.1	CaLOX9	Outro	Citoplasma	Citoplasma	Cloroplasto
	LOC113741416	XM_027268943.1	CaLOX10	Outro	Cloroplasto	Cloroplasto	Cloroplasto
	LOC113741614	XM_027269183.1	CaLOX11	Outro	Cloroplasto	Cloroplasto	Cloroplasto
	LOC113741495	XM_027269026.1	CaLOX12	Outro	Citoplasma	Cloroplasto	Cloroplasto
	LOC113714586	XM_027238534.1	CaLOX13	Peptídeo de transferência - cloroplasto	Citoplasma	Cloroplasto	Cloroplasto
	LOC113716533	XM_027240915.1	CaLOX14	Peptídeo de transferência - cloroplasto	Cloroplasto	Cloroplasto	Cloroplasto
	LOC113700875	XM_027221306.1	CaLOX15	Peptídeo de transferência - cloroplasto	Citoplasma	Cloroplasto	Cloroplasto
	LOC113730976	XM_027255986.1	CaLOX16	Peptídeo de transferência - mitocôndria	Cloroplasto	Cloroplasto	Cloroplasto
	LOC113726022	XM_027249481.1	CaLOX17	Outro	Cloroplasto	Cloroplasto	Cloroplasto
	LOC113719047	XM_027244036.1	CaLOX18.1	Peptídeo de transferência - cloroplasto	Cloroplasto	Cloroplasto	Cloroplasto
	LOC113719047	XM_027244035.1	CaLOX18.2	Peptídeo de transferência - cloroplasto	Cloroplasto	Cloroplasto	Cloroplasto
Coffea eugenioides	LOC113763778	XM_027307697.1	CeLOX1	Outro	Citoplasma	Citoplasma	Citoplasma
	LOC113761107	XM_027303951.1	CeLOX2	Outro	Citoplasma	Citoplasma	Cloroplasto
	LOC113760605	XM_027303251.1	CeLOX3	Outro	Citoplasma	Citoplasma	Citoplasma
	LOC113765110	XM_027309170.1	CeLOX4	Outro	Citoplasma	Citoplasma	Citoplasma-Cloroplasto
	LOC113753037	XM_027297112.1	CeLOX5	Outro	Citoplasma	Citoplasma	Citoplasma-Cloroplasto
	LOC113767484	XM_027311595.1	CeLOX6	Outro	Cloroplasto	Cloroplasto	Cloroplasto
	LOC113759572	XM_027302152.1	CeLOX7	Peptídeo de transferência - cloroplasto	Citoplasma	Cloroplasto	Cloroplasto
	LOC113751729	XM_027295841.1	CeLOX8	Peptídeo de transferência - cloroplasto	Cloroplasto	Cloroplasto	Cloroplasto
	LOC113762616	XM_027306149.1	CeLOX9	Peptídeo de transferência - cloroplasto	Cloroplasto	Cloroplasto	Cloroplasto
Coffea canephora	GSCOCT00013092001	Cc02_g33790.1	CcLOX1	Outro	Citoplasma	Citoplasma	Cloroplasto
	GSCOCG00029162001	Cc02t13400.1	CcLOX2	Peptídeo de transferência - mitocôndria	Cloroplasto	Cloroplasto	Cloroplasto
	GSCOCG00038212001	Cc11t16680.1	CcLOX3	Outro	Cloroplasto	Cloroplasto	Cloroplasto
	GSCOCG00004674001	Cc00_g30760	CcLOX4	Outro	Citoplasma	Citoplasma	Cloroplasto
	GSCOCG00026050001	Cc03t03580.1	CcLOX5	Outro	Cloroplasto	Citoplasma	Citoplasma-Cloroplasto
	GSCOCG00015359001	Cc02t33320.1	CcLOX6	Outro	Citoplasma	Citoplasma	Citoplasma
	GSCOCG00013091001	Cc02t33780.1	CcLOX7	Outro	Citoplasma	Citoplasma	Citoplasma
	GSCOCG00039270001	Cc01t04060.1	CcLOX8	Peptídeo de transferência - cloroplasto	Citoplasma	Cloroplasto	Cloroplasto
	GSCOCG00013093001	Cc02t33800.1	CcLOX9	Outro	Citoplasma	Citoplasma	Citoplasma

Legenda: *Coffea arabica*, representado por: CaLOX. *Coffea eugenioides*, representado por: CeLOX. E *Coffea canephora,* representado por: CcLOX. Sites utilizados: **Plant-mSubP** – com modelo de predição Dipep, **DeepLoc 2.0** – com modelo *High-quality* e formato *Long output* e **Euk-mPLoc 2.0.**

3.2.4 Atividade enzimática de lipoxigenases em *C. arabica* cv Catuaí Vermelho e Obatã

As folhas e raízes de Catuaí Vermelho e Obatã que foram tratadas com ácido hexanoico (HX) apresentaram valores mais elevados de atividade

enzimática de lipoxigenases em comparação com seus controles. Em Catuaí, os valores de folha com HX foram maiores quando comparados aos valores de folha controle, raiz controle e raiz com HX ($p<0,001$). Os valores de folha controle também foram maiores do que os valores de raiz controle ($p<0,01$). Assim como, os valores de raiz com HX, que foram maiores que os de raiz controle ($p<0,05$). Para Obatã, os valores de raiz controle foram significamente maiores que os valores de folha controle ($p<0,05$). Por fim, os valores de raiz com HX foram maiores quando comparados aos valores de raiz controle, folha com HX e folha controle ($p<0,01$).

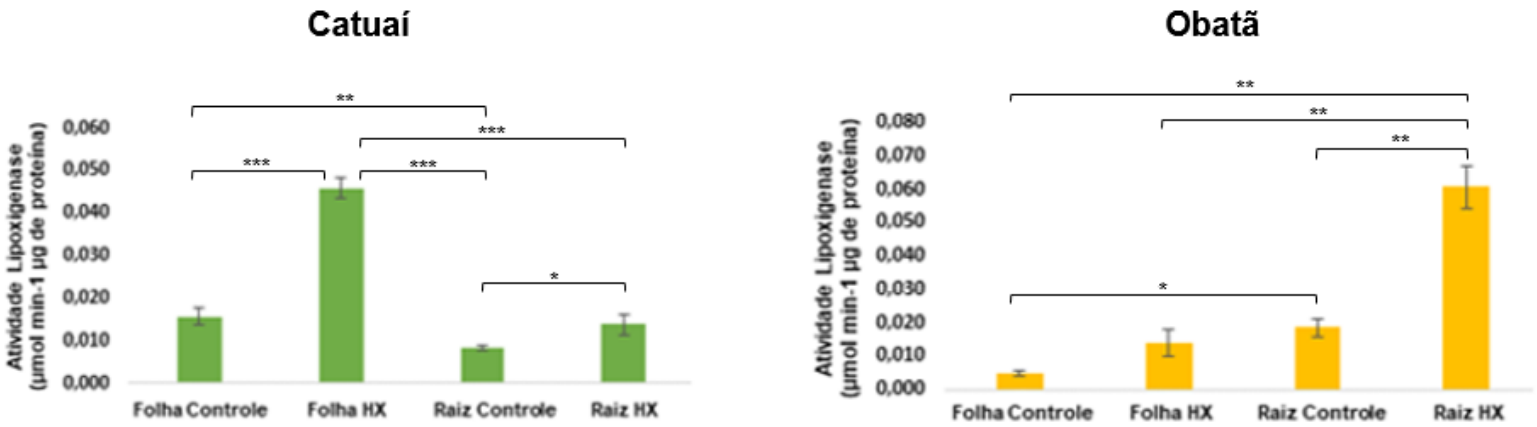


Figura 5. Atividade enzimática de Lipoxigenases em µmol min^{-1}ug. Catuaí folha (controle - tratamento com HX) e raiz (controle - tratamento com HX). Obatã folha (controle - tratamento com HX) e raiz (controle - tratamento com HX). Diferenças estatísticas: * $p<0,05$; ** $p<0,01$; *** $p<0,001$.

3.2.5 Perfis de expressão de genes LOX baseados em RNA-seq

Para entender os perfis de expressão dos genes LOX em *C. arabica* cv. Catuaí Vermelho (CV) e C. arabica cv. Obatã (OB), foram identificados os valores em RPKM (Reads Per Kilobase Million) em folhas e raízes submetidas ao tratamento com HX. Os maiores valores de expressão foram encontrados nos seguintes genes: LOC113727628 (CaLOX1) em raízes de plantas controle tanto em CV como em OB, LOC113732840 (CaLOX2) em raízes controle e com HX, nas duas cultivares (CV e OB), LOC113732839 (CaLOX3), LOC113727627 (CaLOX4) e LOC113741495 (CaLOX12) apresentaram maiores valores de expressão em raiz com HX - Obatã. LOC113714586 (CaLOX13) em Catuaí folha

(controle e com HX) e Obatã folhas e raízes com HX, LOC113700875 (CaLOX15) em folhas Catuaí controle e com HX e em Obatã folha controle e raiz com HX, LOC113730976 (CaLOX16) em raiz Catuaí com HX, LOC113726022 (CaLOX17) em Obatã raiz com HX e LOC113719047 (CaLOX18.1 e CaLOX18.2) para ambas as cultivares em todas as condições (Tabela 3).

Também identificamos os valores de RPKM em furtos verdes, amarelos e vermelhos de *Coffea arabica* em dados públicos de RNA-seq (Cheng et al., 2018). Em frutos verdes e amarelos os maiores valores de expressão foram identidicados em LOC113732840 (CaLOX2), LOC113732838 CaLOX6, e LOC113700875 (CaLOX15). Em frutos vermelhos o maior valor de expressão foi identificado em apenas LOC113732840 (CaLOX2) (Tabela 4).

Tabela 3. Valores de RPKM (*Reads Per Kilobase Million*) dos genes de lipoxigenases em *C. arabica* cv. Catuaí vermelho e *C. arabica* cv. Obatã.

Gene ID	Nome do Gene	mRNA ID	RPKM.CC_folha	RPKM.CHX_folha	RPKM.CC_raiz	RPKM.CHX_raiz	RPKM.OC_folha	RPKM.OHX_folha	RPKM.OC_raiz	RPKM.OHX_raiz
LOC113727628	CaLOX1	XM_027251892.1	3,57 ± 0,65	2,3 ± 0,89	243,43 ± 42,22	56,52 ± 0,25	1,35 ± 0,48	1,8 ± 0,51	138,83 ± 20,58	94,45 ± 14,66
LOC113732840	CaLOX2	XM_027258829.1	29,7 ± 5,11	26,9 ± 0,85	521,43 ± 81,45	64,06 ± 0,48	28,69 ± 12,17	52,02 ± 3,20	472,85 ± 32,05	115,87 ± 38,86
LOC113732839	CaLOX3	XM_027258827.1	9,44 ± 2,05	9,23 ± 1,13	30,66 ± 0,66	6,68 ±1,57	12,74 ± 2,85	9,5 ± 0,98	42,85 ± 6,90	110,43 ± 13,28
LOC113727627	CaLOX4	XM_027251891.1	3,31 ± 0,30	6,23 ± 1,58	41,77 ± 4,01	11,7 ± 1,83	1,4 ± 1,07	2,72 ± 0,80	28,26 ± 1,19	130,56 ± 38,27
LOC113727625	CaLOX5	XM_027251889.1	5,91 ± 1,92	7,15 ± 0,19	2,58 ± 0,38	4,4 ± 4,75	24,16 ± 5,12	15,78 ± 4,04	1,62 ± 0,81	5,18 ± 1,35
LOC113732838	CaLOX6	XM_027258826.1	3,05 ± 0,61	3 ± 0,63	1,28 ± 0,01	3,14 ± 1,76	12,69 ± 4,05	8,82 ± 0,66	2,05 ± 0,03	1,67 ± 0,89
LOC113738240	CaLOX7	XM_027265485.1	16,39 ± 5,85	11,22 ± 2,09	116,75 ± 0,33	18,72 ± 3,50	7,7 ± 0,99	12,9 ± 1,92	44,46 ± 7,35	28,48 ± 13,25
LOC113734115	CaLOX8	XM_027260465.1	5,4 ± 0,60	0 ± 0	0 ± 0	0 ± 0	2 ± 0,34	0 ± 0	0 ± 0	0 ± 0
LOC113727626	CaLOX9	XM_027251890.1	0 ± 0	0 ± 0	3,13 ± 0,02	1,7 ± 0,03	0 ± 0	2,46 ± 0,09	0 ± 0	0 ± 0
LOC113741416	CaLOX10	XM_027268943.1	13,32 ± 1,10	5,24 ± 0,53	4,33 ± 0,59	0 ± 0	4,97 ± 0,82	1,78 ± 2,27	1,23 ± 0,2	69,06 ± 9,80
LOC113741614	CaLOX11	XM_027269183.1	14,67 ± 1,40	6,1 ± 0,45	5,18 ± 1,07	2,19 ± 0,38	4,05 ± 0,34	6,32 ± 06,01	4,45 ± 0,60	48,61 ± 11,61
LOC113741495	CaLOX12	XM_027269026.1	43,32 ± 3,26	1,27 ± 0,10	12,14 ± 2,91	9,15 ± 1,83	6,64 ± 0,62	5,5 ± 0,56	6,13 ± 1,31	163,63 ± 23,93
LOC113714586	CaLOX13	XM_027238534.1	552,46 ± 8,00	316,59 ± 29,17	2,12 ± 0,19	53,63 ± 1,98	79,86 ± 10,06	128,9 ± 10,92	0,31 ± 0	104,91 ± 15,94
LOC113716533	CaLOX14	XM_027240915.1	6,98 ± 0,07	8,63 ± 1,20	8,76 ± 6,52	3,4 ± 1,48	6,66 ± 1,80	7 ± 0,66	15,56 ± 3,23	2,67 ± 1,87
LOC113700875	CaLOX15	XM_027221306.1	676,76 ± 33,67	365,6 ± 22,99	2,93 ± 0,59	58,86 ± 8,30	128,58 ± 34,00	24,2 ± 6,92	0,58 ± 0	135,56 ± 11,43
LOC113730976	CaLOX16	XM_027255986.1	0 ± 0	0 ± 0	3,57 ± 2,76	664,91 ± 40,32	0 ± 0	0 ± 0	2,74 ± 0,54	0 ± 0
LOC113726022	CaLOX17	XM_027249481.1	0 ± 0	0 ± 0	16,11 ± 1,41	3,77 ± 3,78	1 ± 0	0 ± 0	5,8 ± 1,16	832,18 ± 14,13
LOC113719047	CaLOX18.1 e CaLOX18.2	XM_027244035.1;XM_027244036.1	623,99 ± 44,92	686,73 ± 61,64	141 ± 12,15	374,48 ± 30,12	645,87 ± 55,18	732,51 ± 25,26	186,28 ± 16,84	161,72 ± 16,88

Legenda: Tratamento com ácido hexanoico (Hx) em folhas e raízes. Condições: Catuaí controle folha (CC_folha), Catuaí + ácido hexanoico folha (CHX_folha), Catuaí controle raiz (CC_raiz), Catuaí + ácido hexanoico raiz (CHX_raiz), obatã controle folha (OC_folha), obatã + ácido hexanoico folha (OHX_folha), obatã controle raiz (OC_raiz) e obatã + ácido hexanoico raiz (OHX_raiz), ± desvio padrão.

Tabela 4. Valores de RPKM (*Reads Per Kilobase Million*) dos genes de lipoxigenases em *C. arabica* em frutos verdes, amarelos e vermelhos.

ID gene	Nome do gene	ID transcrito	Fruto verde	Fruto amarelo	Fruto vermelho
LOC113727628	CaLOX1	XM_027251892.1	0,44 ± 0,04	0,46 ± 0,23	0,25 ± 0,29
LOC113732840	CaLOX2	XM_027258829.1	124,57 ± 64,80	80,97 ± 13,51	21,04 ± 9,21
LOC113732839	CaLOX3	XM_027258827.1	0,06 ± 0,02	0,20 ± 0,26	0,38 ± 0,33
LOC113727627	CaLOX4	XM_027251891.1	0,24 ± 0,07	0,51 ± 0,33	1,28 ± 1,67
LOC113727625	CaLOX5	XM_027251889.1	5,45 ± 2,08	0,21 ± 0,12	0,00 ± 0,00
LOC113732838	CaLOX6	XM_027258826.1	26,97 ± 10,78	25,64 ± 5,10	8,18 ± 2,74
LOC113738240	CaLOX7	XM_027265485.1	0,96 ± 0,96	0,89 ± 0,07	0,18 ± 0,31
LOC113734115	CaLOX8	XM_027260465.1	1,24 ± 0,37	0,92 ± 0,20	0,18 ± 0,31
LOC113727626	CaLOX9	XM_027251890.1	0,03 ± 0,04	0,00 ± 0,00	0,00 ± 0,00
LOC113741416	CaLOX10	XM_027268943.1	0,08 ± 0,02	0,00 ± 0,00	0,36 ± 0,42
LOC113741614	CaLOX11	XM_027269183.1	0,72 ± 0,74	1,57 ± 0,53	0,27 ± 0,47
LOC113741495	CaLOX12	XM_027269026.1	0,89 ± 0,53	1,96 ± 0,17	1,46 ± 0,36
LOC113714586	CaLOX13	XM_027238534.1	2,54 ± 0,67	0,22 ± 0,39	0,00 ± 0,00
LOC113716533	CaLOX14	XM_027240915.1	5,04 ± 1,80	9,95 ± 1,06	9,09 ± 1,89
LOC113700875	CaLOX15	XM_027221306.1	62,12 ± 32,87	25,68 ± 14,49	1,48 ± 1,21
LOC113730976	CaLOX16	XM_027255986.1	0,00 ± 0,00	0,05 ± 0,05	0,00 ± 0,00
LOC113726022	CaLOX17	XM_027249481.1	0,10 ± 0,14	0,96 ± 0,44	0,34 ± 0,37
LOC113719047	CaLOX18.1	XM_027244035.1	2,63 ± 2,12	4,75 ± 2,33	5,90 ± 2,40
LOC113719047	CaLOX18.2	XM_027244036.1	0,78 ± 0,36	1,30 ± 1,41	0,00 ± 0,00

Legenda: CaLOX – *Coffea arabica* e ± desvio padrão.

3.2.6 Perfis de expressão de genes LOX baseados em qPCR

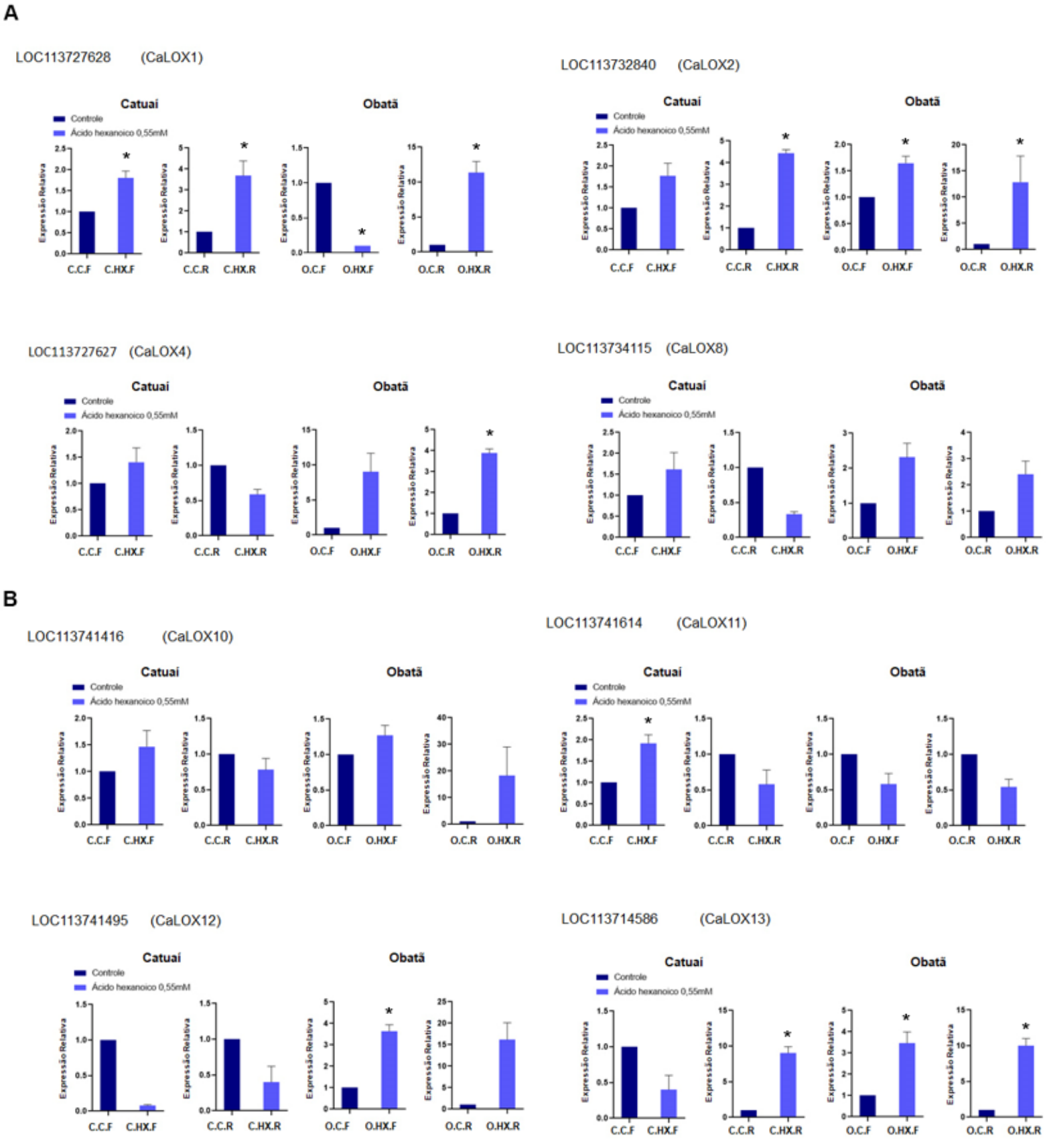
Verificamos se a aplicação exógena de ácido hexanoico pode modular o perfil transcricional de lipoxigenases em folhas e raízes de *C. arabica* cv. Catuaí Vermelho e *C. arabica* cv. Obatã (Figura 6 A, B e C). Para isso, foram escolhidos os seguintes genes: (A) LOC113727628 (CaLOX1), (B) LOC113732840 (CaLOX2), (C) LOC113727627 (CaLOX4), (D) LOC113734115 (CaLOX8), (E) LOC113741416 (CaLOX10), (F) LOC113741614 (CaLOX11), (G) LOC113741495 (CaLOX12), (H) LOC113714586 (CaLOX13), (I) LOC113700875 (CaLOX15), (J) LOC113730976 (CaLOX16), (K) LOC113726022 (CaLOX17) e (L) LOC113719047 (CaLOX18).

A análise do perfil transcricional por qPCR de folhas e raízes das cultivares Catuaí Vermelho (CV) e Obatã (OB) revelou diversos padrões de expressão dos genes de lipoxigenase de *Coffea arabica* (CaLOX) em

resposta ao ácido hexanoico (Figura 6). O gene CaLOX1 apresentou expressão diferencial indutora em folha e raiz de Catuaí e raiz de Obatã. E expressão diferencial repressora em folha de Obatã. O gene CaLOX2 apresentou expressão diferencial indutora em raiz de Catuaí e folha e raiz de Obatã. O gene CaLOX4 apresentou expressão diferencial indutora apenas em Obatã raiz. O gene CaLOX11 apresentou expressão diferencial indutora apenas em Catuaí folha. O gene CaLOX12 apresentou expressão diferencial indutora apenas em Obatã folha. O gene CaLOX13 e CaLOX15 apresentaram expressões diferenciais indutoras em raiz Catuaí e folha e raiz de Obatã. E os genes CaLOX17 e CaLOX18 apresentaram expressões diferenciais indutoras em folhas e raízes de Obatã. Assim, sugerimos que estes dois últimos genes podem estar relacionados às funções específicas para essa cultivar.

Esses padrões variados de expressão gênica sugerem que os genes CaLOX desempenham funções específicas e adaptativas em cada cultivar de café, atuando de forma diferenciada nos processos fisiológicos das plantas de Catuaí Vermelho e Obatã. A presença de genes com padrões similares de indução e repressão pode indicar funções relacionadas ou complementares no desenvolvimento e resposta das plantas a diferentes condições ambientais e estresses.

FIGURA 6. Perfil transcricional de genes lipoxigenases em resposta à aplicação de 0,55mM de ácido hexanoico em solução nutritiva.



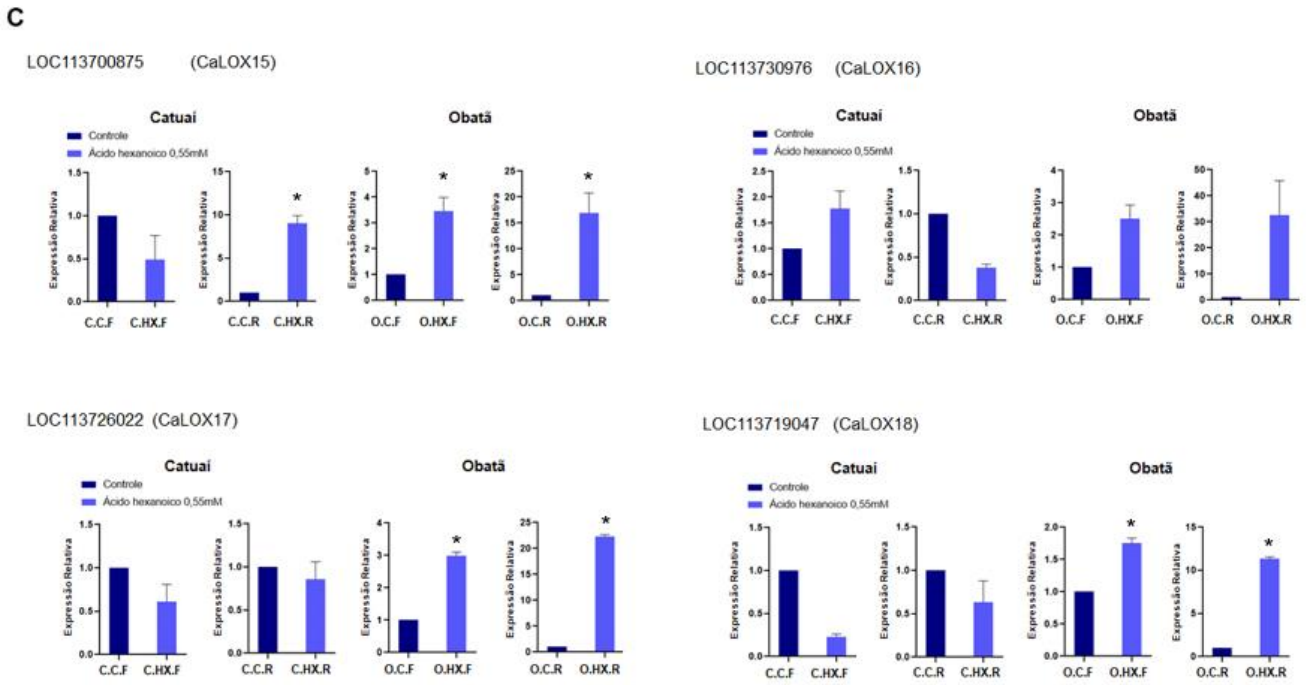


Figura 6. Perfil transcricional dos genes de lipoxigenases em *C. arabica* cv. Catuaí vermelho e *C. arabica* cv. Obatã. Tratamento com ácido hexanoico (HX) em folhas e raízes. (A) LOC113727628 (CaLOX1), LOC113732840 (CaLOX2), LOC113727627 (CaLOX4), LOC113734115 (CaLOX8), (B) LOC113741416(CaLOX10), LOC113741614 (CaLOX11), LOC113741495 (CaLOX12), LOC113714586 (CaLOX13), (C) LOC113700875 (CaLOX15), LOC113730976 (CaLOX16), LOC113726022 (CaLOX17) e LOC113719047 (CaLOX18). Os experimentos foram realizados com a atualização de três réplicas biológicas e três réplicas técnicas cada. Barra: desvio padrão. *Significância de p<0,05 (t Student).

3.2.7 Análises de correlação entre os valores de RNAseq, qPCR e atividade enzimática

Também realizamos a correlação entre os valores de RPKM (*Reads Per Kilobase Million*), - obtidos a partir dos dados de RNAseq, e os valores de qPCR, a fim de verificar se os perfis transcricionais por RPKM e qPCR, são correspondentes (Figura 7 A). Da mesma forma, realizamos a análise de correlação entre os valores de qPCR e os valores de atividade enzimática de lipoxigenase (Figura 7 B). Para a correlação entre os valores de RPKM e RNAseq, os genes CaLOX2 e CaLOX12 apresentaram um valor de **p** de 0,6, o que representa uma correlação moderada. E os genes CaLOX13 e CaLOX15, apresentaram um valor de **p** de 0,8, o que representa uma correlação forte. Para a correlação entre os valores de qPCR e atividade enzimática de lipoxigenases os genes CaLOX13 e CaLOX16, apresentaram

um valor de **p** de 0,6, o que representa uma correlação moderada. E os genes CaLOX2 e CaLOX15 apresentaram um valor de **p** de 0,8, o que representa uma correlação forte.

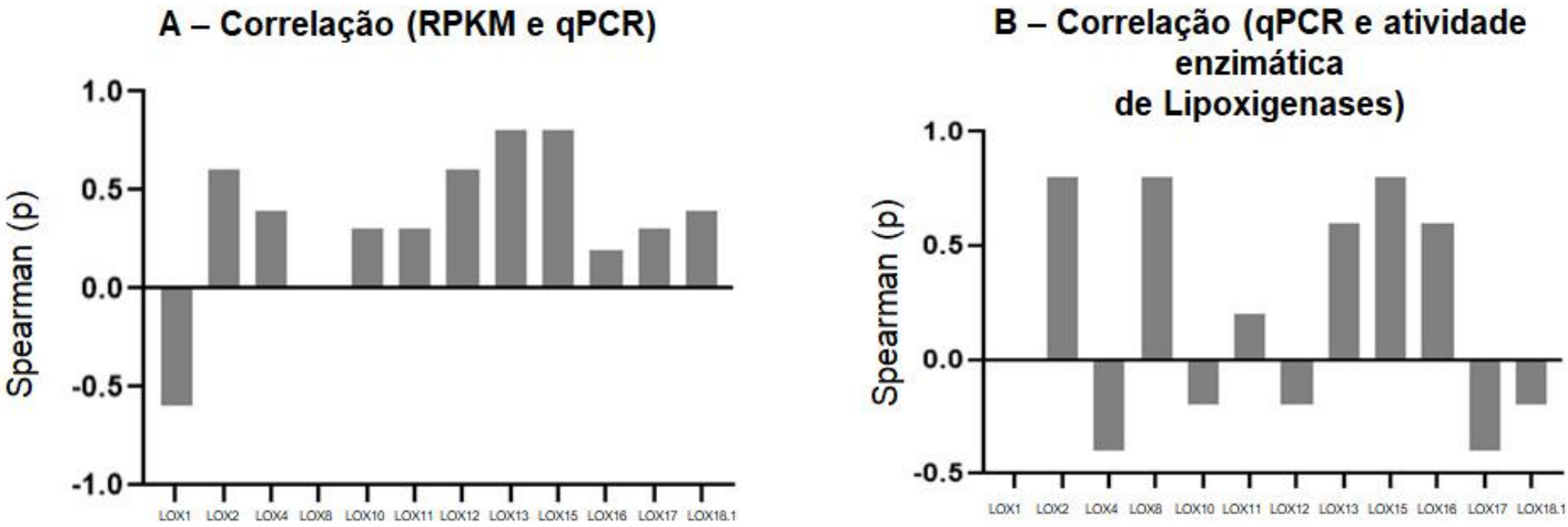


Figura 7. Análise de correcção de Spearman (p), em *Coffea arabica*, entre as comparações: valores de RPKM e valores de qPCR (A) e valores de qPCR e atividade enzimática de Lipoxigenases (B).

3.3 Discussão

É sabido que a espécie *Coffea arabica* é derivada de uma fusão genômica entre as espécies *C. canephora* e *C. eugenioides* (SATTLER et al., 2022). Em nossos resultados o número de genes de lipoxigenases identificados corrobora com a evolução dessas três espécies de *Coffea*. Com tudo, os 9 genes de *Coffea canephora* identificados pelo PLAZA 4.0 nesse estudo, não se enquadram nos critérios de classificação de lipoxigenases estipulados por CAMARGO et al. (2023). De acordo com esses critérios os genes Cc02_g33790.1 (CcLOX1) e Cc00_g30760.1 (CcLOX4) não devem ser incluídos como genes de lipoxigenases, logo a re-anotação desses genes se faz necessária. Em *Coffea arabica*, - genes identificados pelo site NCBI (https://www.ncbi.nlm.nih.gov) -, dos 18 genes identificados 2 apresentaram – se com o mesmo locus, LOC113719047, porém com a expressão de duas proteínas diferentes XM_027244036.1 (CaLOX18.1) e XM_027244035.1 (CaLOX18.2), o que sugere a ocorrência de *splicing* alternativo para esse gene (Tabela 1).

Identificamos dois subgrupos de genes de lipoxigenases, divididos em LOX9 e LOX13 (Figura 2). Esses resultados corroboram com os estudos relatados na literatura. Alguns trabalhos que envolvem a análise da evolução de proteínas lipoxigenases em outras espécies vegetais, como: banana, rabanete e maçã, também demonstraram a subdivisão das proteínas lipoxigenases em dois grupos (LIU et al., 2021; WANG et al., 2019; VOGT et al., 2013).

Ao analisar a diversidade estrutural dos genes lipoxigenases, observamos que a quantidade média de éxons e íntrons não apresenta grandes diferenças entre as espécies aqui estudadas (*C. arabica*, *C. eugenioides* e *C. canephora*). Nesses casos, o número de éxons varia de 7 a 11, enquanto o de íntrons fica entre 6 e 10 (conforme Figura 4). Entretanto, alguns íntrons - nos genes CaLOX5, CaLOX6, CaLOX13, CaLOX15, CcLOX8, CeLOX2, CeLOX3 e CeLOX7 - possuem tamanhos estruturais maiores, com médias entre 3 e 6,5 kilobases (kb). Essas discrepâncias suscitam a necessidade de averiguar se o processamento de RNA é de fato realizado

conforme o predito. Nossos resultados estruturais estão em consonância com os dados de éxons e íntrons de lipoxigenases encontrados em outras espécies vegetais. A literatura descreve que o número médio de éxons varia entre 8 e 11, e o de íntrons, entre 7 e 10 (CAMARGO et al., 2023; LIU et al., 2021; GUO et al., 2016).

Identificamos a localização celular dos genes de lipoxigenases nas três espécies analisadas. É descrito na literatura que proteínas lipoxigenases podem ser classificadas com base em sua localização celular. A proteína LOX tipo I é encontrada no citoplasma e LOX tipo II nos peptídeos de sinal de direcionamento de organelas (SHABAN et al., 2018). Porém, CAMARGO et al. (2023), demonstraram que a localização celular, não é a melhor forma de classificação de proteínas LOX. Contudo, após as análises realizadas pelos sites: Plant-mSubP, DeepLoc 2.0 e Euk-mPLoc 2.0, percebemos que não há concordância entre os métodos e sites para um consenso sobre a localização subcelular de proteínas LOX (Tabela 2).

Alguns trabalhos já demonstraram os posicionamentos de genes de lipoxigenases em mapas cromossômicos (LIU et al., 2021). Os nossos dados obtidos sobre a localização cromossômicas dos genes de lipoxigenases nos genomas de *Coffea arabica*, *Coffea canephora* e *Coffea eugenioides*, evidenciaram que o subgenoma de *Coffea canephora* apresenta uma maior conservação de genes LOX, quando comparado ao subgenoma de *Coffea eugenioides*. Já que, os genes de *Coffea arabica* e de *Coffea canephora* apresentaram 100% de correspondência de posicionamento cromossômico. Porém, para *Coffea arabica* e *Coffea eugenioides* foram identificadas algumas mudanças de posições: Ca4e (*Coffea arabica*), com três genes LOX e Ce4 (*Coffea eugenioides*) com um gene LOX. E Ca11e (*Coffea arabica*), com um gene LOX e Ce11, (*Coffea eugenioides*) com dois genes LOX (Figura 3).

As lipoxigenases estão presentes em diversos órgãos vegetais e exibem diferentes níveis de expressão ao longo do desenvolvimento da planta (VISWANATH et al., 2020; ZHENG et al., 2022). Em nosso estudo, examinamos folhas e raízes de *C. arabica* cv. Catuaí Vermelho (CV) e *C. arabica* cv. Obatã (OB), expostas ao ácido hexanoico exógeno e observamos

expressões distintas em alguns genes de lipoxigenases, conforme os valores de RPKM obtidos a partir dos dados de RNAseq.

O gene LOC113727628 (CaLOX1) apresentou maior expressão em raízes de plantas controle de ambas as cultivares (CV e OB). Já o gene LOC113732840 (CaLOX2) teve maior expressão em raízes, tanto em condições controle quanto com HX, nas duas variedades. Isso indica a atuação desse gene nesse órgão vegetal.

Os genes LOC113732839 (CaLOX3), LOC113727627 (CaLOX4) e LOC113741495 (CaLOX12) exibiram maior expressão em raízes de Obatã tratadas com HX. O gene LOC113714586 (CaLOX13) mostrou maior expressão em folhas de Catuaí (controle e com HX) e em folhas e raízes de Obatã tratadas com HX, sugerindo a atuação desse gene nessas condições. O gene LOC113700875 (CaLOX15) apresentou maior expressão em folhas de Catuaí (controle e com HX) e em folhas controle de Obatã, além de raízes tratadas com HX. O gene LOC113730976 (CaLOX16) mostrou maior expressão apenas em raízes de Catuaí tratadas com ácido hexanoico, indicando a ação desse gene nesse órgão vegetal. O gene LOC113726022 (CaLOX17) exibiu maior expressão somente em raízes de Obatã tratadas com HX. Por fim, o gene LOC113719047 (CaLOX18.1 e CaLOX18.2) apresentou altos valores de expressão em ambas as cultivares, sob todas as condições analisadas.

O ácido hexanoico pode ser usado como um modulador para gerar diferentes respostas de defesa em alguns sistemas vegetais (CAMAÑES et al., 2015; LÓPEZ-GALIANO et al, 2019; ARANEGA-BOU et al., 2014). Proteínas lipoxigenases estão associadas às respostas de defesa e resistência das plantas por meio de seus elementos sinalizadores, como o metil jasmonato (HE et al., 2022). Em nosso estudo, utilizamos RNAs de plantas *C. arabica* cv. Catuaí Vermelho (CV) e *C. arabica* cv. Obatã (OB) submetidas ao tratamento com ácido hexanoico e seus respectivos controles para validar a expressão diferencial usando a técnica de qPCR (Figura 6). Os genes selecionados para validação foram aqueles com valores de RPKM (dados obtidos por meio do sequenciamento – RNAseq) que sugeriam expressão diferencial.

Já foi estudada a atividade de proteínas lipoxigenases em folhas de *Coffea arabica* submetidas ao processo de herbivoria. Foi observado que lipoxigenases podem exercer um importante papel nos processos fisiológicos de defesa vegetal (MERIÑO-CABRERA et al., 2018). Proteínas de lipoxigenases do tipo 13 também foram identificadas em espécies de *Coffea* submetidas à seca (MARQUES et al., 2022). Em nosso estudo, folhas e raízes de Catuaí e folhas e raízes de Obatã, submetidas ao tratamento com HX, apresentaram maiores valores de atividade de lipoxigenases quando comparadas aos seus respectivos tratamentos. Os valores mais altos de atividade foram observados em Catuaí folha com HX e Obatã raiz com HX (Figura 5). Vários mecanismos relacionados à defesa participam do desenvolvimento de respostas das plantas contra estresses bióticos e abióticos. O ácido jasmônico, o metil jasmonato e o ácido abscísico, são moléculas de sinalização que desencadeiam a produção de compostos antiestresse (FABBRI et al., 2000). Esses compostos interagem positivamente com as lipoxigenases para regular vários processos físicos-bioquímicos em plantas (HOU et al., 2018; FEUSSNER et al., 2002; SANTINO et al., 2005). Intermediários e produtos finais do metabolismo da LOX podem desencadear proteínas quinases, transmitir sinais e produzir sua transdução. (SINGH et al., 2022). É sabido que plantas danificadas mecanicamente, aumentam sua síntese de ácido jasmônico, esse, por sua vez, jutamente com seus precursores participaram da via de sinalização multiconstituinte responsável pela imunidade vegetal (ZHOU et al., 2003). É reconhecido também que as vias da lipoxigenases são cruciais para a sinalaização causada por efeitos estressores, assim, essas enzimas servem eficientemente como marcadores moleculares para monitorar a tolerância ao estresse nas plantas (KOSAKIVSKA et al., 2014). Nossos resultados corroboram com os dados descritos na literatura sobre a ação das lipoxigenases em respostas a um modulador, como o HX, e consequentemente ao estresse abiótico. Dessa forma, destacamos às respostas diferencias de LOX ao HX, em folhas e raízes de Catuaí e Obatã. Possivelmente ácido hexanoico absorvido pelas raízes durante o nosso experimento, deva ter iniciado vias de sinalização, envolvendo proteínas LOX, o que resultou em expressões diferenciais de

certos genes de lipoxigenases em folhas e raízes, assim como, uma maior atividade dessas enzimas nesses mesmos órgãos.

Para um maior embasamento dos resultados, e a fim de verificar se há uma relação entre os valores obtidos, de RPKM (RNAseq), qPCR e atividade enzimática de lipoxigenases, os dados gerados neste estudo, foram submetidos à correlação de Spearman. De acordo com MUKAKA et al. (2012), os valores de Spearman (p) entre 0 e 0,3 (ou 0 e -0,3) são biologicamente desprezíveis; entre 0,31 e 0,5 (ou -0,31 e -0,5) são correlações fracas; entre 0,51 e 0,7 (ou -0,51 e -0,7) são moderadas; entre 0,71 e 0,9 (ou -0,71 e 0,9) são correlações fortes; e > 0,9 (ou < -0,9) são consideradas muito fortes. Desta forma, foram realizadas as seguintes correlações: valores de RPKM (RNAseq) versus, os valores de qPCR e os valores de atividade enzimática de lipoxigenases, versus, os valores de qPCR (Figura 7).

O gene CaLOX2, apresentou valor de correlação moderado (0,6 p) para qPCR e RNAseq e forte (8,0 p) para qPCR e atividade de LOX. É importante salientar que esse gene também apresentou valores altos de expressão, medidos em RPKM, para raízes controle – em ambas cultivares, e frutos (verdes, amarelos e vermelhos), sendo um possível candidato para futuros estudos em raiz e frutos. O gene CaLOX13 apresentou valor de correlação forte (0,8) para qPCR e RNAseq e moderado (0,6 p) para qPCR e atividade de LOX e valores altos de RPKM e folhas controle e com HX de Catuaí. O gene CaLOX15 apresentou valores de correlação fortes (8,0 p) para ambas condições e valores, em RPKM, altos para folhas controle e com HX em Catuaí e frutos verdes e amarelos. O gene CaLOX12 apresentou valor de correlação moderado (6,0 p) apenas para comparação entre qPCR e RNAseq e o gene CaLOX16 apresentou valor de correlação moderado (0,6) apenas para comparação entre qPCR e atividade de LOX. Assim, genes que apresentaram valores de correlação moderados ou fortes CaLOX2 (LOX do tipo 9), CaLOX12 (LOX do tipo 13), CaLOX13 (LOX do tipo 13), CaLOX15 (LOX do tipo 13) e CaLOX16 (LOX do tipo 13), são possíveis candidatos para

análises mais específicas em futuros trabalhos que envolvam proteínas lipoxigenases em *Coffea*.

3.4 Material e Métodos

3.4.1 Identificação de Genes da Família LOX inclusos na árvore evolutiva

Os genes LOX de *Arabidopsis thaliana* utilizados neste estudo foram previamente identificados por UMATE et al. (2011). E os genes LOX de *Daucus carota* e *Coffea canephora* foram identificados como descrito por CAMARGO et al. (2023). A seguir, as sequências de genes de *Coffea canephora*, previamente identificadas, foram comparadas (blastn) com as sequências genômicas de *Coffea arabica* (disponível em: https://www.ncbi.nlm.nih.gov/genome/77) e *Coffea eugenioides* (https://www.ncbi.nlm.nih.gov/genome/73741) para a identificação dos genes LOX das respectivas espécies. Genes com *score* acima de 200 e *evalue* acima de e-50 foram selecionados para análise.

3.4.2 Alinhamento de Múltiplas Sequências e Análise Filogenética

As sequências de nucleotídeos (CDS) de *Arabidopsis thaliana*, *Daucus carota, Coffea arabica, Coffea canephora e Coffea eugenioides* foram alinhadas utilizando o programa MUSCLE (EDGAR, 2004) e, em seguida, traduzidas em sequências de aminoácidos por meio da ferramenta TranslatorX (http://translatorx.co.uk/). Os alinhamentos de aminoácidos serviram de base para elaborar o perfil filogenético dos membros da família LOX, empregando o método de Maximum Likelihood (MV) no programa MEGA11 (TAMURA et al., 2021). Foram realizadas 1000 réplicas bootstrap (FELSENSTEIN, 1985), com modelo de Poisson e taxas uniformes para a opção "taxas entre sítios". As lacunas no alinhamento foram tratadas como "exclusão em pares", e o modelo de proteína aplicado foi o LG + G + I + F. A árvore foi visualizada e anotada com o auxílio do iTOL (LETUNIC; BORK 2016).

3.4.3 Determinação das estruturas genéticas em *Coffea arabica, Coffea canephora* e *Coffea eugenioides*

O programa Gene Structure Display Server v2.0 (HU et al., 2015), com parâmetros padrão, foi usado para analisar a estrutura exon-intron dos genes LOX em Coffea. Sequências genômicas e de CDS no formato FASTA correspondentes aos genes de *Coffea arabica, Coffea canephora* e *Coffea eugenioides* foram inseridas para gerar as estruturas gênicas.

3.4.4 Localização Celular de Genes LOX em *Coffea arabica, Coffea canephora* e *Coffea eugenioides*

As localizações subcelulares das proteínas LOX em *Coffea arabica*, *Coffea canephora* e *Coffea eugenioides* foram estimadas utilizando três ferramentas online. A primeira foi o Plant-mSubP, com o modelo de predição Dipep, disponível em http://bioinfo.usu.edu/Plant-mSubP/ (SAHU et al., 2019). A segunda, o DeepLoc 2.0, com o modelo High-quality e formato Long output, acessível em http://services.healthtech.dtu.dk/service.php?DeepLoc-2.0 (THUMULURI et al., 2022). Por fim, a Euk-mPLoc 2.0, encontrada no servidor Euk-mPLoc 2.0 em http://www.sjtu.edu.cn (CHOU et al., 2010).

3.4.5 Mapeamento dos genes LOX

A localização cromossômica dos genes de lipoxigenases, para as espécies *Coffea arabica*, *Coffea eugenioides* e *Coffea canephora*, foi realizada através do site MapGene2Chromosome v2 (http://mg2c.iask.in/mg2c_v2.0/). Dados disponíveis no NCBI (https://www.ncbi.nlm.nih.gov/data-hub/genome/GCF_003713225.1/ - *Coffea arabica*, https://www.ncbi.nlm.nih.gov/data-hub/genome/GCF_003713205.1/ - *Coffea eugenioides* e https://www.ncbi.nlm.nih.gov/data-hub/genome/GCA_900059795.1/ - *Coffea canephora*), com o tamanho dos cromossomos de cada espécie e suas respectivas posições gênicas, foram utilizados como arquivos de entrada para a construção do mapa cromossômico.

3.4.6 Experimento para análise transcricional de genes LOX em *Coffea arabica*: efeito do ácido hexanoico

Foram utilizadas plantas de *C. arabica* das cultivares Catuaí Vermelho IAC 144 e Obatã IAC 1669-20, com cerca de cinco meses de idade (3-4 pares de folhas). Para os experimentos, as plantas foram selecionadas por uniformidade de tamanho e foram transferidas para vasos escuros contendo 3 L de solução nutritiva aerada (SNA), adaptada de CLARK (1975) por (de CARVALHO 2014). A solução foi composta por: K2SO4 (1068 mM), MgSO4.7H2O (332,5 mM), KH2PO4 (266 mM), CaCl22H2O (66 mM) e NH4NO3 (5333 mM). Ferro e micronutrientes foram fornecidos por mistura comercial de sais quelatados (ConMicros Standard, Conplant) nas seguintes concentrações (em µg.L−1): Fe (363), Cu (91), Zn (37), Mn (91), B (91), Mo (18) e Ni (17). O experimento foi conduzido em sala de crescimento vegetal, com temperatura controlada a 23°C (variando de 21°C a 25°C). As plantas foram mantidas em um ciclo dia/noite de 12 horas. A iluminação artificial foi fornecida por painéis de LED com densidade de fluxo de fótons fotossinteticamente ativos de aproximadamente 400µmol.m-2.s-1. O pH da solução nutritiva foi ajustado e mantido entre 5,5 e 5,6 diariamente. As plantas foram mantidas em aclimatação por aproximadamente 96 h. Após o período de aclimatação, a solução nutritiva foi substituída. Foram impostos os seguintes tratamentos: a) SNA (controle); b) SNA + ácido hexanóico (concentração final 0,55 mM) por 48 h. As plantas foram cultivadas em 3 a 6 parcelas (vasos de plástico) com três plantas (repetições) para cada tratamento. Os vasos de plantas foram agrupados e considerados uma réplica. Os experimentos foram repetidos 3 vezes para obter réplicas biológicas. As folhas maduras do terço médio e as raízes principais e laterais das plantas foram coletadas, maceradas em nitrogênio líquido e armazenadas em freezer a -80° até a extração do RNA.

3.4.7 Extração de RNA total para preparo da biblioteca - RNA-Seq e para as análises de RT- qPCR

O RNA total foi extraído de pools de folhas usando o kit *RNeasy Plant* (Qiagen, *Hilden*, *North Rhine-Westphalia*, Alemanha). As amostras de RNA

total foram purificadas usando o kit *RNeasy Minielute Cleanup* (*Qiagen, Hilden, North Rhine-Westphalia*, Alemanha). A pureza do RNA foi determinada usando um espectrofotômetro NanoDrop ND-100 (*Thermo Scientific*, San Jose, CA, EUA), as concentrações foram obtidas usando um fluorímetro Qubit (*Thermo Fisher Scientific, Wilmington*, DE, EUA) e as integridades de RNA foram verificadas em eletroforese em gel de agarose a 1%.

3.4.8 Preparo da biblioteca de RNAseq e sequenciamento

A biblioteca de sequenciamento de RNA Poly (A) foi preparada seguindo o protocolo de preparação de amostra *TruSeq-stranded*-mRNA da Illumina (Illumina *Technologies, San Diego*, CA). O sequenciamento pareado (2 X 150 pb) foi realizado no sistema de sequenciamento NovaSeq 6000 da Illumina na LC *Sciences* (*Houston*, TX, EUA). Os dados foram depositados no *European Nucleotide Archive* (ENA), submissão ERA6282544 e no Zenodo (doi: 10.5281/zenodo.5517785).

3.4.9 Desenho dos primers para sequências de *Coffea arábica*

Para a construção dos primers as sequências de genes LOX de *Coffea arabica*, identificadas no NCBI, foram alinhadas no programa MUSCLE (MUSCLE < Multiple Sequence Alignment < EMBL-EBI). Para a distinção dos genes considerados ambíguos, as sequências alinhadas foram analisadas pelo site *HIV sequence database* (https://www.hiv.lanl.gov/content/sequence/CONSENSUS/AmbigCon.html). Os alinhamentos de nucleotídeos foram obtidos com códigos de ambiguidade IUPAC (*International Union of Pure and Applied Chemistry*). Os primers específicos para cada gene foram desenhados, através do software Primer3Plus (https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi), para amplificar produtos que variam de 95 - 140 pb.

3.4.10 Análises do perfil transcricional de genes LOX por RT-qPCR

Os RNAs extraídos foram submetidos ao kit *High Capacity cDNA Reverse Transcriptase* (*Thermo Fisher Scientific, Wilmington*, DE, EUA) para a construção dos DNAs complementares (cDNAs), seguindo as instruções do fabricante, em um volume final de 20 µl, usando 5 µg de RNA total. O RT-qPCR foi realizado no QuantiStudio3 (*Applied Biosystems, Carlsbad*, CA, EUA) e seguindo procedimentos básicos relatados em publicação anterior do grupo (IVAMOTO et al., 2017). A mistura de reação continha 7,5 µl de GoTaq® *Hot Start Polymerase (Promega, Madison, Wisconsin, EUA)*, 0,3 µl de cada primer (3 µM), 1 µl de cDNA (10 ng/µL) e 5,5 µl de água MilliQ em um volume final de 15 µl. As condições de RT-qPCR foram 95°C por 2 min; 40 ciclos de 94°C por 30 s, 60°C por 60s. As curvas de *melting* foram analisadas para verificar a presença de um único produto incluindo um controle negativo. Todas as reações foram realizadas usando 3 réplicas biológicas e técnicas. As expressões relativas foram calculadas como $2^{-\Delta\Delta Ct}$, onde −ΔCt = Ct condição tratamento – Ct condição controle e ΔΔCq = ΔCq alvo - Δ Cq da média geométrica dos genes dede referência GAPDH e EF (Tabela 5) conforme recomendado para *C. arabica* (CARVALHO et al., 2013).

Tabela 5. Sequências de primers desenhados para a análise de expressão diferencial de genes LOX em *Coffea Arabica*.

Nome do Gene	Locus	Sequências primer *forward*	Sequências primer *reverse*	Tamanho do primer (pb*)
CaLOX1	LOC113727628	5' GACTAGAGCAGTTCGCTCCTT 3'	3' ATAGGAAAACGACAACTGCCC 5'	21
CaLOX2	LOC113732840	5' GACTTGAGCAGTTCGCTCCCC 3'	3' ATAGGAAAACGACAACTGCCC 5'	21
CaLOX4	LOC113727627	5' CACTTCCGTGACACAATGAACA 3'	3' TGCTTGTTCAGGAAAGATCCAA 5'	22
CaLOX8	LOC113734115	5' GTCGTTAACTTCTTACCCAGTC 3'	3' CTGTCCACCATGACTGGAGTTC 5'	22
CaLOX10	LOC113741416	5' CCTTCCCTCCTGTTAGCAAGCT 3'	3' TGGTAAGTAGGCATCGTGGTGG 5'	22
CaLOX11	LOC113741614	5' AGTGGCTAACAAGCACCAACAA 3'	3' TCTTAATGCTTTCAACCCAGCA 5'	22
CaLOX12	LOC113741495	5' GAATAAGCTTCCATTGCCCAAG 3'	3' GGCCATAGATTTCAGGGTCAAG 5'	22
CaLOX13	LOC113714586	5' TTTTCGGCTGACAGTTCCGAAT 3'	3' GTCCAAAACCCTGCTATCCTCG 5'	22
CaLOX15	LOC113700875	5' TTTTCGGCTGACAGTTCCGAAT 3'	3' TACCAAAACCCTGCTATCCTCG 5'	22
CaLOX16	LOC113730976	5' TGTGAAATGTGCTAATCCAACG 3'	3' GCTAACAAGCTGTAGCACAACAC 5'	22
CaLOX17	LOC113726022	5' CTGCTCACCGTCAGTTAAGCAT 3'	3' TGGTACCATTTGCATTGAGGAG 5'	22
CaLOX18.1 e CaLOX18.2	LOC113719047	5' GAACCATCCATTCGTTGTTGAA 3'	3' TCAGGTGTTTGAGATGGCAAGT 5'	22

Legenda: Genes de *Coffea arabica* representados por - CaLOX. Os primers foram desenhados através do software Primer3Plus.

3.4.11 Determinação da atividade enzimática de lipoxigenase – extrato foliar

Amostras de folhas e raízes de todas as condições de análise foram congeladas em nitrogênio líquido e trituradas com o auxílio de almofariz e pistilo. A extração proteica foi realizada com a adição de 1 mL de tampão 0,05 M fosfato de sódio (pH 6,5) e 2% (p/v) PVPP (polivinil polipirrolidona), de acordo com Meriño-Cabrera (2018). Pequenas modificações foram realizadas a fim de evitar a oxidação do material foliar, como a adição de 1mM EDTA (ácido etilenodiamino tetra-acético) e 2% β-mercaptoetanol. Em seguida, o extrato foi centrifugado por 40 minutos, 20000 g, 4 °C. O sobrenadante foi utilizado para as análises em espectrofotômetro e a determinação da concentração proteica foi obtida pelo método de Bradford (1976).

3.4.12 Determinação da atividade de lipoxigenases

A atividade da lipoxigenase em folhas e raízes de café foi determinada em espectrofotômetro com a utilização de ácido linoleico como substrato (AXELROD et al., 1981). O preparo do ácido linoleico foi realizado de acordo com MERIÑO-CABRERA, 2015. Para tanto, foi preparada solução estoque de linoleato de sódio 10 mM, utilizando-se ácido linoléico, aproximadamente 99% (SIGMA), como se segue: a um erlenmeyer, envolvido por papel alumínio, contendo aproximadamente 10 mL de água deionizada, previamente fervida, adicionou-se 78 µL de ácido linoléico e 90 µL de tween 20 (SIGMA). Em seguida, a solução foi homogeneizada. Para clareamento da solução foram adicionadas gotas de 0,5 M NaOH. A solução estoque de linoleato de sódio foi armazenada em microtubos de 1,5 mL, envolvidos em papel alumínio e armazenados a -20^{o}C.

Para cada análise foram utilizados 20 µL e 30 µL do extrato de folhas e raízes, respectivamente e 4,0 µL da solução estoque de linoleato de sódio em 1,0 mL de tampão fosfato 50,0 mM, pH 6,5, como descrito por MERIÑO-CABRERA et al. (2018). A absorbância da reação foi determinada a cada 30 segundos a 234 nm por período de 1 minuto. A partir dos valores de absorbância obtidos a 234 nm foram calculadas as velocidades de formação dos produtos, usando a seguinte fórmula: onde: A234 = absorbância a 234 nm; ε = 25000 M -1 cm -1 (coeficiente de extinção molar dos hidroperóxidos

do ácido linoléico a 234 nm); l = 1,0 cm (caminho ótico) e t = tempo de incubação.

3.4.13 Análises de correlação para proteínas Lipoxigenases

Foram realizadas análises de correlação com as seguintes comparações: valores de expressão de RNAseq em RPKM (*Reads Per Kilobase Million*), versus os valores de expressão relativa em RT-qPCR. E valores de atividades de lipoxigenases, versus os valores de expressão relativa em RT- qPCR. Para isso, foi utilizado o coeficiente de correlação de Spearman, uma medida não paramétrica da correlação entre duas variáveis.

Cálculo do coeficiente:

Log_2 (tratamento/controle)

Os valores de correlação de cada gene – CaLOX1 a CaLOX18.1 foram atribuídos a partir das médias e dos valores dos cálculos de coeficiente obtidos para as duas cultivares (Catuaí e Obatã) e para os dois tecidos (folha e raíz).

3.5 Conclusões

Os genes de *Coffea arabica, Coffea eugenioides e Coffea canephora* identificados em nosso estudo corroboram os estudos de processos evolutivos de *Coffea arabica*, uma espécie alotetraploide. Os padrões de estrutura éxon-íntron dos genes LOX no gênero Coffea seguem um valor médio de 7 à 9 éxons e um valor médio de 7 à 10 íntrons. Os genes de *Coffea arabica* apresentaram alta correspondência de posicionamentos, quando comparados aos seus subgenomas*, Coffea eugenioides* e *Coffea canephora*. Nossos resultados demonstram que aplicação exógena de ácido hexanoico pode modular o perfil transcricional de lipoxigenases, em alguns genes de *C. arabica* cv. Catuaí Vermelho (CV) e *C. arabica* cv. Obatã (OB). Por fim, o nosso trabalho indica que alguns genes de lipoxigenases apresentam moderada ou forte correlação entre os valores de RPKM e expressão por qPCR e valores de atividade enzimática e expressão por qPCR.

4. CONSIDERAÇÕES FINAIS

Nossos estudos fornecem base para o estudo funcional e evolutivo das lipoxigenases em angiospermas, onde identificamos 247 genes LOX entre 23 espécies de angiospermas e plantas basais. Nossas análises filogenéticas identificaram clados que suportam LOX13 e dois clados principais para LOX9: LOX9_A e LOX9_B. Também foi possível identificar que há potenciais para novos padrões de localização subcelular e resíduos conservados de oxidação para LOX9 e LOX13, ainda inexplorados. Folhas e raízes submetidas à aplicação exógena de ácido hexanoico em *C. arabica* cv. Catuaí Vermelho (CV) e *C. arabica* cv. Obatã (OB), apresentaram expressão diferencial para alguns genes de lipoxigenases, quando comparados os valores de KPKM gerados a partir dos dados de RNAseq, assim como, para os dados gerados em qPCR. As atividades enzimáticas de lipoxigenases também apresentaram valores diferenciais, sendo que folhas e raízes de Obatã e Catuaí, submetidas ao ácido hexanoico, obtiveram maiores valores, quando comparados aos valores dos seus repectivos controles. Por fim, a partir dos valores de correlação de RPKM, qPCR e atividade enzimática, nossos estudos sugerem que, alguns genes de *Coffea arabica*: CaLOX2, CaLOX12, CaLOX13, CaLOX15 e CaLOX16, são possíveis candidatos para análises mais específicas em futuros trabalhos que envolvam proteínas lipoxigenases em genêro Coffea.

REFERÊNCIAS

Axelrod B, Cheesbrough T, Laasko S. Lipoxygenase from soybeans: EC 1.13.11.12. Linoleate: oxygen oxidoreductase. *Methods in Enzymology* **1981**, 71, 441–451. doi.org/10.1016/0076-6879(81)71055-3

Bertrand, B.; Bardil, A.; Baraille, H.; Dussert, S.; Doulbeau, S.; Dubois, E.; Severac, D.; Dereeper A.; Etienne, H. The Greater Phenotypic Homeostasis of the Allopolyploid *Coffea arabica* Improved the Transcriptional Homeostasis Over that of Both Diploid Parents. *Plant Cell Physiol* **2015**, 10, 2035-51. doi: 10.1093/pcp/pcv117

Bradford M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Analytical Biochemistry* **1976**, 72, 248–254. doi: 10.1006/abio.1976.9999.

Bustin, S.A.; Benes, V.; Garson, J.; et al. The need for transparency and good practices in the qPCR literature. *Nat Methods* **2013**, 10, 1063-1067. doi: 10.1038/nmeth.2697

Camañes, G.; Scalschi, L.; Vicedo, B.; González-Bosch, C.; García-Agustín, P. An untargeted global metabolomic analysis reveals the biochemical changes underlying basal resistance and priming in Solanum lycopersicum, and identifies 1-methyltryptophan as a metabolite involved in plant responses to Botrytis cinerea and Pseudomonas syringae. *Plant J* **2015** 1,125-39. doi: 10.1111/tpj.12964

Camargo, P. O.; Calzado, N. F.; Budzinski, I. G. F.; Domingues, D. S. Genome-Wide Analysis of Lipoxygenase (LOX) Genes in Angiosperms. *Plants* **2023**, *12*, 398. doi.org/10.3390/plants12020398

Cheng, B.; Furtado, A.; Henry, R.J. The coffee bean transcriptome explains the accumulation of the major bean components through ripening. *Scientific Reports* **2018**, 8, 11414. doi.org/10.1038/s41598-018-29842-4

Chou, K. C.; Shen, H. B. A new method for predicting the subcellular localization of eukaryotic proteins with both single and multiple sites: Euk-mPLoc 2.0. *PLoS One* **2010,** 4, e9931. doi: 10.1371/journal.pone.0009931

CONAB. Acompanhamento da safra Brasileira: Café. Disponível em: https://www.conab.gov.br/info-agro/safras/cafe Acesso em: 19 de janeiro de 2023.

Christie, W. W. Harwood, J. L. Oxidation of polyunsaturated fatty acids to produce lipid mediators. *Essays Biochem* **2020**, 3, 401-421. doi: 10.1042/EBC20190082

Cruz, F.; Kalaoun, S.; Nobile, P.; Colombo, C.; Almeida, J.; Barros, L. M. G.; Romano, E.; Grossi-de-Sá, M. F.; Vaslin, M.; Alves-Ferreira, M. Evaluation of coffee reference genes for relative expression studies by quantitative real-time RT-PCR. *Molecular Breeding* **2009**, 23, 607–616. doi.org/10.1007/s11032-009-9259-x

de Carvalho, K.; Petkowicz, C. L.; Nagashima, G. T.; Bespalhok Filho, J. C.; Vieira, L. G.; Pereira, L. F.; Domingues, D. S. Homeologous genes involved in mannitol synthesis reveal unequal contributions in response to abiotic stress in Coffea arabica. *Molecular genetics and genomics* **2014**, 289, 951-963. doi: 10.1007/s00438-014-0864-y

de Carvalho, K.; Bespalhok Filho, J. C.; dos Santos, T. B.; de Souza, S. G.; Vieira, L. G.; Pereira, L. F.; Domingues, D. S. Nitrogen starvation, salt and heat stress in coffee (Coffea arabica L.): identification and validation of new genes for qPCR normalization. *Molecular Biotechnology* **2013**, 53, 3, 315-25. doi: 10.1007/s12033-012-9529-4

Del Grossi, L., Sera, T., Sera, G. H., Fonseca, I., Ito, D. S., Shigueoka, L. H.; Andreazi E.; Carvalho, F. G. Rust resistance in arabic coffee cultivars in

Northern Paraná. Braz. *Brazilian Archives of Biology and Technology* **2013**, 56, 27–33. doi: 10.1590/S1516-89132013000100004

Edgar, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **2004**, *32*, 1792–1797. doi.org/10.1093/nar/gkh340

Fabbri, A. A.; Fanelli, C.; Reverberi, M.; Ricelli, A.; Camera, E.; Urbanelli, S.; Rossini, A.; Picardo, M.; Altamura, M.M. Early physiological and cytological events induced by wounding in potato tuber. *J. Exp. Bot* **2000**, 51, 1267–1275. doi.org/10.1093/jexbot/51.348.1267

Feussner, I.; Wasternack, C. The lipoxygenase pathway. *Annu. Rev. Plant Biol* **2002**, *53*, 275–297. doi: 10.1146/annurev.arplant.53.100301.135248

Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **1985**, *39*, 783–791. doi.org/10.2307/2408678

Finiti, I.; de la O Leyva, M.; Vicedo, B.; Gómez-Pastor R.; López-Cruz, J.; García-Agustín, P.; Real, M. D.; González-Bosch, C. Hexanoic acid protects tomato plants against Botrytis cinerea by priming defence responses and reducing oxidative stress. *Mol Plant Pathol* **2014**, 6, 550-62. doi: 10.1111/mpp.12112

Guo, S.; Song, Z.; Ma, R.; Yang, Y.; Yu, M. Genome-wide identification and expression analysis of the lipoxygenase gene family during peach fruit ripening under different postharvest treatments. Acta Physiol Plant 2017, 39, 111. doi.org/10.1007/s11738-017-2409-6

He, S.; Zhi, F.; Min, Y.; Ma, R.; Ge, A.; Wang, S.; Wang, J.; Liu, Z.; Guo, Y.; Chen, M. The MYB59 Transcription factor negatively regulates salicylic acid- and jasmonic acid-mediated leaf senescence. *Plant Physiol* **2022**, kiac589

Hou, S.; Lin, L.; Lv, Y.; Xu, N.; Sun, X. Responses of lipoxygenase, jasmonic acid, and salicylic acid to temperature and exogenous phytohormone treatments in Gracilariopsis lemaneiformis (Rhodophyta). J. *Appl. Phycol* **2018**, 30, 3387–3394. doi:10.1007/s10811-018-1514-2

Hu, B.; Jin, J.; Guo, A.Y.; Zhang, H.; Luo, J.; Gao, G. GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics* **2015**, *31*, 1296–1297. doi.org/10.1093/bioinformatics/btu817

Ivamoto, S.T.; Sakuray, L.M.; Ferreira, L.P.; Kitzberger, C.S.G.; Scholz, M.B.S.; Pot, D.; Leroy, T.; Vieira, L.G.E.; Domingues, D.S.; Pereira, L. F. P. Diterpenes biochemical profile and transcriptional analysis of cytochrome P450s genes in leaves, roots, flowers, and during *Coffea arabica* L. fruit development. *Plant Physiol Biochem* **2017**, *111*, 340-347. doi: 10.1016/j.plaphy.2016.12.004

Kosakivska, I.; Babenko, L.; Skaterna, T.; Ustinova, A. Influence of hypo- and hyperthermia on lipoxygenase activity, content of pigments and soluble

proteins in *Triticum aestivum* L. cv. Yatran 60 seedlings. *Genet. Plant Physiol.* **2014**, *46*, 212–220

Ksouri, N.; Jiménez, S.; Wells, C.E.; Contreras-Moreira, B.; Gogorcena, Y. Transcriptional Responses in Root and Leaf of Prunus persica under Drought Stress Using RNA Sequencing. *Front. Plant Sci* **2016**, 7, 1715. doi: 10.3389/fpls.2016.01715

Letunic, I.; Bork, P. Interactive tree of life (iTOL) v3: An online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* **2016**, 44, 242–245. doi.org/10.1093/nar/gkw290

Liu, F.; Li, H.; Wu, J.; Wang, B.; Tian, N.; Liu, J.; Sun, X.; Wu, H.; Huang, Y.; Lü, P.; Cheng, C. Genome-wide identification and expression pattern analysis of lipoxygenase gene family in banana. *Sci Rep* **2021**, 1, 9948. doi: 10.1038/s41598-021-89211-6

López-Galiano, M. J.; García-Robles, I.; Ruiz-Arroyo, V. M.; Oltra, S. S.; Marko Petek, M.; Rausell, C.; Real, M. D. Colorado potato beetle chymotrypsin genes are differentially regulated in larval midgut in response to the plant defense inducer hexanoic acid or the Bacillus thuringiensis Cry3Aa toxin. *Journal of Invertebrate Pathology* **2019**, 166, 107224. doi.org/10.1016/j.jip.2019.107224

MapGene2Chromosome v2. Disponível em: MapGene2Chrom web v2 (iask.in) Acesso em: 15/12/2022

Marques, I.; Rodrigues, A. P.; Gouveia, D.; Lidon, F. C.; Martins S.; Semedo, M. C.; Gaillard, J. C.; Pais, I. P.; Semedo, J. N.; Scotti, P. C.; Reboredo, F. H.; Partelli, F. L.; Da Matta, F. M.; Armengaud, J.; Barros, A. I. R.; Ramalho, J. C. High-resolution shotgun proteomics reveals that increased air [CO2] amplifies the acclimation response of *Coffea* species to drought regarding antioxidative, energy, sugar, and lipid dynamics, **2022**, *Journal of Plant Physiology* 276, 153788. doi.org/10.1016/j.jplph.2022.153788

Meriño-Cabrera, Y.; Zanuncio, J. C.; da Silva, R. S.; Solis-Vargas, M.; Cordeiro, G.; Rainha, F. R.; Campos, W. G.; Picanço, M. C.; de Almeida Oliveira, M. G. Biochemical response between insects and plants: an investigation of enzyme activity in the digestive system of Leucoptera coffeella (Lepidoptera: Lyonetiidae) and leaves of Coffea arabica (Rubiaceae) after herbivory. *Annals of Applied Biology* **2018**, 172, 236–243. doi:org/10.1111/aab.12416

Meriño-Cabrera, Y. Resposta de plantas de *Coffea arabica* (Rubiaceae) às injurias de Leucoptera coffeella (Lepidoptera: Lyonetiidae) através das vias das lipoxigenases e atividade de polifenoloxidases. **2015**. Tese (Doutorado) - Programa de Pós-graduação em Bioquímica Agrícola, Universidade Federal de Viçosa, Viçosa, 2015

Mukaka, M. M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Med J **2012**. 3, 69 -71. PMid:23638278

Sahu, S. S.; Loaiza, C. D.; Kaundal, R. Plant-mSubP: a computational framework for the prediction of single- and multi-target protein subcellular

localization using integrated machine-learning approaches. *AoB Plants* **2019**, 12 (3). doi: 10.1093/aobpla/plz068

Santino, A.; Iannacone, R.; Hughes, R.; Casey, R.; Mita, G. Cloning and characterisation of an almond 9-lipoxygenase expressed early during seed development. *Plant Sci.* **2005**, *168*, 699–706. doi.org/10.1016/j.plantsci.2004.10.001

Sattler, M. C., de Oliveira, S. C.; Mendonça, M. A. C.; Clarindo, W. R. Coffea cytogenetics: from the first karyotypes to the meeting with genomics. *Plants* **2022**, 6, 112. doi: 10.1007/s00425-022-03898-z

Scalschi, L.; Vicedo, B.; Camañes, G.; Fernandez-Crespo, E.; Lapeña, L.; González-Bosch, C.; García-Agustín, P. Hexanoic acid is a resistance inducer that protects tomato plants against Pseudomonas syringae by priming the jasmonic acid and salicylic acid pathways. *Mol Plant Pathol*, **2013**, 4, 342-55. doi: 10.1111/mpp.12010

Shaban, M.; Ahmed, M.M.; Sun, H.; Ullah, A.; Zhu, L. Genome-wide identification of lipoxygenase gene family in cotton and functional characterization in response to abiotic stresses. *BMC Genom*. **2018**, 19, 599. doi: 10.1186/s12864-018-4985-2

Singh, P.; Arif, Y.; Miszczuk, E.; Bajguz, A.; Hayat, S. Specific Roles of Lipoxygenases in Development and Responses to Stress in Plants. *Plants* **2022**, *11*, 979. doi.org/10.3390/plants11070979

Rivera-Contreras, I.K.; Zamora-Hernández, T.; Huerta-Heredia, A.A.; Capataz-Tafur, J.; Barrera-Figueroa, B.E.; Juntawong, P.; Peña-Castro, J.M. Transcriptomic analysis of submergence-tolerant and sensitive Brachypodium distachyon ecotypes reveals oxidative stress as a major tolerance factor. *Sci. Rep* **2016**, 6, 27686. doi: 10.1038/srep27686

Singh, P.; Arif, Y.; Miszczuk, E.; Bajguz, A.; Hayat, S. Specific Roles of Lipoxygenases in Development and Responses to Stress in Plants. *Plants* (Basel). **2022**, 7, 979. doi: 10.3390/plants11070979

Talhinhas, P., Batista, D., Diniz, I., Vieira, A., Silva, D.N., Loureiro, A., Tavares, S., Pereira, A.P., Azinheira, H.G., Guerra-Guimarães, L., Várzea, V. and Silva, M.d.C. The coffee leaf rust pathogen Hemileia vastatrix: one and a half centuries around the tropics. *Molecular Plant Pathology* **2017**, 18, 1039-1051. doi: https://doi.org/10.1111/mpp.12512

Tamura, K.; Stecher, G.; Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution* **2021**, 38, 3022 – 3027. doi.org/10.1093/molbev/msab120

Thumuluri, V.; Armenteros, J. J. A.; Johansen, A. R.; Nielsen, H.; Winther, O. DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. *Nucleic Acids Research* **2022**, 50, W228 – W234. doi.org/10.1093/nar/gkac278

Umate, P. Genome-wide analysis of lipoxygenase gene family in Arabidopsis and rice. *Plant Signal Behav* **2011**, *6*, 335–338. doi.org/10.4161/psb.6.3.13546

Viswanath, K.K.; Varakumar, P.; Pamuru, R.R.; Basha, S.J.; Rao, A.D. Plant Lipoxygenases and their role in plant physiology. *J. Plant Biol* **2020**, 63, 83–95. doi.org/10.1007/s12374-020-09241-x

Vogt, J.; Schiller, D.; Ulrich, D.; Schwab, W.; Dunemann, F. Identification of lipoxygenase (LOX) genes putatively involved in fruit flavour formation in apple (Malus × domestica). *Tree Genet. Genomes* **2013**, 9, 1493–1511. doi: 10.1007/s11295-013-0653-5

Wang, J.; Hu, T.; Wang, W.; Hu, H.; Wei, Q.; Wei, X.; Bao, C. Bioinformatics analysis of the lipoxygenase gene family in radish (Raphanus sativus) and functional characterization in response to abiotic and biotic stresses. *Int. J. Mol. Sci* **2019**, 20, 6095. doi:10.3390/ijms20236095

Zheng, K.; Wang, Z.; Pang, L.; Song, Z.; Zhao, H.; Wang, Y.; Wang, B.; Han, S. Systematic identification of methyl jasmonate-responsive long noncoding rnas and their nearby coding genes unveils their potential defence roles in tobacco BY-2 cells. Int. J. *Mol. Sci* **2022**, 23, 15568. doi: 10.3390/ijms232415568

Zhou, Y.; Wang, H.; Gilmer, S.; Whitwill, S.; Fowke, L.C. Effects of co-expressing the plant CDK inhibitor ICK1 and D-type cyclin genes on plant growth, cell size and ploidy in *Arabidopsis thaliana*. *Plant* **2003**, *216*, 604–613. doi: 10.1007/s00425-002-0935-x

Apêndice I
Artigo 1

frontiers | Frontiers in Genetics

Data Report
Published: 28 October 2022
doi: 10.3389/fgene.2022.935811

OPEN ACCESS

EDITED BY
Andrew H Paterson,
University of Georgia, United States

REVIEWED BY
Isabel Marques,
University of Lisbon, Portugal
Marcio Alves-Ferreira,
Federal University of Rio de Janeiro, Brazil

*CORRESPONDENCE
Douglas S. Domingues,
dougsd@usp.br

SPECIALTY SECTION
This article was submitted to Plant Genetics, a section of the journal Frontiers in Genetics

RECEIVED 21 April 2022
ACCEPTED 13 October 2022
PUBLISHED 28 October 2022

CITATION
Budzinski IGF, Camargo PO, Lemos SMC, Guyot R, Calzado NF, Ivamoto-Suzuki ST and Domingues DS (2022) Transcriptomic alterations in roots of two contrasting Coffea arabica cultivars after hexanoic acid priming. Front. Genet. 13:935811. doi: 10.3389/fgene.2022.935811

Transcriptomic alterations in roots of two contrasting *Coffea arabica* cultivars after hexanoic acid priming

Ilara G. F. Budzinski[1], Paula O. Camargo[2], Samara M. C. Lemos[1,2], Romain Guyot[3], Natália F. Calzado[1], Suzana T. Ivamoto-Suzuki[1] and Douglas S. Domingues[1,4]*

[1]Group of Genomics and Transcriptomics in Plants, Department of Biodiversity, Institute of Biosciences, São Paulo State University, UNESP, Rio Claro, Brazil, [2]Graduate Program in Biological Sciences (Genetics), Institute of Biosciences, São Paulo State University, UNESP, Botucatu, Brazil, [3]Institut de Recherche pour le Développement (IRD), Université Montpellier, Montpellier, France, [4]Department of Genetics, "Luiz de Queiroz" College of Agriculture, University of São Paulo, ESALQ/USP, Piracicaba, Brazil

KEYWORDS
RNA-seq, coffea, hexanoic acid, priming agent, elicitation, root

Introduction

Plants have the capacity to enter a state of alert that enables them to respond rapidly and robustly after exposure to stress (Aranega-Bou et al., 2014). This phenomenon is known as priming and can be described as an induced state whereby plants are pre-exposed to an inducing agent (elicitor), thus improving their perception and/or amplification of defense response-inducing signals (Aranega-Bou et al., 2014; Tugizimana et al., 2019). Hexanoic acid (Hx), a monocarboxylic acid, is a natural priming agent with proven efficiency in a wide range of host plants and pathogens (Llorens et al., 2016), including coffee pathogens. Coffee (*Coffea* spp.) is one of the most important agricultural commodities in the world. Brazil is the largest producer and exporter of *Coffea arabica* L. (Brazilian Coffee Exporters Council, 2021). The genus *Coffea* comprises 124 species (Davis et al., 2011). The most planted one is *C. arabica*, the only allotetraploid species in the genus. As many other plants, *Coffea* spp. are sensitive to a diverse range of biotic and abiotic stress. It is known that priming leads to changes at the transcriptional, physiological, metabolic and epigenetic levels (Bacilli et al., 2016). A transcriptional reprogramming may occur after priming stimulation, affecting a huge number of genes (Cervantes-Gámez et al., 2016; Bacilli et al., 2020). Within this context, our aim was to investigate the effect *per se* of Hx application. We hypothesize if Hx application could modulate genes related to defense response, in *C. arabica*, being a potential eliciting agent to this crop. To test this, Hx was applied in the roots of two Brazilian *C. arabica* cultivars: Catuaí Vermelho and Obatã. Cultivars were chosen based on their distinct breeding histories and contrasting resistance to rust, the major disease in Arabica coffee worldwide (Talhinhas et al., 2017). Catuaí Vermelho is susceptible to rust, and is one of the most planted cultivars in Brazil, while Obatã is described as a moderately resistant cultivar (Del Grossi et al., 2013). In the present work, transcriptomic analysis of roots were performed, revealing different molecular responses. Based on FPKM ratio and

statistical analysis, 1,545 differentially expressed genes (DEGs) were found. Functional annotation of DEGs through Blast2GO showed that primary, organic substance and cellular metabolic processes were mainly affected by priming, in both cultivars. Here, we present an RNA-seq dataset containing raw files and an initial exploration of differentially expressed genes in two *C. arabica* cultivars. Besides, these data could contribute to the identification of key genes differentially expressed in response to Hx.

Material and methods

Plant material

Plant material and experimental setup used in this work was the same described in a previous publication from our group (Rodrigues et al., 2021).

Two commercial cultivars of *C. arabica* (five-month-old plants) were used, Catuaí Vermelho IAC 144 and Obatã IAC 1669-20. Both cultivars are inbred lines of *C. arabica* (Maluf et al., 2005); however, Catuaí is derived from a cross between Catuaí Amarelo 476 × Mundo Novo 374-19, while Obatã is derived from interspecific crosses between (Villa Sarchi × Hybrid of Timor) × Catuaí Vermelho; clarifying that Villa Sarchi is a *C. arabica* cultivar and Hybrid of Timor is a natural *C. arabica* x *C. canephora* hybrid (Lashermes et al., 2000; Maluf et al., 2005). These cultivars were chosen due to their contrasting response to rust, with Obatã being the resistant one (Maluf et al., 2005; Krohling et al., 2018). Plants were selected based on size uniformity and were transferred to pots containing 3 L of aerated nutrient solution (ANS), adapted from (Clark, 1975) by de Carvalho et al. (2013). The experiment was carried out as described in Silva et al. (2020), under controlled temperature (25 ± 2°C) and light/dark cycle (12h/12h, photosynthetically active photon flux density of ~400 μmol $m^{-2}.s^{-1}$). The following treatments were applied: (a) ANS (control); (b) ANS + hexanoic acid (Merck, final concentration 0.55 mM) for 48 h. Three plants per pot were grown into six plastic pots in which three pots received each treatment. The experiments were repeated 3 times to obtain biological replicates. The potted plants were grouped in "pools" (made of 9–18 plants), which were considered a biological replicate. Three biological replicates were used. We collected plant secondary roots within the 3rd hour of the light period and stored at -80°C for further analyses.

Total RNA extraction and quality control

All steps from total RNA extraction until gene expression analysis were the same as described in Rodrigues et al. (2021).

Total RNA from root pools were isolated using the RNeasy Plant kit (Qiagen, Hilden, North Rhine-Westphalia, Germany). Total RNA samples were purified using the RNeasy MinElute Cleanup kit (Qiagen, Hilden, North Rhine-Westphalia, Germany). The purity of RNA was determined using a NanoDrop ND-300 spectrophotometer (Thermo Scientific, San Jose, CA, United States). RNA concentrations were measured by a Qubit fluorometer (Thermo Fisher Scientific, Wilmington, DE, United States).

Library preparation, and RNA-seq

Poly(A) RNA sequencing library was prepared following Illumina's TruSeq-stranded-mRNA sample preparation protocol (Illumina Technologies, SanDiego, CA). Paired-end sequencing (2 X 150 bp) was performed on Illumina's NovoSeq 6000 sequencing system at LC Sciences (Houston, TX, United States). Data was deposited into the European Nucleotide Archive (ENA), submission PRJEB52366.

RNAseq analysis and gene expression analysis

All steps mentioned here are the same as described in Rodrigues et al. (2021). Adaptor contamination, low quality bases and undetermined bases were removed by using Cutadapt (Martin, 2011) and in house PERL scripts. Sequence quality was verified using FastQC (Andrews, 2010). HISAT2 (Kim et al., 2015) was used to map reads to the *Coffea arabica* genome (ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/003/713/225/GCF_003713225.1_Cara_1.0/).

StringTie (Pertea et al., 2015) was used to assemble the mapped reads and to detect the expression level for mRNAs by calculating FPKM. The differentially expressed genes (DEGs) were selected with log2 (fold change) >1 or log2 (fold change) <-1 and with statistical significance (p value <0.05) by R package edgeR (Robinson et al., 2010). A second analysis was done on the differentially expressed mRNAs and only the ones with FPKM (ratio) ≥ 2 or FPKM (ratio) ≤ -2; coefficient of variation ≤30% and average FPKM ≥5 were selected for further analyses. Genes found specifically in one condition (control or plants exposed to Hx) were also described as DEGs.

Sequence annotation and gene ontology (GO) enrichment analysis of DEGs were performed using *Blast2GO* (Conesa et al., 2005), at the *BioMaas* (Götz et al., 2008) platform. Sequences were annotated by blasting nucleotide sequences against the NCBI NR database (BLASTX, evalue ≤1.10^{-3}). The hypergeometric distribution was used to test whether the GO function set was significantly enriched ($p < 0.05$). Pathway mapping was done using MapMan software (Thimm et al., 2004) with the *Arabidopsis thaliana* mapping file (http://mapman.gabipd.org/). TAIR IDs were retrieved from NCBI (https://www.ncbi.nlm.nih.gov).

10.3389/fgene.2022.925821

TABLE 1 Summary of sequencing data quality

Sample	Raw data	Valid data	Valid data gb (G)	Valid ratio (reads)	Q30%	GC content%
CC_root1	35,459,648	34,232,894	5.73	71.32	99.03	46
CC_root2	32,484,378	34,309,650	5.75	73.1	99.02	45.5
CC_root3	31,672,842	34,455,566	5.47	70.55	98.99	45
CIC_root1	42,811,370	37,029,650	5.55	88.13	97.6	45
CIC_root2	42,591,098	37,388,950	5.64	88.69	97.56	45
CIC_root3	41,533,681	36,290,882	5.44	87.58	97.56	45.5
CHX_root1	44,422,470	32,992,840	4.94	74.87	97.45	51
CHX_root2	51,270,376	50,082,294	7.5	97.53	97.84	51
CHX_root3	46,166,432	45,680,326	6.1	88.12	98.59	52
CHX_rad1	35,581,294	32,733,990	4.91	97.48	98.12	51
CHX_rad2	41,431,214	35,239,206	5.29	85.1	98.57	51
CHX_rad3	32,863,356	31,991,596	4.8	97.03	98.18	52

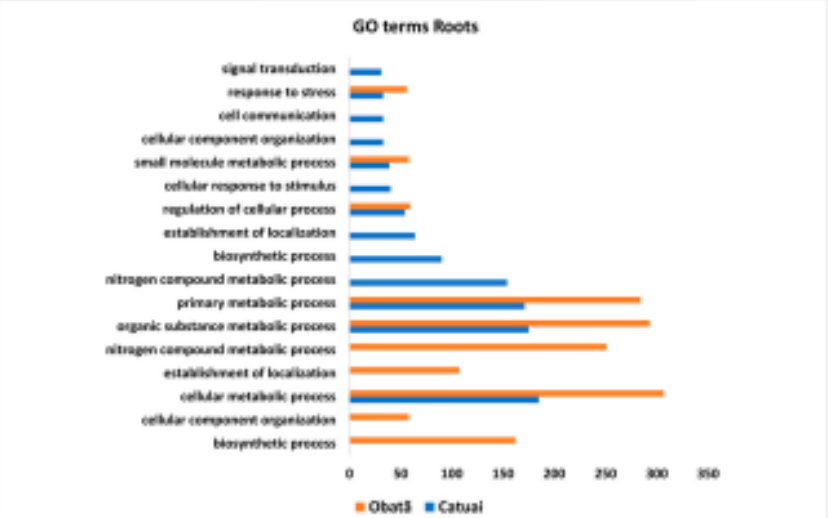


FIGURE 1
Gene ontology (GO) functional enrichment analysis of DEGs from C. arabica Catuai and Obatã cultivars.

Overall data annotation, differentially expressed genes and gene ontology analysis

Quality control and mapping information are available in Table 1. About 67.12 Gb total clean bases were obtained by RNA-seq after quality check, with an average of 5.6 Gb for each sample. The lowest value of Q30 (percentage of bases with sequencing error rate lower than 1%) was 97.36%. The GC content ranged from 45 to 52%.

As a preliminary analysis to identify genes and functional categories potentially modulated by Hx application, the first step of our work was to identify the DEGs based on FPKM and statistical analysis. Based on FPKM ratio and statistical analysis, 1,545 DEGs were found in total, 557 and 988 in Catuaí and Obatã, respectively (Supplementary Table S1). From these, 157 DEGs were found in both cultivars, while 400 and 831 DEGs were specifically found in Catuaí and Obatã cultivars, respectively (Supplementary Tables S2, S3). We hypothesize that the discrepancy between the number of specific DGEs, found in each cultivar, is related to differences in rust resistance, reinforcing that molecular mechanisms of defense are differentially recruited depending on cultivar tolerance. Most of the DEGs have a role in plant defense, indicating the modulation of this mechanism in roots by priming. Blast2GO analysis showed that primary, organic substance and cellular metabolic processes were mainly affected by priming, followed by response to stress, small molecule metabolic process and regulation of cellular process (Figure 1, Supplementary Table S5). Pathway analysis of DEGs using MapMan showed differences in the activity of cellular metabolisms due to Hx (Supplementary Table S3). The dataset presented here indicates that hexanoic acid modulates plant defense mechanisms in *C. arabica*. Moreover, we are providing useful data for further investigations on *C. arabica* root responses to Hx.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ebi.ac.uk/ena, PRJEB52366. All supplementary files are available on https://doi.org/10.5281/zenodo.6467813.

Author contributions

Conceptualization, Project Administration, Funding Acquisition, Supervision: DD. Data Curation, Investigation: PC, SL, RG, NC, STI-S. Formal Analysis, Validation, Visualization: IB. Writing—Original Draft Preparation, Writing—Review and Editing: IB, DD.

Funding

This research was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), grant number #2016/10896-0 and CAPES-PrInt Program 2346/2018 (process 88881.310767/2018-01). IB, NC and SL were financed in part by the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES)—Finance Code 001.

Acknowledgments

IB acknowledges the scholarship granted from the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES), in the scope of the Program CAPES-PrInt, process number 88887.310463/2018-00, International Cooperation Project number 88887.512173/2020-00. SL also acknowledges a CAPES fellowship, process number 88887.570128/2020-00. STI-S. acknowledges FAPESP for providing a post-doctoral fellowship, process number #2017/01455-2. DD also acknowledges CNPq for a research productivity fellowship (process number #312823/2019-3).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.925811/full#supplementary-material

References

Artigo 2

frontiers in Sustainable Food Systems

DATA REPORT
published: 25 October 2021
doi: 10.3389/fsufs.2021.735660

Transcriptome Analyses of Leaves Reveal That Hexanoic Acid Priming Differentially Regulate Gene Expression in Contrasting *Coffea arabica* Cultivars

Ilara G. F. Budzinski[1], Paula O. Camargo[1], Raissa S. Rosa[1], Natália F. Calzado[1], Suzana T. Ivamoto-Suzuki[1,2] and Douglas S. Domingues[1]*

[1] Group of Genomics and Transcriptomes in Plants, Department of Biodiversity, Institute of Biosciences, UNESP, São Paulo State University, Rio Claro, Brazil, [2] Department of Agronomy, UEL, State University of Londrina, Londrina, Brazil

OPEN ACCESS

Keywords: RNA-seq, hexanoic acid, priming, abiotic stress, Coffea

Edited by:
Pushp Sheel Shukla,
Dalhousie University, Canada

Reviewed by:
Rakesh Upadhyay,
United States Department of Agriculture (USDA), United States
Marie-Christine Combes,
Institut de recherche pour le Développement, France

***Correspondence:**
Douglas S. Domingues
douglas.domingues@unesp.br

Specialty section:
This article was submitted to Crop Biology and Sustainability, a section of the journal Frontiers in Sustainable Food Systems

Received: 03 July 2021
Accepted: 27 September 2021
Published: 25 October 2021

Citation:
Budzinski IGF, Camargo PO, Rosa RS, Calzado NF, Ivamoto-Suzuki ST and Domingues DS (2021) Transcriptome Analyses of Leaves Reveal That Hexanoic Acid Priming Differentially Regulate Gene Expression in Contrasting *Coffea arabica* Cultivars. Front. Sustain. Food Syst. 5:735660. doi: 10.3389/fsufs.2021.735660

INTRODUCTION

Coffee (*Coffea* spp.) is one of the most important traded commodities in the international market [International Coffee Organization (ICO), 2018]. *C. arabica*, the only allotetraploid species in the genus, is the most planted one. Despite its economic importance, evaluation of physiological changes taking account molecular responses to biostimulants are still scarce. When plants recognize potential biotic/abiotic challenges, they often switch to a primed state of enhanced defense. This mechanism enables plants to respond robustly after exposure to stress (Aranega-Bou et al., 2014). The response of *C. arabica* plants to priming, in terms of transcriptional profiles, is a big gap in this area. Hexanoic acid (Hx) is a natural priming agent with proven efficiency in a wide range of host plants and pathogens (Llorens et al., 2016), including coffee pathogens. In this study we aimed to investigate the effect of Hx priming in *C. arabica* leaves transcriptome. We hypothesize if Hx application could modulate genes related to defense responses, being a potential eliciting agent in *C. arabica*. To test this effect, we applied Hx in roots of two Brazilian *C. arabica* cultivars with distinct breeding histories and contrasting resistance to the major disease in Arabica coffee, coffee leaf rust. While Catuaí Vermelho is among the most used cultivars in Brazil, but it is susceptible to leaf rust, Obatã is a moderately resistant cultivar (Del Grossi et al., 2013). We performed transcriptome analysis of leaves. Reads were mapped to the *C. arabica* public genome and up to 94% of reads were mapped. Transcript expression level was quantified and differentially expressed genes (DEGs) were identified based on FPKM ratio and statistical analyses. A total of 57 and 63 DEGs were found in Catuaí Vermelho and Obatã, respectively. Most DEGs correspond to upregulated genes in response to Hx, in both cultivars (86% Catuaí Vermelho and 73% Obatã). Eight DEGs were found modulated in both cultivars, including ferredoxin-NADP reductase and phenylalanine ammonia-lyase. DEGs were functionally annotated through Blast2GO. Biological process and molecular function categorizations revealed that DEGs related to cellular, oxidation-reduction, organic substance and primary metabolic process, as well as transferase and ion binding activity might play a dominant role in the leaf response to priming. These data contribute to identify key genes differentially expressed in response to Hx as well as indicate pathways modulated by this eliciting agent.

VALUE OF DATA

- The species *C. arabica* is responsible for 60% of the world coffee production due to its fine flavor and aroma. Brazil is the largest producer and exporter of this commodity.
- Hexanoic acid (Hx) is a priming agent which leads plant to a physiological state that enables them to respond more rapidly and/or robustly to biotic or abiotic stress scenarios.
- To date, there is no transcriptome analysis for *C. arabica* cultivars Catuaí Vermelho and Obatã in response to the priming agent hexanoic acid (Hx) application.
- These data will contribute to identify key genes differentially expressed in response to Hx application and it can also indicates which pathways are modulated by Hx.

MATERIALS AND METHODS

Plant Material

Five-month-old plants of *C. arabica* cv. Catuaí Vermelho IAC 144 and Obatã IAC 1669-20 (4–5 leaf pairs) were used in this study. Plants were selected based on size uniformity and were transferred to pots containing 3 L of aerated nutrient solution (ANS), adapted from Clark (1975) by de Carvalho et al. (2013). The hexanoic acid experiment was carried out as described in Silva et al. (2020), under controlled temperature (23 ± 2°C) and light/dark cycle (12/12 h, photosynthetically active photon flux density of ~400 μmol.m^{-2}.s^{-1}). The following treatments were assessed: (a) ANS (control); (b) ANS + hexanoic acid (Merck, final concentration 0.55 mM) for 48 h. Plants were grown in 3 to 6 plastic pots in which three pots received each treatment. The experiments were repeated 3 times to obtain biological replicates. The potted plants were grouped in "pools" (made of 9–18 plants), which were considered a biological replicate. Three biological replicates were used. The mature leaves of the middle third of the plants were collected within the 3rd hour of the light period and stored at −80°C to further analyses.

Total RNA Extraction and Quality Control, Library Preparation, and RNA-Seq

Total RNA was extracted from leaves pools using the RNeasy Plant kit (Qiagen, Hilden, North Rhine-Westphalia, Germany). Total RNA samples were purified using the RNeasy MinElute Cleanup kit (Qiagen, Hilden, North Rhine-Westphalia, Germany). The purity of RNA was determined using a NanoDrop ND-100 spectrophotometer (Thermo Scientific, San Jose, CA, USA) and concentrations were obtained using Qubit fluorimeter (Thermo Fisher Scientific, Wilmington, DE, USA). Poly(A) RNA sequencing library was prepared following Illumina's TruSeq-stranded-mRNA sample preparation protocol (Illumina Technologies, SanDiego, CA). Paired-ended sequencing (2 X 150 bp) was performed on Illumina's NovaSeq 6000 sequencing system at LC Sciences (Houston, TX, USA). Data were deposited into the European Nucleotide Archive (ENA), submission ERA6282544 and in Zenodo (doi: 10.5281/zenodo.5517785).

TABLE 1 | Statistics of *C. arabica* cv. Catuaí Vermelho and Obatã RNA-seq analysis.

Cultivar/Treatment	Raw reads	Valid reads	
Pre-Assembly			
C. arabica cv. Catuaí Control	151,329,548	118,851,828	
C. arabica cv. Catuaí Hx	150,026,378	119,069,580	
C. arabica cv. Obatã Control	160,857,610	129,342,302	
C. arabica cv. Obatã Hx	162,917,402	124,809,996	
Cultivar/Treatment	**Mapped reads**	**Uniquely mapped reads**	**Spliced reads**
Post-Assembly			
C. arabica cv. Catuaí Control	109,447,764	88,080,059	30,007,564
C. arabica cv. Catuaí Hx	111,251,338	70,544,331	31,056,743
C. arabica cv. Obatã Control	114,541,661	64,017,674	30,597,588
C. arabica cv. Obatã Hx	115,845,730	62,705,577	31,583,481

RNA-Seq Analysis and Differential Transcript Abundance

Adaptor contamination, low quality bases and undetermined bases were removed by using Cutadapt (Martin, 2011) and in house PERL scripts. Sequence quality was verified using FastQC (Andrews, 2010). HISAT2 (Kim et al., 2015) was used to map reads to the *C. arabica* genome (ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/003/713/225/GCF_003713225.1_Cara_1.0/Mapped) reads were assembled using StringTie (Pertea et al., 2015). **Table 1** shows the statistics of the transcriptome analysis. StringTie was also used to detect expression level for mRNAs by calculating FPKM. The differentially expressed genes (DEGs) were selected with log2 (fold change) >1 or log2 (fold change) <-1 and with statistical significance (p value < 0.05) by R package edgeR (Robinson et al., 2010). A second analysis was done on the differentially expressed mRNAs and only the ones with FPKM (ratio) $\geq$ 2 or FPKM (ratio) $\leq$ −2; coefficient of variation $\leq$ 30% and average FPKM $\geq$ 5. These genes are reported in **Supplementary Table 1**.

DEGs Data Annotation and Gene Ontology (GO) Analysis

Sequence annotation and functional analysis were done for the DEGs by using *Blast2GO* (Conesa et al., 2005), at the OmicsBox (Götz et al., 2008) platform. Sequences were annotated by blasting nucleotide sequences against the NCBI

NR database (BLASTX, evalue ≤ 1.10^{-5}). A total of 57 and 63 DEGs were found in Catuaí Vermelho and Obatã, respectively. Most DEGs correspond to upregulated genes in response to Hx, in both cultivars (86% Catuaí Vermelho and 73% Obatã). Eight DEGs were commonly found in the Catuaí Vermelho and Obatã cultivars (AAA-ATPase, auxin-induced protein (two transcripts), ferredoxin-NADP reductase, premnaspirodiene oxygenase-like, glutathionyl-hydroquinone reductase, flavonol 3-O-glucosyltransferase and phenylalanine ammonia-lyase). DEGs were analyzed by Gene Ontology and were functionally assigned to the relevant terms. Biological process and molecular function categorizations revealed that DEGs related to cellular, oxidation-reduction, organic substance, and primary metabolic process, as well as transferase and ion binding activity might play a dominant role in the leaf response to priming. Most of the DEGs have a role in plant defense, corroborating to our hypothesis. These data contribute to identify key genes differentially expressed in response to Hx as well as indicate pathways modulated by this eliciting agent.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ebi.ac.uk/ena, PRJEB39901, https://doi.org/10.5281/zenodo.5517786.

AUTHOR CONTRIBUTIONS

DD conceived the idea and acquired funding. PC, RR, NC, and SI-S conducted the experiment. IB performed analysis on the data. IB and DD wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), grant number #2016/10896-0. IB and NC were financed in part by the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES) – Finance Code 001. IB acknowledge the scholarship granted from the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES), in the scope of the Program CAPES-PrInt, process number 88887.310463/2018-00, International Cooperation Project number 88887.512173/2020-00. SI-S acknowledges FAPESP for providing a post-doctoral fellowship, process number #2017/01455-2. DD also acknowledges CNPq for a research productivity fellowship (process number #312823/2019-3).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fsufs.2021.735893/full#supplementary-material

REFERENCES

Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data*. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc (accessed July 1, 2021).

Aranega-Bou, P., Leyva, M., Finiti, I., García-Agustín, P., and González-Bosch, C. (2014). Priming of plant resistance by natural compounds. Hexanoic acid as a model. *Front. Plant Sci.* 5:488. doi: 10.3389/fpls.2014.00488

Clark, R. B. (1975). Differential magnesium efficiency in corn inbreds: I. dry-matter yields and mineral element composition. *Soil Sci. Soc. Am. J.* 39, 488–491. doi: 10.2136/sssaj1975.03615995003900030032x

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610

de Carvalho, K., Bespalhok Filho, J. C., Dos Santos, T. B., de Souza, S. G. H., Vieira, L. G. E., Pereira, L. F. P., et al. (2013). Nitrogen starvation, salt and heat stress in coffee (*Coffea arabica* L.): identification and validation of new genes for qPCR normalization. *Mol. Biotechnol.* 53, 315–325. doi: 10.1007/s12033-012-9529-4

Del Grossi, L., Sera, T., Sera, G. H., Fonseca, I., Ito, D. S., Shigueoka, L. H., et al. (2013). Rust resistance in arabic coffee cultivars in Northern Paraná. *Braz. Arch. Biol. Technol.* 56, 27–33. doi: 10.1590/S1516-89132013000100004

Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., et al. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36, 3420–3435. doi: 10.1093/nar/gkn176

International Coffee Organization (ICO) (2019). *Statistics of Production*. Available online at: http://www.ico.org/ (accessed July 1, 2021).

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Llorens, E., Camañes, G., Lapeña, L., and García-Agustín, P. (2016). Priming by hexanoic acid induce activation of mevalonic and linolenic pathways and promotes the emission of plant volatiles. *Front. Plant Sci.* 7:495. doi: 10.3389/fpls.2016.00495

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* 17, 10–12. doi: 10.14806/ej.17.1.200

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Robinson, M., McCarthy, D., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Silva, N., Ivamoto-Suzuki, S. T., Camargo, P. O., Rosa, R. S., Pereira, L. F. P., and Domingues, D. S. (2020). Low-copy genes in terpenoid metabolism: the evolution and expression of MVK and DXR genes in angiosperms. *Plants (Basel)* 9:525. doi: 10.3390/plants9040525

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Apêndice II

MATERIAL SUPLEMENTAR (ARTIGO 1)

Table S1: Lipoxygenase genes in angiosperms.

Species	Gene code	Number of nucleotides	Number of amino acids	Site	Number of gene
	AT1G55020	2.580	859	Plaza	
	AT1G17420	2.750	919	Plaza	
Arabidopsis thaliana	AT3G45140	2.691	896	Plaza	
	AT1G72520	2.781	926	Plaza	6
	AT1G67560	2.754	917	Plaza	
	AT3G22400	2.661	886	Plaza	
	CS00075G00190	2.562	853	Plaza	
	CS00101G00510	2.694	897	Plaza	
	CS00101G00520	2.673	890	Plaza	
	CS00101G00530	2.604	867	Softberry	
Citrus sinensis	CS00103G00550	2.625	874	Plaza	
	CS00033G00970	2.703	900	Plaza	10
	CS00033G00990	2.685	894	Plaza	
	CS00011G01030	2.649	882	Plaza	
	CS00003G00910	2.781	926	Plaza	
	CS00033G01010	2.685	894	Plaza	
	Bradi5g11590	2.748	915	Plaza	
	Bradi1g11680	2.583	860	Plaza	
	Bradi1g11670	2.640	879	Plaza	
	Bradi3g39980	2.805	934	Plaza	
Brachypodium distachyon	Bradi1g09270	2.592	863	Plaza	10
	Bradi1g09260	2.595	864	Plaza	
	Bradi3g59710	2.538	845	Plaza	
	Bradi3g07000	2.694	897	Plaza	
	Bradi3g07010	2.688	895	Plaza	
	Bradi1g72690	2.763	920	Plaza	
	Carubv10019745m.g	2.763	920	Plaza	
	Carubv10019783m.g	2.580	859	Plaza	
	Carubv10019743m.g	2.778	925	Plaza	7
Capsella rubella	Carubv10012957m.g	2.610	869	Plaza	
	Carubv10016667m.g	2.703	900	Plaza	
	Carubv10016666m.g	2.709	902	Plaza	
	Carubv10008244m.g	2.766	921	Plaza	

	Cc02_g13400	2.817	938	Plaza	
	Cc11_g16680	2.820	939	Plaza	
Coffea canephora	Cc03_g03580	2.649	798	Plaza	7
	Cc02_g33320	2.571	856	Plaza	
	Cc02_g33780	2.574	857	Plaza	
	Cc01_g04060	2.514	837	Plaza	
	Cc02_g33800	2.400	799	Softberry	
	Gorai.006G238200	2.703	900	Plaza	
	Gorai.004G059500	2.727	908	Plaza	
	Gorai.004G059900	2.916	971	Plaza	
	Gorai.004G241400	2.580	859	Plaza	
	Gorai.009G071200	2.715	904	Plaza	
	Gorai.011G060400	2.760	919	Plaza	
Gossypium raimondii	Gorai.004G060100	2.646	881	Plaza	
	Gorai.001G228100	2.559	852	Plaza	
	Gorai.004G059700	3.000	999	Plaza	17
	Gorai.005G170200	2.739	912	Plaza	
	Gorai.013G124500	2.598	865	Plaza	
	Gorai.006G087200	2.736	911	Plaza	
	Gorai.010G255700	2.661	886	Plaza	
	Gorai.010G255600	2.601	866	Plaza	
	Gorai.005G041900	2.634	877	Plaza	
	Gorai.004G241100	2.475	824	Plaza	
	Gorai.004G241300	2.580	859	Plaza	
	Prupe.3G039200	2.739	912	Plaza	
	Prupe.8G189000	2.520	839	Plaza	
	Prupe.6G324400	2.589	862	Plaza	
	Prupe.1G011400	2.646	881	Plaza	
	Prupe.2G005800	2.718	905	Plaza	
Prunus persica	Prupe.4G047800	2.721	906	Plaza	
	Prupe.6G324600	2.751	916	Plaza	14
	Prupe.2G005300	2.754	917	Plaza	
	Prupe.2G005500	2.763	920	Plaza	
	Prupe.1G232400	2.763	920	Plaza	
	Prupe.6G324100	2.772	923	Plaza	
	Prupe.6G324300	2.802	933	Plaza	
	Prupe.6G324200	2.841	946	Plaza	
	Prupe.6G018700	2.844	947	Plaza	
	Eucgr.H04662	2.787	928	Plaza	
	Eucgr.H04499	2.616	871	Plaza	
	Eucgr.H00713	2.352	783	Plaza	

	Eucgr.H04496	2.616	871	Plaza	
	Eucgr.H03535	2.699	899	Plaza	16
	Eucgr.H04498	2.616	871	Plaza	
Eucalyptus grandis	Eucgr.J02807	2.736	911	Plaza	
	Eucgr.J00832	2.592	863	Plaza	
	Eucgr.H04495	2.595	864	Plaza	
	Eucgr.H04500	2.634	877	Plaza	
	Eucgr.J00820	2.712	903	Plaza	
	Eucgr.J00824	2.709	902	Plaza	
	Eucgr.L01891	2.709	902	Plaza	
	Eucgr.J00825	2.709	902	Plaza	
	Eucgr.H04497	2.424	807	Softberry	
	Eucgr.J00819	2.709	902	Softberry	
	RCO.g.30152.000070	2.709	902	Plaza	
	RCO.g.29851.000064	2.514	837	Plaza	
	RCO.g.30128.000242	2.616	871	Plaza	
	RCO.g.29808.000009	2.496	831	Plaza	
Ricinus communis	RCO.g.29726.000003	2.607	868	Plaza	
	RCO.g.29680.000053	2.598	865	Plaza	11
	RCO.g.29929.000202	2.739	912	Plaza	
	RCO.g.29851.000066	2.361	786	Plaza	
	RCO.g.30169.000166	2.703	899	Plaza	
	RCO.g.30178.000021	2.559	852	Plaza	
	RCO.g.30169.000164	2.370	789	Plaza	
	Seita.3G294500	2.544	847	Plaza	
	Seita.9G127800	2.595	864	Plaza	
	Seita.9G270500	2.694	897	Plaza	
Setaria italica	Seita.5G411600	2.604	867	Plaza	
	Seita.4G215400	2.763	920	Plaza	12
	Seita.9G127600	2.664	887	Plaza	
	Seita.9G127700	2.664	887	Plaza	
	Seita.9G518800	2.754	917	Plaza	
	Seita.5G411700	2.595	864	Plaza	
	Seita.6G205300	2.922	973	Plaza	
	Seita.1G050700	2.685	894	Plaza	
	Seita.7G113700	2.769	922	Plaza	
	Potri.014G177200	2.583	860	Plaza	
	Potri.001G015600	2.697	898	Plaza	
	Potri.017G046200	2.697	898	Plaza	
	Potri.013G022100	2.586	861	Plaza	
	Potri.001G015500	2.697	898	Plaza	

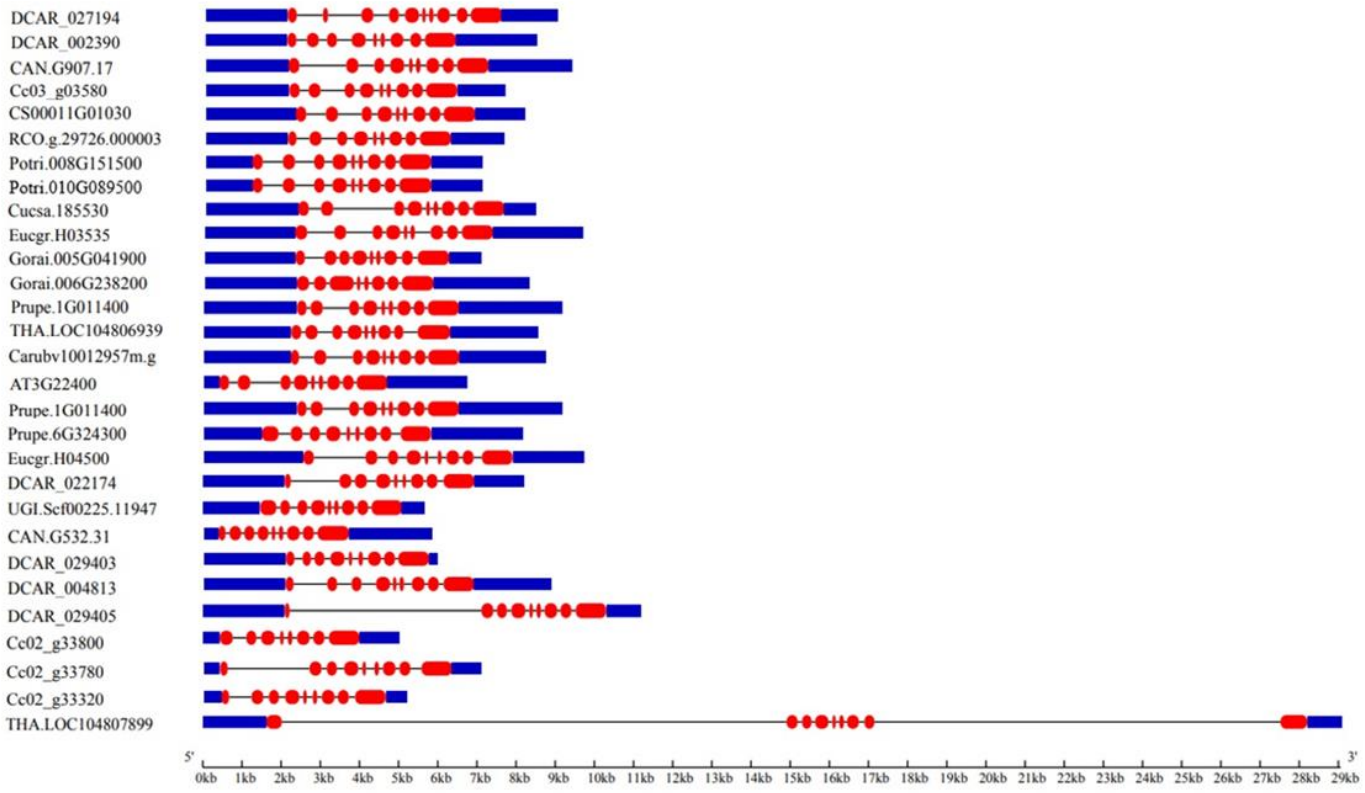
	Potri.005G032600	2.592	863	Plaza	
	Potri.005G032700	2.592	863	Plaza	
	Potri.005G032400	2.601	866	Plaza	
	Potri.005G032800	2.592	863	Plaza	20
Populus trichocarpa	Potri.009G022400	2.706	901	Plaza	
	Potri.001G015400	2.709	902	Plaza	
	Potri.001G015300	2.598	865	Plaza	
	Potri.010G089500	2.712	903	Plaza	
	Potri.008G178000	2.781	926	Plaza	
	Potri.010G057100	2.781	926	Plaza	
	Potri.008G151500	2.643	880	Softberry	
	Potri.001G167700	2.772	923	Plaza	
	Potri.013G022000	2.616	871	Plaza	
	Potri.014G018200	2.583	860	Plaza	
	Potri.003G067600	2.778	925	Plaza	
	ATR0743G022	2.496	831	Plaza	
	ATR0824G023	2.739	912	Plaza	4
	ATR0605G039	2.715	904	Plaza	
Amborella trichopoda	ATR0605G100	2.721	906	Plaza	
	LOC_Os03g49260	2.592	863	Plaza	
	LOC_Os04g37430	2.394	797	Plaza	
	LOC_Os02g10120	2.781	926	Plaza	
	LOC_Os12g37260	2.769	922	Plaza	
Oryza sativa ssp. japonica	LOC_Os08g39840	2.775	924	Plaza	
	LOC_Os05g23880	2.544	847	Plaza	11
	LOC_Os03g49380	2.634	877	Plaza	
	LOC_Os08g39850	2.826	941	Plaza	
	LOC_Os03g52860	2.613	870	Plaza	
	LOC_Os11g36719	2.607	868	Plaza	
	LOC_Os03g08220	2.757	918	Plaza	
	Ma02_g07800	2.562	853	Plaza	
	Ma01_g18060	2.586	861	Plaza	
	Ma06_g26890	2.622	873	Plaza	
	Ma06_g26850	2.622	873	Plaza	
	Ma03_g07770	2.745	914	Plaza	
Musa acuminata	Ma01_g18040	2.544	847	Plaza	
	Ma10_g17560	2.760	919	Plaza	
	Ma09_g12090	2.712	903	Plaza	17
	Ma06_g30170	2.739	912	Plaza	
	Ma09_g15420	2.733	910	Plaza	

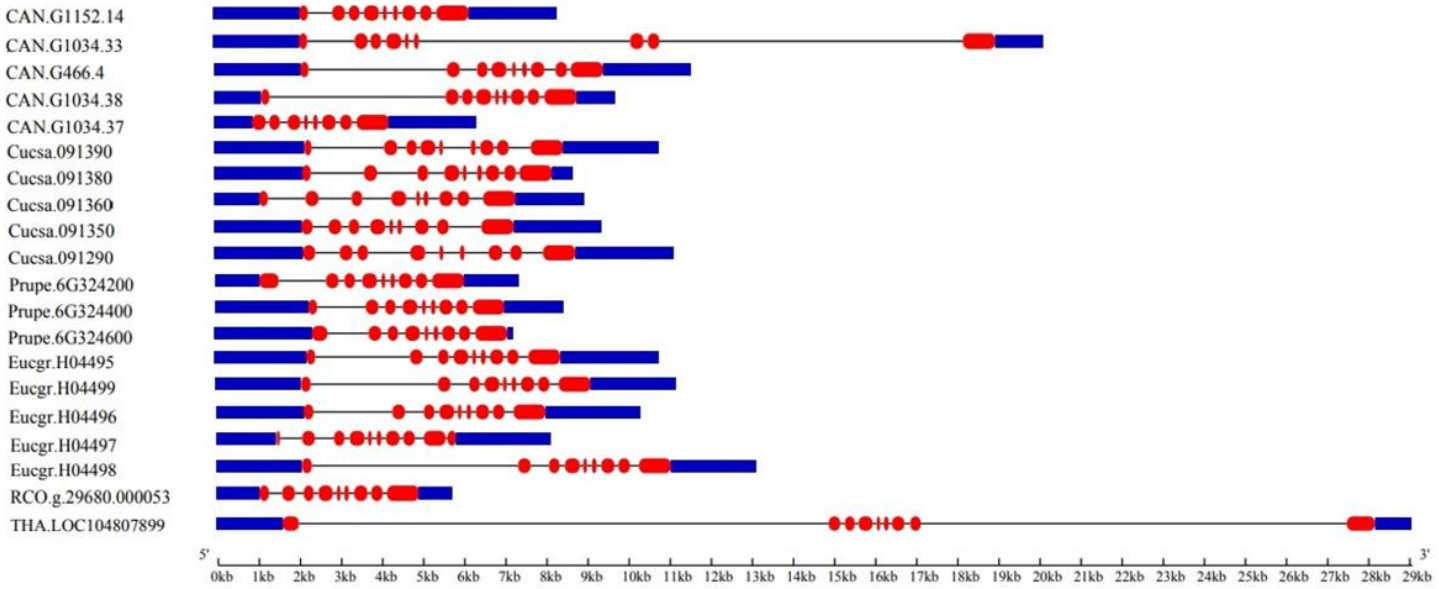
	Ma03_g11520	2.724	907	Plaza	
	Ma01_g16400	2.619	872	Plaza	
	Ma09_g19130	2.856	951	Plaza	
	Ma01_g18020	2.544	847	Plaza	
	Ma08_g23400	2.736	911	Plaza	
	Ma06_g26840	2.616	871	Softberry	
	Ma06_g26870	2.736	911	Softberry	
	Sobic.003G385500	2.607	868	Plaza	
	Sobic.007G210400	2.598	865	Plaza	
	Sobic.001G125900	2.601	866	Plaza	
	Sobic.004G078600	2.754	917	Plaza	
	Sobic.001G125800	2.667	888	Plaza	
	Sobic.006G248300	2.724	907	Plaza	11
Sorghum bicolor	Sobic.001G125700	2.487	828	Plaza	
	Sobic.003G385900	2.622	873	Plaza	
	Sobic.008G191000	2.574	857	Plaza	
	Sobic.006G095600	2.775	924	Plaza	
	Sobic.001G483400	2.775	924	Plaza	
	Cucsa.185530	2.646	881	Plaza	
	Cucsa.065250	2.469	822	Plaza	
	Cucsa.360190	2.724	907	Plaza	
	Cucsa.153620	2.493	830	Plaza	
	Cucsa.065240	2.562	853	Plaza	
	Cucsa.065260	2.709	902	Plaza	
	Cucsa.091390	2.565	854	Plaza	17
Cucumis sativus	Cucsa.153660	2.493	830	Plaza	
	Cucsa.153610	2.496	831	Plaza	
	Cucsa.153630	2.703	900	Plaza	
	Cucsa.153650	2.457	818	Plaza	
	Cucsa.091290	2.673	890	Plaza	
	Cucsa.091360	2.583	860	Plaza	
	Cucsa.153640	2.442	813	Softberry	
	Cucsa.091350	2.637	878	Plaza	
	Cucsa.075520	2.787	928	Plaza	
	Cucsa.091380	2.595	864	Plaza	
	CAN.G1034.37	2.370	789	Softberry	
	CAN.G1034.33	2.580	859	Plaza	
	CAN.G1152.14	2.586	861	Plaza	
	CAN.G907.17	2.661	886	Plaza	10
	CAN.G466.4	2.580	859	Plaza	
Capsicum annuum	CAN.G1931.1	2.748	915	Plaza	

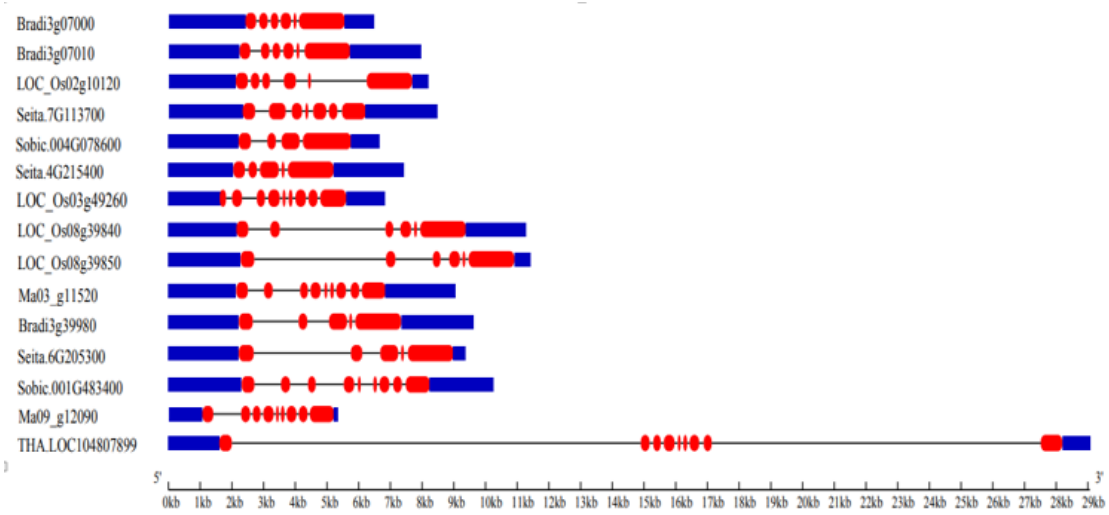
	CAN.G862.55	2.730	909	Plaza	
	CAN.G649.19	2.697	897	Plaza	
	CAN.G1034.38	2.589	862	Plaza	
	CAN.G532.31 (artigo)	2.463	820	Softberry	
Utricularia gibba	UGI.Scf00006.1017	2.769	892	Plaza	2
	UGI.Scf00225.11947	2.724	907	Plaza	
	DCAR_029405	2.514	837	Plaza	
	DCAR_029403	2.511	836	Plaza	
	DCAR_007013	2.874	957	Plaza	
	DCAR_007649	2.574	857	Plaza	
	DCAR_001406	2.661	886	Plaza	
Daucus carota	DCAR_024393	2.706	901	Plaza	
	DCAR_002390	2.616	871	Plaza	13
	DCAR_004813	2.535	844	Plaza	
	DCAR_027194	2.742	913	Plaza	
	DCAR_022174	2.562	853	Plaza	
	DCAR_016115	2.742	913	Plaza	
	DCAR_024394	2.721	906	Plaza	
	DCAR_026417	2.724	907	Plaza	
	PAB00043353	2.571	856	Plaza	
	PAB00025444	2.751	916	Plaza	
	PAB00015721	2.373	790	Softberry	
Picea abies	PAB00007111	2.397	798	Plaza	
	PAB00009052	2.613	870	Plaza	8
	PAB00029662	2.778	925	Plaza	
	PAB00038246	2.793	930	Plaza	
	PAB00025825	2.715	904	Plaza	
	Mapoly0063s0050	2.907	968	Plaza	
	Mapoly0026s0148	2.877	958	Plaza	
	Mapoly0055s0019	2.922	973	Plaza	
	Mapoly0420s0001	2.874	957	Plaza	
	Mapoly0001s0529	2.838	945	Plaza	
	Mapoly0079s0056	2.682	893	Plaza	
	Mapoly0063s0014	2.868	955	Plaza	16
Marchantia polymorpha	Mapoly0006s0130	2.868	955	Plaza	
	Mapoly0028s0085	2.868	955	Plaza	
	Mapoly0083s0028	2.868	955	Plaza	
	Mapoly0064s0034	2.679	892	Plaza	
	Mapoly0106s0031	2.862	953	Plaza	
	Mapoly0026s0145	2.754	917	Plaza	

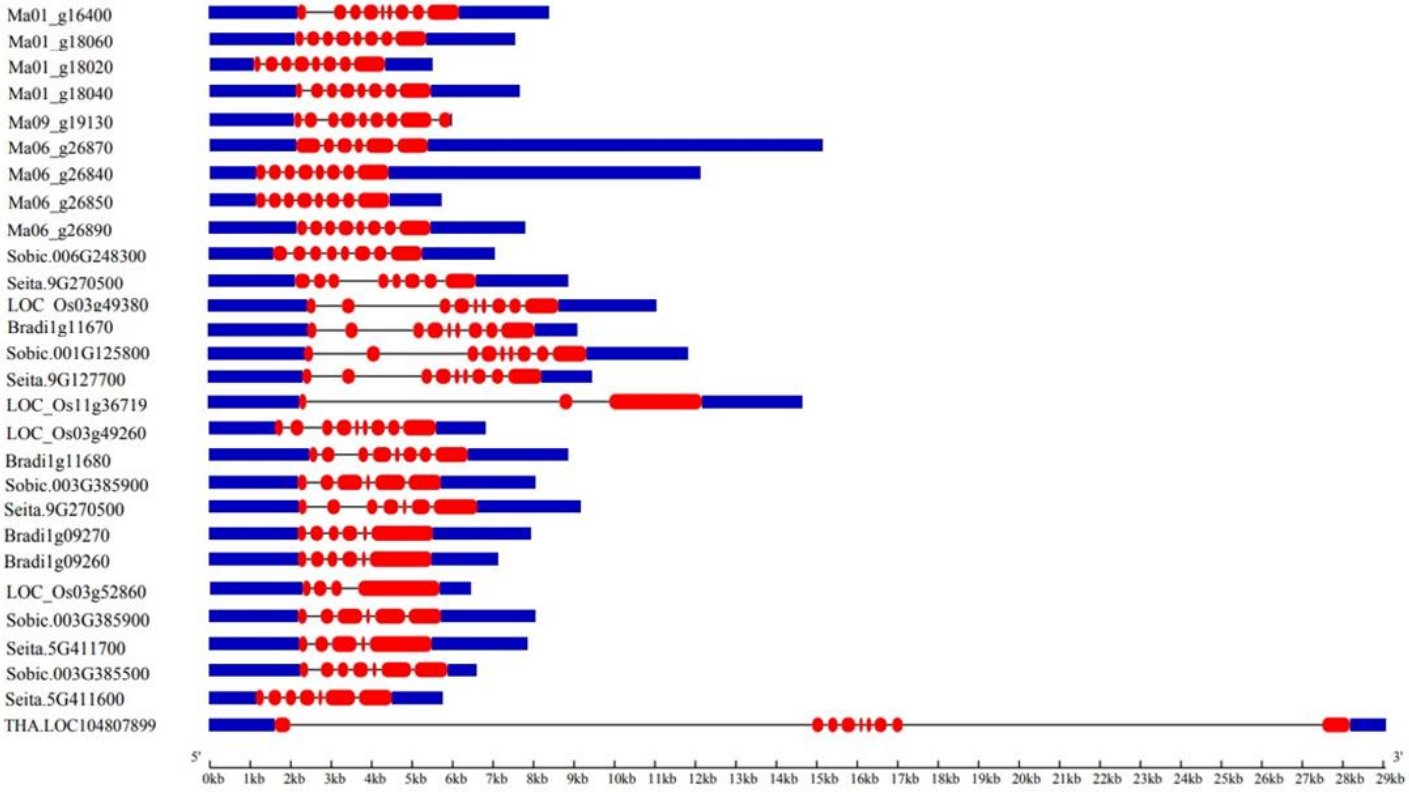
	Mapoly0191s0014	2.958	985	Plaza	
	Mapoly0191s0015	2.907	968	Plaza	
	Mapoly0023s0182	2.823	940	Plaza	
Chlamydomonas reinhardtii	Cre12.g512300	2.967	988	Softberry	1
	THA.LOC104808057	2.775	924	Plaza	
	THA.LOC104812216	2.700	899	Plaza	
	THA.LOC104820855	2.586	861	Plaza	
Tarenaya hassleriana	THA.LOC104807899	2.667	888	Plaza	8
	THA.LOC104808056	2.724	913	Plaza	
	THA.LOC104823730	2.580	859	Plaza	
	THA.LOC104814043	2.658	885	Plaza	
	THA.LOC104806939	2.652	883	Plaza	

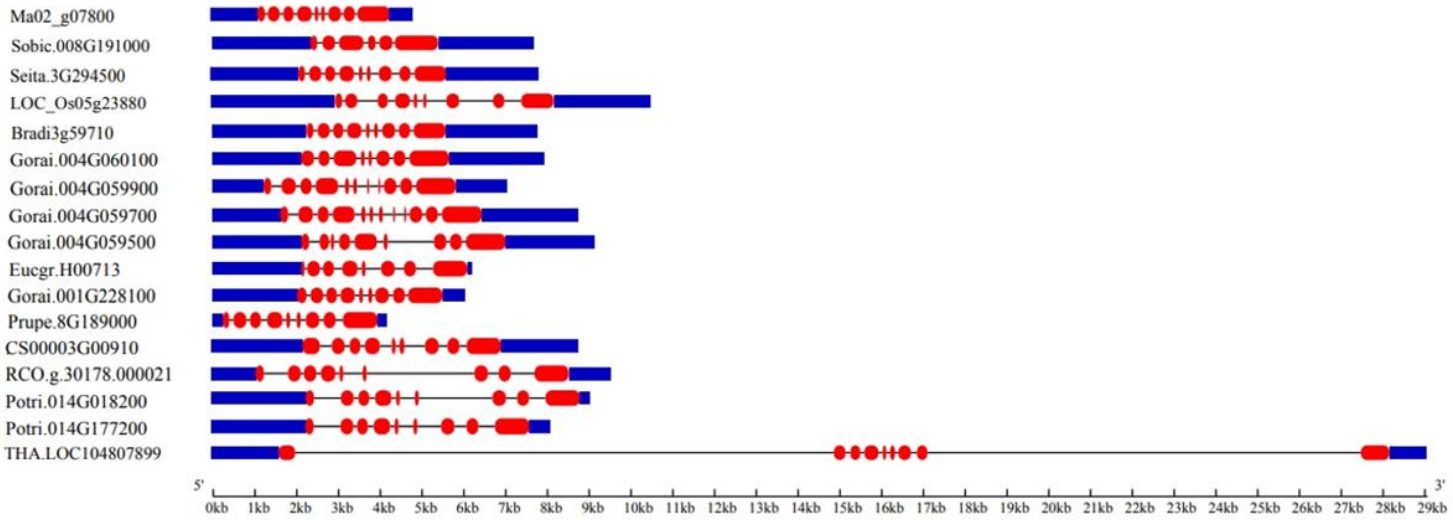
S2: Exon and intron structure. Exons are represented by a red color and introns are represented by a dark-gray line.

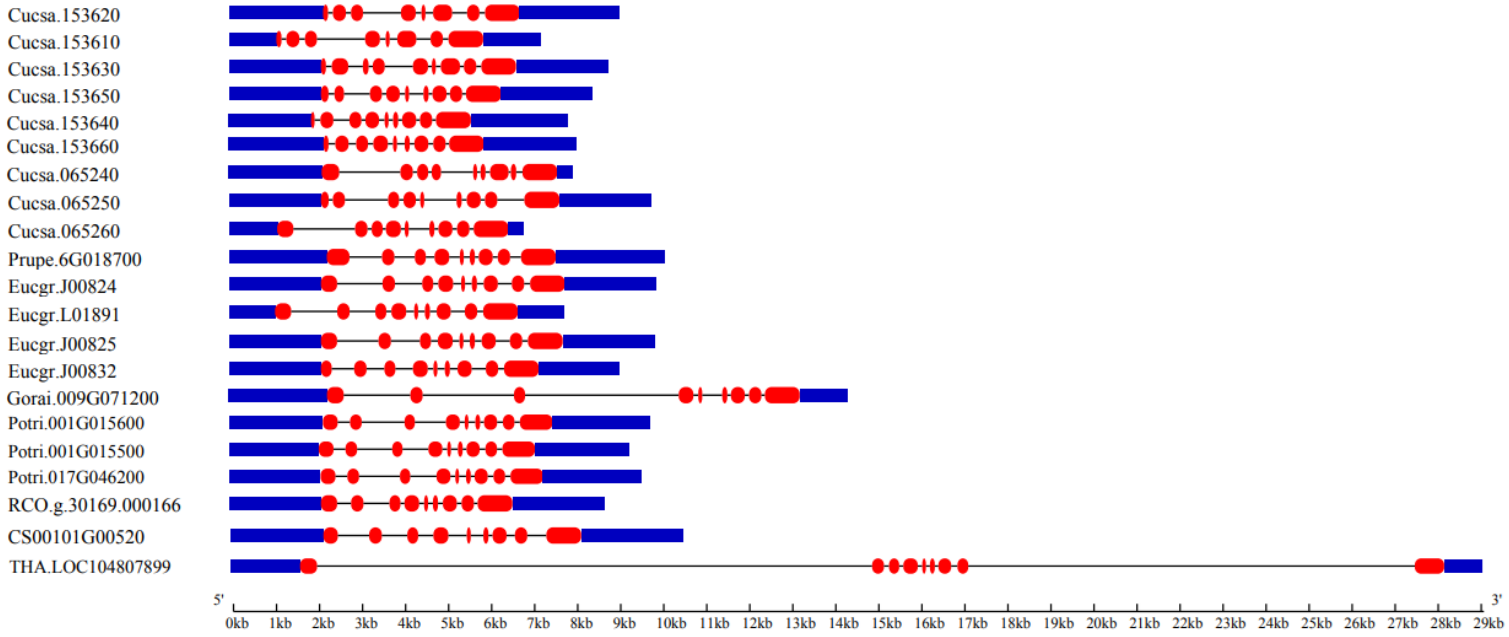


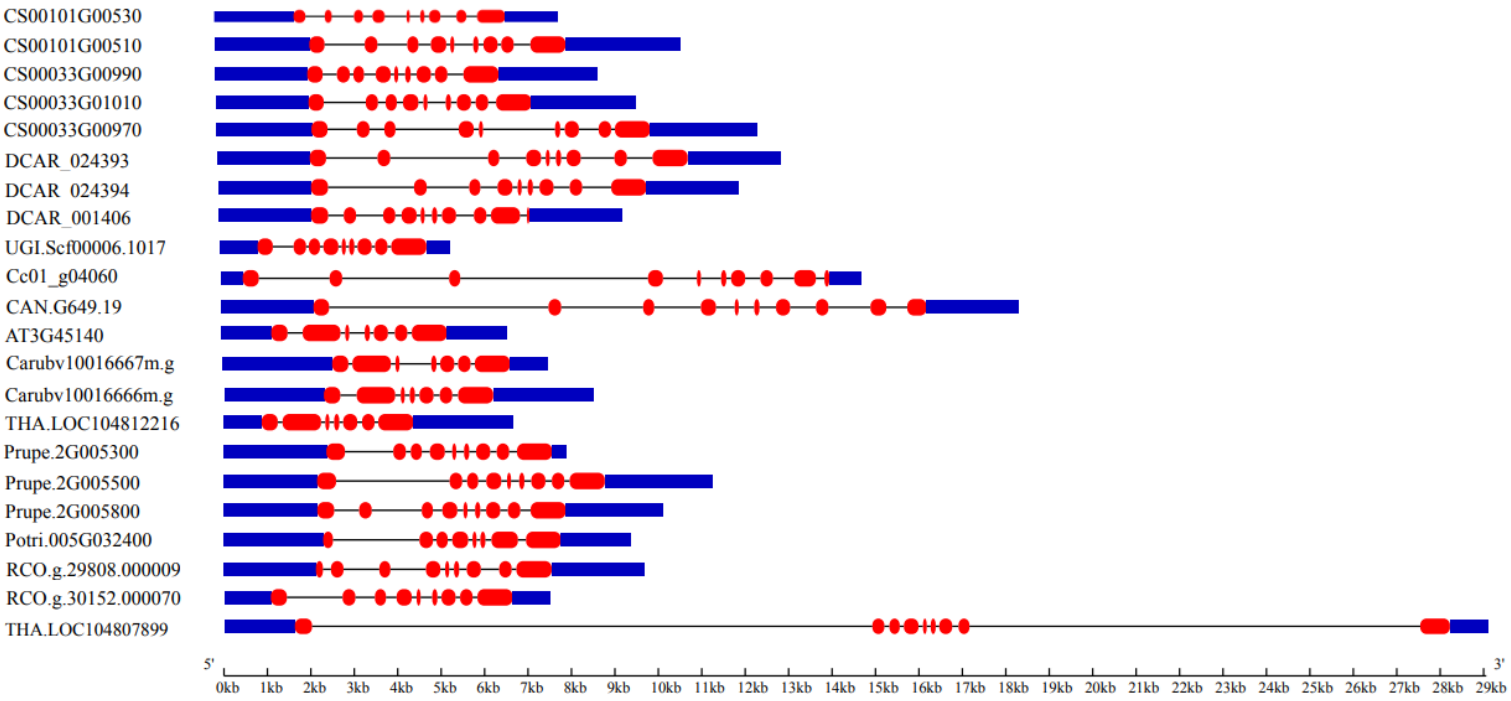

CAN.G1152.14
CAN.G1034.33
CAN.G466.4
CAN.G1034.38
CAN.G1034.37
Cucsa.091390
Cucsa.091380
Cucsa.091360
Cucsa.091350
Cucsa.091290
Prupe.6G324200
Prupe.6G324400
Prupe.6G324600
Eucgr.H04495
Eucgr.H04499
Eucgr.H04496
Eucgr.H04497
Eucgr.H04498
RCO.g.29680.000053
THA.LOC104807899
5'
3'
0kb 1kb 2kb 3kb 4kb 5kb 6kb 7kb 8kb 9kb 10kb 11kb 12kb 13kb 14kb 15kb 16kb 17kb 18kb 19kb 20kb 21kb 22kb 23kb 24kb 25kb 26kb 27kb 28kb 29kb

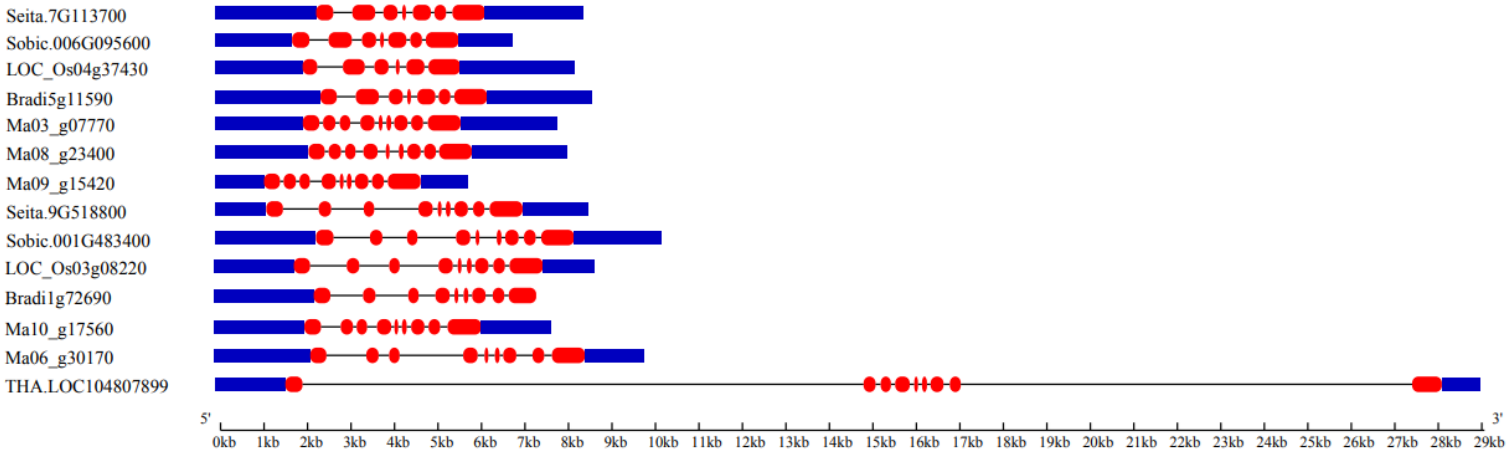

Bradi3g07000
Bradi3g07010
LOC_Os02g10120
Seita.7G113700
Sobic.004G078600
Seita.4G215400
LOC_Os03g49260
LOC_Os08g39840
LOC_Os08g39850
Ma03_g11520
Bradi3g39980
Seita.6G205300
Sobic.001G483400
Ma09_g12090
THA.LOC104807899
5'
3'
0kb 1kb 2kb 3kb 4kb 5kb 6kb 7kb 8kb 9kb 10kb 11kb 12kb 13kb 14kb 15kb 16kb 17kb 18kb 19kb 20kb 21kb 22kb 23kb 24kb 25kb 26kb 27kb 28kb 29kb

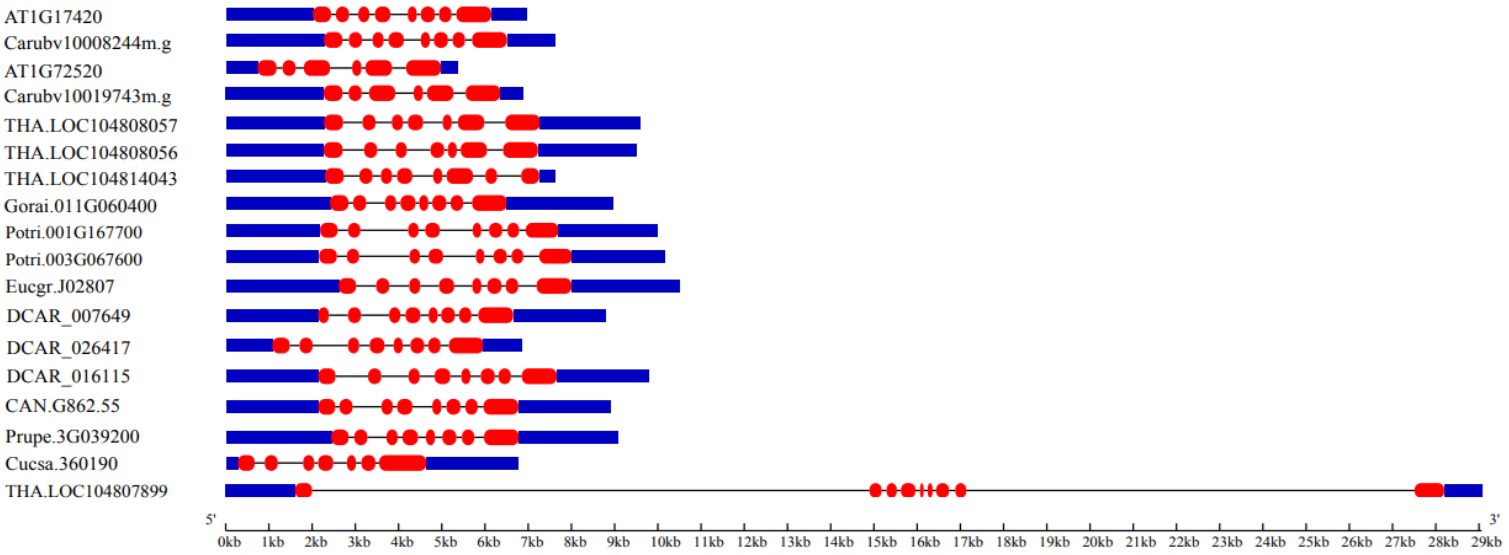

Ma01_g16400
Ma01_g18060
Ma01_g18020
Ma01_g18040
Ma09_g19130
Ma06_g26870
Ma06_g26840
Ma06_g26850
Ma06_g26890
Sobic.006G248300
Seita.9G270500
LOC_Os03g49380
Bradi1g11670
Sobic.001G125800
Seita.9G127700
LOC_Os11g36719
LOC_Os03g49260
Bradi1g11680
Sobic.003G385900
Seita.9G270500
Bradi1g09270
Bradi1g09260
LOC_Os03g52860
Sobic.003G385900
Seita.5G411700
Sobic.003G385500
Seita.5G411600
THA.LOC104807899
5'
3'
0kb 1kb 2kb 3kb 4kb 5kb 6kb 7kb 8kb 9kb 10kb 11kb 12kb 13kb 14kb 15kb 16kb 17kb 18kb 19kb 20kb 21kb 22kb 23kb 24kb 25kb 26kb 27kb 28kb 29kb

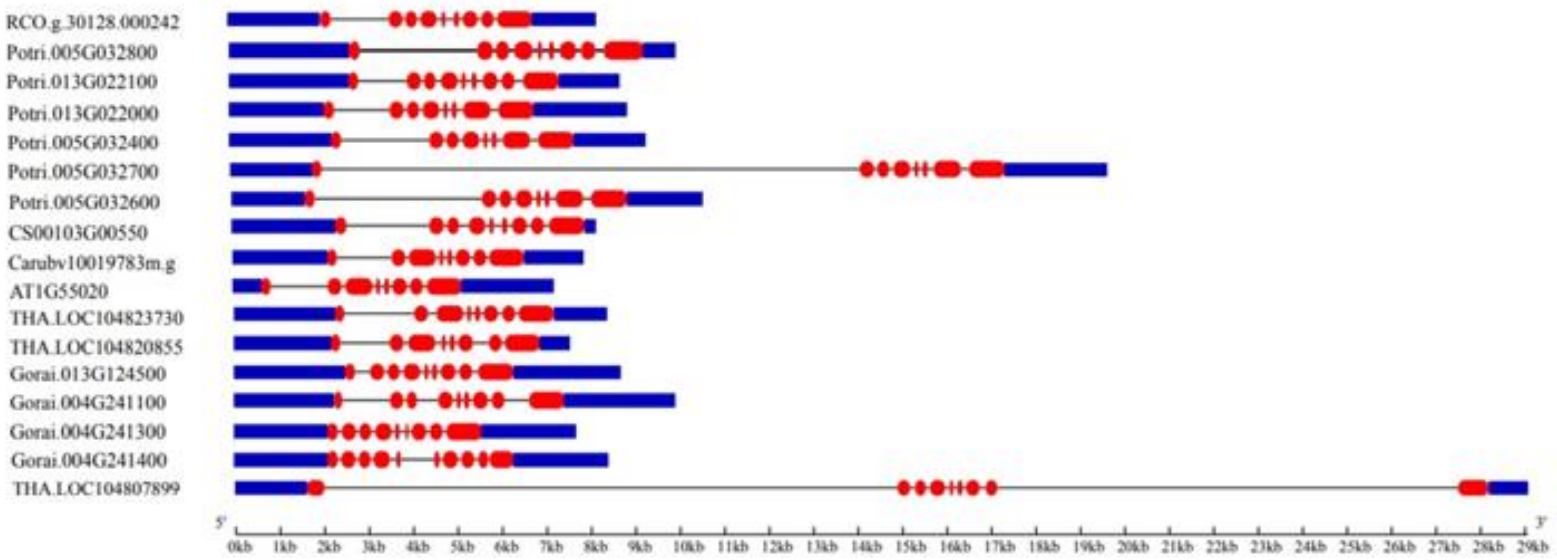

Ma02_g07800
Sobic.008G191000
Seita.3G294500
LOC_Os05g23880
Bradi3g59710
Gorai.004G060100
Gorai.004G059900
Gorai.004G059700
Gorai.004G059500
Eucgr.H00713
Gorai.001G228100
Prupe.8G189000
CS00003G00910
RCO.g.30178.000021
Potri.014G018200
Potri.014G177200
THA.LOC104807899
5'
3'
0kb 1kb 2kb 3kb 4kb 5kb 6kb 7kb 8kb 9kb 10kb 11kb 12kb 13kb 14kb 15kb 16kb 17kb 18kb 19kb 20kb 21kb 22kb 23kb 24kb 25kb 26kb 27kb 28kb 29kb

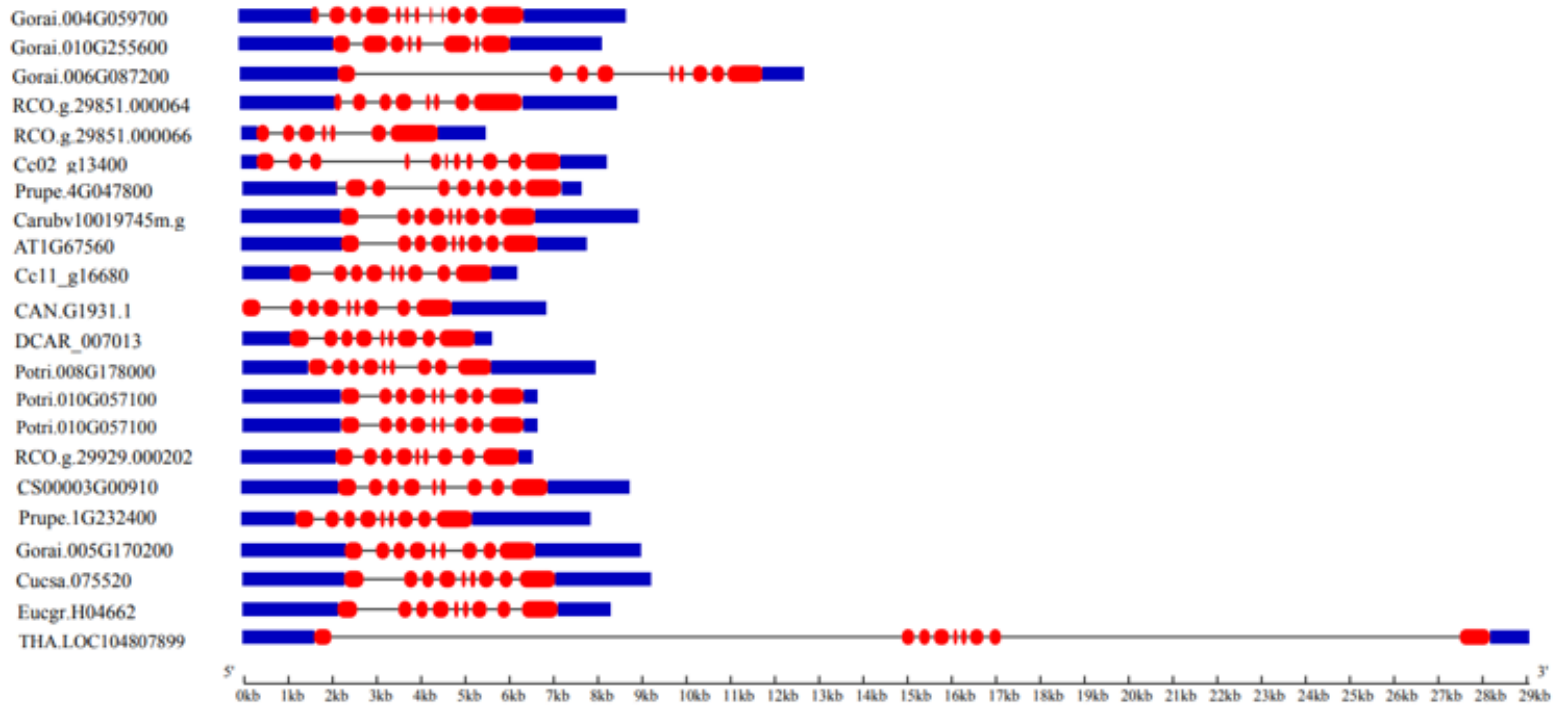

Cucsa.153620
Cucsa.153610
Cucsa.153630
Cucsa.153650
Cucsa.153640
Cucsa.153660
Cucsa.065240
Cucsa.065250
Cucsa.065260
Prupe.6G018700
Eucgr.J00824
Eucgr.L01891
Eucgr.J00825
Eucgr.J00832
Gorai.009G071200
Potri.001G015600
Potri.001G015500
Potri.017G046200
RCO.g.30169.000166
CS00101G00520
THA.LOC104807899
5'
3'
0kb 1kb 2kb 3kb 4kb 5kb 6kb 7kb 8kb 9kb 10kb 11kb 12kb 13kb 14kb 15kb 16kb 17kb 18kb 19kb 20kb 21kb 22kb 23kb 24kb 25kb 26kb 27kb 28kb 29kb


CS00101G00530
CS00101G00510
CS00033G00990
CS00033G01010
CS00033G00970
DCAR_024393
DCAR_024394
DCAR_001406
UGI.Scf00006.1017
Cc01_g04060
CAN.G649.19
AT3G45140
Carubv10016667m.g
Carubv10016666m.g
THA.LOC104812216
Prupe.2G005300
Prupe.2G005500
Prupe.2G005800
Potri.005G032400
RCO.g.29808.000009
RCO.g.30152.000070
THA.LOC104807899
5'
3'
0kb 1kb 2kb 3kb 4kb 5kb 6kb 7kb 8kb 9kb 10kb 11kb 12kb 13kb 14kb 15kb 16kb 17kb 18kb 19kb 20kb 21kb 22kb 23kb 24kb 25kb 26kb 27kb 28kb 29kb

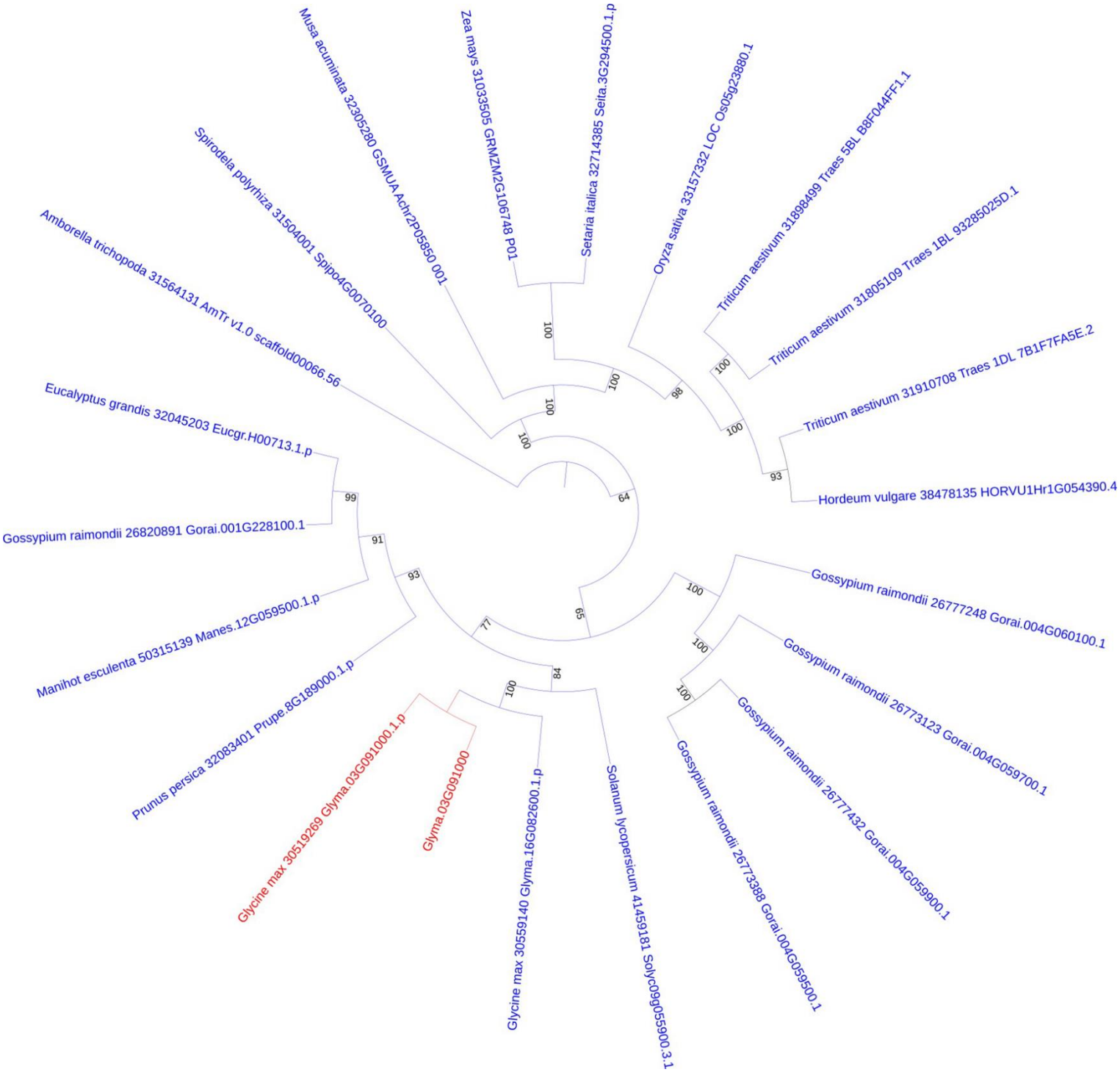

Seita.7G113700
Sobic.006G095600
LOC_Os04g37430
Bradi5g11590
Ma03_g07770
Ma08_g23400
Ma09_g15420
Seita.9G518800
Sobic.001G483400
LOC_Os03g08220
Bradi1g72690
Ma10_g17560
Ma06_g30170
THA.LOC104807899
5'
3'
0kb 1kb 2kb 3kb 4kb 5kb 6kb 7kb 8kb 9kb 10kb 11kb 12kb 13kb 14kb 15kb 16kb 17kb 18kb 19kb 20kb 21kb 22kb 23kb 24kb 25kb 26kb 27kb 28kb 29kb


AT1G17420
Carubv10008244m.g
AT1G72520
Carubv10019743m.g
THA.LOC104808057
THA.LOC104808056
THA.LOC104814043
Gorai.011G060400
Potri.001G167700
Potri.003G067600
Eucgr.J02807
DCAR_007649
DCAR_026417
DCAR_016115
CAN.G862.55
Prupe.3G039200
Cucsa.360190
THA.LOC104807899
5'
3'
0kb 1kb 2kb 3kb 4kb 5kb 6kb 7kb 8kb 9kb 10kb 11kb 12kb 13kb 14kb 15kb 16kb 17kb 18kb 19kb 20kb 21kb 22kb 23kb 24kb 25kb 26kb 27kb 28kb 29kb


RCO.g.30128.000242
Potri.005G032800
Potri.013G022100
Potri.013G022000
Potri.005G032400
Potri.005G032700
Potri.005G032600
CS00103G00550
Carubv10019783m.g
AT1G55020
THA.LOC104823730
THA.LOC104820855
Gorai.013G124500
Gorai.004G241100
Gorai.004G241300
Gorai.004G241400
THA.LOC104807899
5'
3'
0kb 1kb 2kb 3kb 4kb 5kb 6kb 7kb 8kb 9kb 10kb 11kb 12kb 13kb 14kb 15kb 16kb 17kb 18kb 19kb 20kb 21kb 22kb 23kb 24kb 25kb 26kb 27kb 28kb 29kb


Gorai.004G059700
Gorai.010G255600
Gorai.006G087200
RCO.g.29851.000064
RCO.g.29851.000066
Cc02_g13400
Prupe.4G047800
Carubv10019745m.g
AT1G67560
Cc11_g16680
CAN.G1931.1
DCAR_007013
Potri.008G178000
Potri.010G057100
Potri.010G057100
RCO.g.29929.000202
CS00003G00910
Prupe.1G232400
Gorai.005G170200
Cucsa.075520
Eucgr.H04662
THA.LOC104807899
5'
3'
0kb 1kb 2kb 3kb 4kb 5kb 6kb 7kb 8kb 9kb 10kb 11kb 12kb 13kb 14kb 15kb 16kb 17kb 18kb 19kb 20kb 21kb 22kb 23kb 24kb 25kb 26kb 27kb 28kb 29kb

S3: LOX9_B distribution in angiosperms.



Apêndice III

MATERIAL SUPLEMENTAR (ARTIGO 2)