JÉFYNE CAMPOS CARRÉRA

INTEGRATED PHENOTYPIC, METABOLIC, AND PROTEOMIC PROFILING TO ELUCIDATE PRODUCTIVITY AND QUALITY OF *C. arabica* CULTIVARS FROM CERRADO MINEIRO

LAVRAS – MG
2025

JÉFYNE CAMPOS CARRÉRA

INTEGRATED PHENOTYPIC, METABOLIC, AND PROTEOMIC PROFILING TO ELUCIDATE PRODUCTIVITY AND QUALITY OF *C. arabica* CULTIVARS FROM CERRADO MINEIRO

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-graduação em Botânica Aplicada, área de concentração em Botânica Aplicada, para a obtenção do título de Doutor.

Prof. Dr. Fábio Akira Mori
Orientador

Dra. Vânia Aparecida Silva
Coorientadora

LAVRAS – MG
2025

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

Campos Carréra, Jéfyne.
Integrated phenotypic, metabolic, and proteomic profiling to elucidate productivity and quality of C. Arabica cultivars from Cerrado Mineiro / Jéfyne Campos Carréra. - 2024.
201 p. : il.

Orientador: Fábio Akira Mori
Coorientadora: Vânia Aparecida Silva

Tese (Doutorado) - Universidade Federal de Lavras, 2024.
Bibliografia.

1. anatomia foliar. 2. ecometabolômica. 3. metabolômica não direcionada. 4. proteômica. 5. cafeeiro. I. Akira Mori, Fábio. II. Aparecida Silva, Vânia. III. Universidade Federal de Lavras. IV. Título.

JÉFYNE CAMPOS CARRÉRA

PERFIL FENOTÍPICO, METABÓLICO E PROTEÔMICO INTEGRADO PARA ELUCIDAR A PRODUTIVIDADE E A QUALIDADE DAS CULTIVARES DE *C. arabica* DO CERRADO MINEIRO

INTEGRATED PHENOTYPIC, METABOLIC, AND PROTEOMIC PROFILING TO ELUCIDATE PRODUCTIVITY AND QUALITY OF *C. arabica* CULTIVARS FROM CERRADO MINEIRO

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-graduação em Botânica Aplicada, área de concentração em Botânica Aplicada, para a obtenção do título de Doutor.

APROVADA em 13 de dezembro de 2024.
Dr. Fábio Akira Mori UFLA
Dra. Leonor de Castro Esteves Guerra Guimarães UNIVERSIDADE DE LISBOA
Dra. Marinês Ferreira Pires Lira UFLA
Dra. Vânia Aparecida Silva EPAMIG
Dr. Vitor de Laia Nascimento UFLA

Prof. Dr. Fábio Akira Mori
Orientador

Dra. Vânia Aparecida Silva
Coorientadora

LAVRAS – MG
2025

Aos meus queridos companheiros diários,

Leonardo e Bartholomeo.

Dedico

AGRADECIMENTOS

A Deus, por me guiar, fortalecer e permitir que eu vivenciasse momentos únicos ao longo desta jornada.

Aos meus pais e irmãos – Grazyanne, Graziele e Jeferson –, que, mesmo à distância, sempre me apoiaram; e às minhas avós, Inesila e Dina, bem como às minhas tias, por me incentivarem, desde a infância, a valorizar os estudos.

Ao meu noivo, Leonardo, pelo apoio incondicional durante todo o doutorado. Sua paciência, compreensão e constante incentivo foram fundamentais para a conclusão desta tese.

À Universidade Federal de Lavras, em especial ao Programa de Pós-Graduação em Botânica Aplicada, pela oportunidade concedida.

Às minhas companheiras de quarto, Vitória e Gislayne; aos colegas de doutorado, Lucas Munõz, Luana, Orivaldo, Joabe, Camila e Fernanda; e aos amigos do Laboratório de Anatomia da Madeira, pelas conversas, pelo companheirismo e pelos momentos de descontração ao longo do curso.

Aos meus orientadores, Fábio Akira e Vânia Silva, e à pesquisadora Margarete Volpato, pela orientação e suporte na realização desta pesquisa.

À professora Leonor Guerra-Guimarães, minha orientadora durante o doutorado no exterior, por toda a ajuda, tanto pessoal quanto profissional, ao longo do desenvolvimento deste trabalho.

Ao professor Mário Lúcio, por sua valiosa contribuição nas análises proteômicas.

À Dra. Carla Pinheiro, pela colaboração nas análises laboratoriais e na revisão dos trabalhos.

Aos colegas do Centro de Investigação das Ferrugens do Cafeeiro (CIFC) e a todos aqueles que conheci durante minha estadia no exterior, pelos momentos agradáveis de convivência.

Ao Dr. John Charles D'Auria e a todos os integrantes do grupo de pesquisa Metabolic Diversity, do Leibniz Institute of Plant Genetics and Crop Plant Research (IPK Gatersleben), pela calorosa recepção e pelo auxílio nas análises metabolômicas.

À Dra. Céline Leclercq, do Instituto de Ciência e Tecnologia de Luxemburgo, por seu apoio nas análises proteômicas.

Ao professor Paulo Guimarães, por toda a assistência nas análises estatísticas.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, além de recursos do Programa de Doutorado Sanduíche no Exterior (PDSE).

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), ao Consórcio Brasileiro de Pesquisa e Desenvolvimento do Café, ao Instituto Nacional de Ciência e Tecnologia do Café (INCT Café/CNPq) e à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), pelo apoio concedido ao projeto.

Esta pesquisa contou com apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (código de financiamento: 001), do Programa Horizon2020 da União Europeia (EPPN2020 Grant Agreement 731013), da Fundação para a Ciência e Tecnologia (FCT) de Portugal, da Unidade de Investigação LEAF (UID/AGR/04129/2020) e da UCIBIO (FCT UIDP/04378/2020; FCT UIDB/04378/2020).

A todos que, de alguma forma, contribuíram para a realização desta tese, meus mais profundos e sinceros agradecimentos!

"…During its life, be it weeks, months, or years
Each leaf serves, through adversity, perseveres
Though some may go while others appear
All play a role through every tier
Removing CO_2 from atmosphere
Supporting life in the biosphere
To all life's relief
Whether long or brief
The life of a leaf
Is nature's motif."

(Poema *The Life of a Leaf*
por William W. Adams III)

RESUMO

O Brasil é o maior exportador mundial de café arábica, com Minas Gerais respondendo por mais da metade da produção nacional. No Estado, a região do Cerrado Mineiro é reconhecida pela produção de cafés especiais, sendo a primeira região brasileira a obter denominação de origem. Por essa razão, avaliações periódicas em campo são essenciais para verificar o desempenho de cultivares *Coffea arabica* L. que apresentem resiliência e mantenham padrão constante de qualidade da bebida. Para a realização desse estudo foram coletadas amostras de folhas de cinco cultivares *C. arabica* cultivadas em dois locais experimentais distintos na região do Cerrado Mineiro. Essas cultivares foram selecionadas com base em suas características de produtividade, qualidade de bebida e resistência à ferrugem do cafeeiro e a nematoides. O material foi submetido a uma análise abrangente, incluindo análises anatômicas, determinação de área foliar específica, potencial hídrico foliar e índices espectrais, além da quantificação de amido e proteínas totais. Também foram realizadas análises metabolômicas não direcionadas, proteômicas e de atividade enzimática (especialmente invertases). Foram levantados dados de clima e características de solo, altitude e regime hídrico dos locais experimentais, assim como as características agronômicas das cultivares: qualidade de bebida e produtividade média. Para analisar esse conjunto de dados complexo, foi aplicada uma combinação de métodos estatísticos, incluindo análises univariadas e algoritmos de aprendizado de máquina, para extrair informações significativas sobre o desenvolvimento do campo de cultivares. As cultivares analisadas apresentaram maiores variações no metaboloma e proteoma do que na anatomia foliar. Apesar das semelhanças climáticas entre os locais experimentais, as variáveis ambientais e o estado hídrico das folhas no momento da coleta influenciaram mais o metabolismo das plantas. A análise do metaboloma e do proteoma permitiu identificar vias metabólicas relacionadas ao estresse, especialmente nas cultivares Catiguá MG2 e Paraíso 2. A cultivar Paraíso 2 inclusive demonstrou maior estabilidade agronômica, mantendo alta produtividade e qualidade da bebida nos dois locais de avaliação. A pesquisa contribui para o entendimento da aclimatação de cultivares de café em campo, orientando estratégias de manejo e melhoramento para aumentar a produtividade, a qualidade da bebida e a resiliência frente aos desafios climáticos futuros.

Palavras-chave: anatomia foliar; ecometabolômica; metabolômica não direcionada; proteômica; cafeeiro.

ABSTRACT

Brazil is the world's largest exporter of arabica coffee, with the State of Minas Gerais accounting for more than half of the national production. Within the state, the Cerrado Mineiro region is renowned for specialty coffee production and was the first Brazilian region to receive a designation of origin. For this reason, periodic field evaluations are essential to assess the performance of Coffea arabica L. cultivars that exhibit resilience and maintain consistent cup quality standards. In this study, leaf samples from five C. arabica cultivars were collected from two experimental sites within the Cerrado Mineiro region. These cultivars were selected based on their productivity, cup quality, and resistance to coffee leaf rust and nematodes. The material underwent comprehensive analysis, including anatomical analysis, specific leaf area, leaf water potential, and spectral indices determination and quantification of starch and total proteins. In addition, untargeted metabolomic, proteomic, and enzymatic activity analyses (particularly invertases) were performed. Climate data, soil characteristics, altitude, water regime, and agronomic traits of the cultivars (cup quality and mean productivity) were also retrieved. To analyze this complex dataset, a combination of statistical methods was applied, including univariate analyses and machine learning algorithms, to extract meaningful insights into the field performance of the cultivars. The analyzed cultivars showed higher metabolome and proteome profile variation than in leaf anatomy. Despite similar climatic conditions at the experimental sites, environmental variables and the leaf water status at the sampling time had more influence on plant metabolism. Metabolomic and proteomic analyses revealed stress-related metabolic pathways, particularly in the Catiguá MG2 and Paraíso 2 cultivars. Notably, Paraíso 2 demonstrated higher agronomic stability, maintaining high productivity and cup quality in both environments. This research enhances the understanding of coffee cultivar acclimation in field conditions, guiding management and breeding strategies to improve productivity, beverage quality, and resilience to future climate challenges.

Keywords: leaf anatomy; ecometabolomics; untargeted metabolomics; proteomics; coffee plant.

INDICADORES DE IMPACTO

As mudanças climáticas exercem influência significativa sobre a produtividade agrícola, tornando a seleção de cultivares mais tolerantes uma estratégia crucial para garantir o suprimento contínuo de alimentos no Brasil. No Estado de Minas Gerais, o café, além de sua relevância econômica, possui um profundo significado cultural. A organização de cooperativas na região do Cerrado Mineiro tem sido fundamental para a produção de cafés especiais, elevando o valor agregado do produto. A fim de assegurar a produtividade e a qualidade do café nessa região, é imprescindível avaliar o desempenho de diferentes cultivares de *Coffea arabica* L. em condições de campo. Os resultados obtidos ao longo desta pesquisa evidenciam que diferentes cultivares apresentam estratégias adaptativas distintas para lidar com a variabilidade ambiental. Em particular, a cultivar Paraíso 2 demonstrou maior resiliência em campo, aliada a uma produção e qualidade consistentes nos locais avaliados. É importante destacar que essa foi a primeira vez que análises do metaboloma e proteoma foram desenvolvidas com cultivares de *Coffea arabica* L. no Cerrado Mineiro, fornecendo subsídios valiosos para o programa de melhoramento genético do café. Os resultados obtidos contribuirão para o avanço da área da tecnologia e produção cafeeira no Brasil.

IMPACT INDICATORS

Climate change has a significant influence on agricultural productivity, highlighting the selection of more tolerant cultivars as an essential strategy to secure the continued food supply in Brazil. In the state of Minas Gerais, coffee has a profound cultural significance in addition to its economic importance. The organization of cooperatives in the Cerrado Mineiro region has been fundamental to the production of specialty coffees, increasing the added value of the product. To ensure the productivity and quality of coffee in this region, it is essential to evaluate the performance of different *Coffea arabica* L. cultivars under field conditions. The results obtained during this research show that different cultivars have different adaptive strategies for dealing with environmental variability. In particular, the Paraíso 2 cultivar showed greater resilience in the field, combined with consistent production and cup quality in the sites evaluated. It is important to note that this was the first time that metabolome and proteome analyses were carried out with *Coffea arabica* L. cultivars in the Cerrado Mineiro, providing valuable information for the coffee breeding program. The results obtained will contribute to advances in the field of coffee technology and production in Brazil.

SUMÁRIO

PRIMEIRA PARTE

INTRODUÇÃO GERAL

O café é uma das commodities agrícolas de maior importância econômica no cenário global, sendo majoritariamente representado pelas espécies *Coffea arabica* L. e *Coffea canephora* Pierre ex A. Froehner (Davis *et al*., 2019; International Coffee Organization, 2024). As plantas de *C. arabica* desenvolvem-se de maneira ideal em temperaturas que variam entre 18 °C e 23 °C (DaMatta *et al*., 2018). O cultivo ocorre predominantemente na faixa intertropical do globo, onde fatores ambientais como precipitação, temperatura, radiação solar e concentração de dióxido de carbono (CO_2) afetam de forma significativa o crescimento vegetativo e a produtividade da cultura (DaMatta *et al*., 2018). O Brasil, reconhecido como o maior produtor mundial, apresenta uma ampla diversidade de cultivares, as quais diferem quanto a características agronômicas e comerciais, incluindo qualidade da bebida, resistência a doenças e tolerância ao déficit hídrico (De Carvalho *et al*., 2022; International Coffee Organization, 2024).

A trajetória histórica do cultivo de café no Brasil está profundamente vinculada às condições climáticas, as quais exerceram influência decisiva na migração da produção cafeeira para distintas regiões do território nacional (Nagay, 1999). A partir da década de 1970, a expansão da cafeicultura foi impulsionada por avanços nos programas de melhoramento genético do cafeeiro, bem como pela incorporação de novas áreas de cultivo, com destaque para a região do Cerrado Mineiro, situada no noroeste do estado de Minas Gerais (Nagay, 1999; Caixeta *et al*., 2022). Esses fatores foram determinantes para o expressivo aumento da produção de café arábica em Minas Gerais.

A região do Cerrado Mineiro tem se destacado pela notável capacidade de organização cooperativa, o que possibilita a valorização de sua Indicação Geográfica (Denominação de Origem) no mercado interno, bem como o atendimento às exigências específicas do mercado internacional de cafés especiais (Moura; Bueno, 2023). As condições climáticas dessa região são fortemente influenciadas por sua localização geográfica em área continental, a qual reduz a atuação direta das massas de ar oceânicas e resulta em padrões característicos de precipitação (Fernandes *et al*., 2012). Em decorrência dessa configuração, o Cerrado Mineiro apresenta duas estações climáticas bem definidas: um verão predominantemente úmido e um inverno marcadamente seco (Santana; Chelotti, 2021). Durante o período seco, que se estende, em geral, de maio a setembro, os índices pluviométricos são consideravelmente reduzidos (Voltolini *et al*., 2025). Para mitigar os efeitos do estresse hídrico e otimizar o desempenho das lavouras, práticas de irrigação controlada são amplamente empregadas na região. Tais práticas

desempenham papel estratégico na regulação do desenvolvimento vegetativo e reprodutivo de diversas cultivares de café (Guerra *et al*., 2021).

Em 2016, foi iniciado, na região do Cerrado Mineiro, um projeto colaborativo intitulado "Unidades Demonstrativas de Cultivares de Café", desenvolvido pela Empresa de Pesquisa Agropecuária de Minas Gerais (EPAMIG) em parceria com a Fundação de Desenvolvimento do Cerrado Mineiro (FUNDACCER). O principal objetivo dessa iniciativa consistiu em apresentar aos cafeicultores mineiros novas cultivares com atributos agronômicos aprimorados, notadamente quanto à resistência a estresses, à produtividade e à qualidade da bebida. As unidades demonstrativas foram implantadas em propriedades comerciais distribuídas por 12 municípios da região do Cerrado Mineiro, além da Fazenda Experimental da EPAMIG, localizada no município de Patrocínio (EPAMIG, 2021).

A avaliação do desempenho em campo das cultivares de café costuma envolver a análise de características estruturais e metabólicas das folhas, uma vez que tais parâmetros fornecem informações relevantes sobre as respostas fisiológicas das plantas frente a condições de estresse (Coelho *et al*., 2022; Marques *et al.*, 2022) e sobre os impactos decorrentes na produtividade e na qualidade da bebida (Dos Santos *et al.*, 2022; Pérez-Molina *et al*., 2021).

Nesta tese são apresentados três artigos científicos desenvolvidos com base em dados relativos às características foliares — incluindo análises anatômicas, metabolômicas, proteômicas, entre outras — e agronômicas de cultivares de *Coffea arabica* L. cultivadas na região do Cerrado Mineiro, bem como em variáveis ambientais, como clima e solo. Os materiais foram coletados em duas unidades experimentais localizadas no município de Monte Carmelo (fazenda particular, com uso de irrigação) e Patrocínio (Campo experimental da EPAMIG, sem uso de irrigação).

O primeiro artigo teve como objetivo caracterizar as cultivares presentes no banco de germoplasma da EPAMIG (Patrocínio, MG) quanto às principais classes metabólicas, com vistas à identificação de assinaturas metabólicas relevantes para a diferenciação entre essas cultivares e à investigação de suas possíveis relações com preditores de desempenho, como produtividade, qualidade da bebida e sanidade vegetal.

No segundo artigo, os dados metabolômicos das cultivares cultivadas em duas unidades experimentais (Monte Carmelo e Patrocínio) foram analisados em conjunto com informações foliares referentes ao potencial hídrico, acúmulo de amido, concentração de proteínas totais e índices vegetativos, além de dados climáticos e edáficos. Para a avaliação desse conjunto complexo de informações, foram testados modelos gerados por algoritmos de aprendizado de máquina (*machine learning*), com o objetivo de identificar variáveis-chave para a classificação

das cultivares com base no contexto ambiental, bem como compreender as interações entre essas variáveis e sua influência sobre a produtividade e a qualidade da bebida de cafeeiros cultivados em condições de campo.

Por fim, o terceiro artigo teve como propósito aprofundar a análise das folhas das cultivares de *C. arabica* por meio da investigação de possíveis variações anatômicas, metabolômicas e proteômicas, além da atividade das enzimas invertases. Essa abordagem considerou o estado hídrico foliar de plantas cultivadas com e sem irrigação, permitindo avaliar seus efeitos sobre os mecanismos estruturais e bioquímicos associados ao desempenho das cultivares em condições de campo.

REFERÊNCIAS

CAIXETA, Eveline Teixeira *et al*. **Aceleração do melhoramento do cafeeiro via seleção genômica:** agilidade e eficácia no lançamento de novas cultivares, Documentos 17. Brasília: Embrapa Café, 2022.

COELHO, Larissa Sousa *et al*. Morphological, physiological, and agronomic traits of crossings of 'Icatu' x 'Catimor' coffee tree subjected to water deficit. **Pesquisa Agropecuária Brasileira**, v.57, e02788, 2022. DOI: https://doi.org/10.1590/S1678-3921.pab2022.v57.02788.

DaMATTA, Fábio M. *et al*. Physiological and Agronomic Performance of the Coffee Crop in the Context of Climate Change and Global Warming: A Review. **J Agric Food Chem.** v. 66, n. 21, p. 5264-5274. 2018. DOI: https://doi.org/10.1021/acs.jafc.7b04537.

DAVIS, Aaron P. *et al*. High extinction risk for wild coffee species and implications for coffee sector sustainability. **Science Advances**, v. 5, n. 1, 2019. DOI: https://doi.org/10.1126/sciadv.aav3473.

DE CARVALHO, Carlos Henrique S. *et al*. **Catálogo de cultivares de café arábica**, Documentos 16. Brasília: Embrapa Café, 2022.

DOS SANTOS, Cynthia Stephânia *et al.* Adaptations to the drought season and impacts on the yield of 'Híbrido de Timor' coffee tree in the Minas Gerais State Cerrado (Brazilian Savanna). **Pesq. Agropec. Trop.**, Goiânia, v. 52, e72448, 2022.

EMPRESA DE PESQUISA AGROPECUÁRIA DE MINAS GERAIS, EPAMIG. **Projeto que avalia o desempenho de cultivares de café no Cerrado Mineiro chega ao segundo ano de colheita,** 2020. Available in: https://www.epamig.br/projeto-que-avalia-o-desempenho-de-cultivares-de-cafe-no-cerrado-mineiro-chega-ao-segundo-ano-de-colheita/. Acessed in: 01 dez. 2024.

FERNANDES, André Luís Teixeira *et al*. A moderna cafeicultura dos Cerrados Brasileiros**. Pesq. Agropec. Trop., Goiânia**, v.42, n.2, p.231-240, 2012.

GUERRA, Antônio Fernando *et al*. Cafés do Brasil: Pesquisa, sustentabilidade e inovação. *In:* TELHADO, Samuel Filipe Pelicano e; CAPDEVILLE, Guy de (Ed.). **Tecnologias poupa-terra 2021**. Brasília, DF: Embrapa, 2021. Cap. 5, p. 63-75.

INTERNATIONAL COFFEE ORGANIZATION, ICO. **Trade Statistic Tables, Production, Cofffee Trade Status**, 2024. Available in: http://www.ico.org/ trade_statistics.asp. Acessed in: 20 jun. 2024.

MARQUES, Isabel *et al*. High-resolution shotgun proteomics reveals that increased air [CO_2] amplifies the acclimation response of *coffea* species to drought regarding antioxidative, energy, sugar, and lipid dynamics. **J Plant Physiol**, v. 276, 2022 DOI: https://doi.org/10.1016/j.jplph.2022.153788.

MOURA, Erick de Freitas; BUENO, Janaina Maria. Dinâmicas de expansão das Cooperativas agrícolas de café da região do Cerrado Mineiro. **Desafio Online**, Campo Grande, v.11, n.1, 2023. DOI: https://doi.org/10.55028/don.v11i1.12884.

NAGAY, Julio Hidemitsu Côrrea. Café do Brasil: dois séculos de história. **Formação econômica**, Campinas, v.3, p. 17-23, 1999.

PÉREZ-MOLINA, Junior Pastor *et al*. Treasured exceptions: Association of morphoanatomical leaf traits with cup quality of *Coffea arabica* L. Cv. “Catuaí”. **Food Research International**, v. 141, 110118, 2021. DOI:https://doi.org/10.1016/j.foodres.2021.110118.

SANTANA, Guilherme Henrique dos Santos; CHELOTTI, Marcelo Cervo. EXPRESSÕES SOBRE A INDICAÇÃO GEOGRÁFICA DO CAFÉ DA REGIÃO DO CERRADO MINEIRO. *In*: Menezes, Sônia de Souza Mendonça *et al*. (org.). **Geografia dos Alimentos**: territorialidades, identidades e valorização dos saberes e fazeres, 1. ed. Aracaju, SE: Criação Editora, 2021. Cap. 7, p. 117-140.

VOLTOLINI, Giovani Belutti. *et al*. Agronomic Performance of Irrigated and Rainfed Arabica Coffee Cultivars in the Cerrado Mineiro Region. **Agronomy**, v. 15, n. 1, 222. 2025. DOI: https://doi.org/10.3390/agronomy15010222.

SEGUNDA PARTE
ARTIGOS

ARTIGO 1: NON-TARGETED METABOLOMIC ANALYSIS OF FIELD-GROWN *Coffea arabica* CULTIVARS REVEALS DISTINCT LEAF METABOLIC SIGNATURES

Artigo aceito pelo *Theoretical and Experimental Plant Physiology Journal*
(Thematic Collection: Metabolomics and plant metabolism regulation and integration)

NON-TARGETED METABOLOMIC ANALYSIS OF FIELD-GROWN *Coffea arabica* CULTIVARS REVEALS DISTINCT LEAF METABOLIC SIGNATURES

Jéfyne Campos Carréra[1], Leonor Guerra-Guimarães[2,3], John Charles D'Auria[4], Luana de Jesus Sartori[1], Carla Pinheiro[5,6], Vânia Aparecida Silva[7], Margarete Lordelo Volpato[7], Gladyston Rodrigues Carvalho[7], Fabio Akira Mori[1]

[1]UFLA - Universidade Federal de Lavras, 37200-000 Lavras, Minas Gerais, Brasil.
[2]CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal.
[3]LEAF - Linking Landscape, Environment, Agriculture and Food Research Centre and TERRA Associated Laboratory, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017, Lisboa, Portugal.
[4]Research Group Metabolic Diversity, Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK Gatersleben), OT Gatersleben, Corrensstraße 3, 06466 Seeland, Germany
[5]UCIBIO Applied Molecular Biosciences Unit, Department of Life Sciences, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.
[6]Associate Laboratory i4HB Institute for Health and Bioeconomy, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.
[7]EPAMIG – Empresa de Pesquisa Agropecuária de Minas Gerais, 37200-000 Lavras, Minas Gerais, Brasil.

Correspondence author: Leonor Guerra-Guimarães, CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal. E-mail: leonorguimaraes@isa.ulisboa.pt. https://orcid.org/0000-0002-9676-1036.

ORCID List

Jéfyne Campos Carréra: https://orcid.org/0000-0001-6597-7152
John Charles D'Auria: https://orcid.org/0000-0002-4865-3938
Luana de Jesus Sartori: https://orcid.org/0000-0003-0764-977X
Carla Pinheiro: https://orcid.org/0000-0002-7799-0183
Vânia Aparecida Silva: https://orcid.org/0000-0001-6640-3644

Margarete Lordelo Volpato: https://orcid.org/0000-0002-6905-6876
Gladyston Rodrigues Carvalho: https://orcid.org/0000-0002-4466-674X
Fabio Akira Mori: https://orcid.org/0000-0002-7468-018X

Abstract

This study conducted a non-targeted metabolomic analysis of five *Coffea arabica* L. cultivars grown in the field experimental areas of Cerrado Mineiro (MG, Brazil) to identify their metabolic fingerprints. The five cultivars selected for this study were chosen based on their specific genetic backgrounds and traits, including disease resistance, productivity, and cup quality. A total of 463 metabolic features were detected in the overall *C. arabica* metabolome, with the major metabolic classes comprising sugars, amino acids, lipids, phenylpropanoids, and phenolic compounds. Among these, 41 metabolites were identified as key discriminators among the five cultivars. Partial least squares discriminant analysis (PLS-DA) revealed distinct metabolic profiles, highlighting ferulic acid, theobromine, octopamine, rosmarinic acid, and gibberellin as key metabolites. These findings emphasize the importance of phenolic compounds and alkaloids in cultivar discrimination. The most relevant metabolic markers associated with environmental stress tolerance suggest their potential as biochemical indicators for selecting resilient cultivars, thereby contributing to coffee breeding programs. Notably, this study is the first documented characterization of the leaf metabolome of field-grown *C. arabica* cultivars, with Catiguá MG2 emerging as the most distinct. Our findings demonstrate the efficacy of metabolomic fingerprinting via non-targeted metabolomic as a powerful tool for differentiating coffee cultivars and for precision breeding strategies.

Keywords: GC-TOF-MS, phenolics, alkaloids, Cerrado Mineiro (Brazil).

1. INTRODUCTION

Brazil is currently the world's leading producer and exporter of coffee, accounting for over one-third of global coffee production (International Coffee Organization, 2024). Within Brazil, Minas Gerais state is the primary region for cultivating *Coffea arabica* L., a position largely attributed to breeding research initiatives with developed numerous cultivars since the 1970s (Caixeta et al. 2022). The establishment of breeding programs, coupled with government incentives, has enabled the expansion of coffee cultivation into the Cerrado Mineiro, a region situated in northwest Minas Gerais renowned for its high-quality coffee (Nagay 1999, Caixeta

et al. 2022). Field characterization has a crucial role in identifying cultivars with superior yield potential, excellent cup quality, and optimal adaptation to the region's specific edaphoclimatic conditions (De Carvalho et al. 2022). Additionally, metabolic profiling provides a biochemical perspective on cultivar performance, linking field traits to the underlying metabolic processes that drive plant growth, tolerance to stressful conditions, and quality conditions.

Metabolomics enables the assessment of chemical compound variations in plants, stemming from species differences and environmental changes (Shen et al. 2023). Specifically, non-targeted metabolomics enables the identification of metabolic signatures by providing a comprehensive snapshot of the metabolome, encompassing both primary and specialized metabolites (Shen et al. 2023). By examining the specialized metabolism, distinct metabolic fingerprints can be identified within a given species, providing valuable insights into metabolic differentiation and adaptation mechanisms (Dussarrat et al. 2023).

The identification of metabolic signatures is crucial for understanding the biochemical differences among *C. arabica* cultivars. These signatures, shaped by genetic, physiological, and environmental factors (Kitashova et al. 2023), offer valuable tools for cultivar differentiation. They are characterized by distinct metabolic profiles, which are influenced by plant genetics, cultivation practices, and stress responses (Martins et al. 2014, Mihai et al. 2024). Since the leaf metabolome provides insights into key plant processes – such as health, stress responses, and development -, its evaluation is essential for understanding factors that impact coffee bean quality and yield (Da Silva et al. 2021, Chekol et al. 2024). In the context of coffee cultivars, metabolic signatures can help identify variations in key biochemical pathways, such as those governing alkaloid production and polyphenol biosynthesis (Cangeloni et al. 2022, Castro-Moretti et al. 2023). In addition, studies on coffee leaf metabolism have identified diverse classes of metabolites, including chlorogenic acids, benzophenones, flavonoids, xanthones, and alkaloids, that are critical to plant physiology and defense mechanisms (Souard et al. 2017, Montis et al. 2022).

By linking metabolic profiles to specific cultivars, a deeper understanding of the biochemical foundations of phenotypic diversity can be achieved. This knowledge is fundamental for breeding programs, precision agriculture, and the development of targeted strategies, aimed at enhancing both plant resilience and the sensory attributes of coffee products (Martins et al. 2014, Guimarães et al. 2021, Chekol et al. 2024).

While coffee metabolomics has traditionally focused on species differences within the Rubiaceae family, particularly the coffee bean's role in beverage quality (Amalia et al. 2023, Castro-Moretti et al. 2023, Zhai et al. 2024), recent studies have begun to explore how plant

tissues, including leaves, influence overall plant health and productivity (Castro-Moretti et al. 2023). As a key indicator of plant stress responses, nutrient status, and growth, the leaf metabolome provides valuable insights that indirectly affect coffee bean quality (Sardans et al. 2011, Bollen et al. 2024). To address this gap, this study applied non-targeted polar metabolomic analysis to the leaves of five field-grown *C. arabica* cultivars. By characterizing the unique metabolome of each cultivar and identifying potential metabolic signatures relevant to differentiation, this study aims to shed light on how leaf metabolic profiles serve as predictors of plant performance, contributing to optimized coffee production, plant health and enhanced beverage quality.

2. MATERIAL AND METHODS

2.1.Plant Material and Field Experiments

Seven-year-old *C. arabica* plants used in this study were grown since 2016 in an EPAMIG experimental germplasm field located in the municipality of Patrocínio, MG, Brazil (latitude: 18°57’ S, longitude: 46°54’ W) located in Cerrado Mineiro region. The selection of five cultivars used in the present study was based on their specific traits, resistance to *Hemileia vastatrix* (coffee rust) and *Meloidogyne exigua* (nematode), as well as their productivity and cup quality (Table 1). The study utilized rainfed plants (in dry season, august 2022). The plants were arranged in a 0.7 m spacing in the same crop row (plot). The soil conditions in terms of macro and micronutrient content and soil texture were standardized across the plot (Supplementary Table 1). A pair of fully expanded leaves from the middle third of five individual plants per cultivar were collected. The biological replicates were obtained from pooled individuals (Sumner et al. 2007) and the biological sampling was n = 5 for each coffee cultivar. Subsequently, the leaf midribs were excised and discarded, and the remaining material was preserved in liquid nitrogen. Leaf collection occurred between 07:00 and 09:00 h considering the sun-exposed side of the plant (Supplementary Figure 1).

2.2.Leaf extract

The leaf samples were ground with liquid nitrogen, followed by lyophilization, and subsequently stored in vacuum-sealed conical centrifuge tubes (Supplementary Figure 1). The extraction procedure was carried out with adaptations based on the method proposed by Salem et al. (2016, 2020), designed for the extraction of polar metabolites, lipids, starch and proteins from a single sample of leaf tissue. This involved transferring 30 mg of lyophilized material into microtubes containing 1.5 mL of the extraction solvent mixture, Methyl-tert-butyl-ether:

methanol (MTBE: MeOH, 3:1, v:v), along with internal standards U-^{13}C sorbitol and L-Alanine-d4 (50 µL/100 mL).

The microtubes were vortexed and incubated on an orbital shaker (40 rpm) for 45 min at 4 °C followed by a 15 min sonication step in a water bath at 4 °C. The microtubes were then centrifuged (10,000 x *g*) for 10 min at 4 °C. From the soluble fractions, aliquots of 0.5 mL were transferred to new microtubes and dried in the speed vacuum for 45 min for the metabolomic analyses.

2.3.Metabolite profiling

Dried metabolic extracts in sealed, crimped vials were derivatized using 10 µL of a 20 mg mL^{-1} solution of methoxyamine hydrochloride in pyridine and 10 µL of *N*-trimethylsilyl-*N*-methyl trifluoroacetamide (MSTFA) following the protocol of Erban et al. (2007) and injected using a Gerstel MPS2-XL (Gerstel, Muhlheim/Ruhr, Germany) autosampler. A LECO Pegasus BT time-of-flight mass spectrometer (LECO, St. Joseph, MI, USA) hyphenated with an Agilent 8890 gas chromatograph and helium as the carrier gas at 1.0 mL min^{-1} flow and linear velocity as flow control mode. The front inlet temperature was set at 230 °C, and samples were run in splitless mode. The capillary column used was an Agilent DB-35MS (30 m × 0.25 mm × 0.25 µm). The temperature program started at 85 °C in isothermal mode for 2 minutes. The isothermal step was followed by a 15°C/min ramp to 360°C. Fatty acid methyl esters (FAME MIX: C8-C30) were used in the determination of the retention time index (Kind et al. 2009). The ion source was set to 250 °C. The recorded mass range is m/z = 50-600 at 30 scans/sec. A solvent delay of 3 minutes was used. Filament bias current was -70V, and detector voltage ranged from approximately 1900-2000 V depending on detector age. The instrument was tuned and validated according to EPA tune compliance.

Metabolic features were identified by employing the LECO ChromaTOF software in conjunction with the GOLM Metabolome Database. Peak intensities were determined using the TargetSearch package (Cuadros-Inostroza et al. 2009) within the R software (R Core Team 2023). For normalisation, the content of the internal standard L-Alanine-d_4 was utilized. All metabolic class identifications were manually validated based on GOLM (Hummel et al. 2013), KEGG (Kanehisa et al. 2023), and HMDB (Wishart et al. 2022) metabolite databases. After annotating metabolic classes, 123 metabolic features were excluded from subsequent analysis. These features included reagents (44), unassigned or unknown features (10), and those not clearly identified as plant metabolites (69).

2.4.Statistical analysis

Metabolic classes were analyzed using Microsoft Office Excel (Microsoft Corp., Redmond, WA, USA) and R software (R Core Team 2023). Univariate and multivariate analyses were conducted using MetaboAnalyst 6.0 (Pang et al. 2024). Metabolic features with constant or single value across samples were excluded from the statistical analysis. Additionally, features with more than 50% missing values were removed, and for the remaining features, missing values were imputed with 1/5 of the minimum positive value for each variable. All data were log-transformed (base 10), and differences among the five groups (cultivars) were evaluated using a one-way analysis of variance (ANOVA). The significance level was set at $p < 0.05$. Significant metabolic features were further analyzed using partial least squares discriminant analysis (PLS-DA) and hierarchical cluster heatmap to cluster the cultivar samples based on auto-scaled variables. The cross-validation method was a 5-fold cross-validation (CV), with a maximum of five components considered for selection and evaluation based on three distinct accuracy metrics (Supplementary Figure 2). A variable importance plot (VIP) was generated to identify the most significant variables for distinguishing the five groups (cultivars) in the PLS-DA model (VIP > 1).

3. RESULTS

Non-targeted metabolomic analysis identified 463 metabolic features (MFs) in the leaves of *C. arabica*, consistently present in all the cultivars examined (Supplementary Figure 3). As multiple MFs can correspond to the same metabolite, duplicates were then removed, resulting in the identification of 398 unique metabolites (Supplementary Table 2). Metabolites were categorized by their chemical and functional metabolic characteristics using databases, such as, GOLM and KEGG pathways, enabling the identification of key classes in the *C. arabica* metabolome. Major metabolic classes identified were sugars (21%), amino acids (20%), lipids (12%) phenylpropanoids and phenolics (11%), and other organic acids (11%) as shown in Fig. 1. The compositional breakdown provides valuable insights into the metabolic diversity and complexity of *C. arabica* leaves, highlighting the roles of primary and secondary metabolites in coffee plant physiology.

Among the 463 metabolic features (MFs) identified, only 41 MFs were statistically significant in distinguishing the five cultivars assessed. Amino acids, sugars, alkaloids, and organic acids collectively accounted for 68% of these significant MFs (Table 2). In the cross-validation of the PLS-DA model, the accuracy exceeded 80% up to the third component (Supplementary Figure 2). The PLS-DA analysis distinguished Catiguá cultivars (MG2 and MG3) from the others (Catuaí, Paraíso 2, and Sarchimor), explaining 44.7% of the total variance

in the first component (Fig. 2). The second component accounted for 10.1% of the variance, effectively separating the Sarchimor cultivar from the others.

A hierarchical cluster heatmap analysis allowed visualisation of the relative values of the 41 metabolites that significantly change between cultivars (Fig. 3). Based on the variable importance plot (VIP), 15 out of the 41 significant metabolites were identified as the most relevant variables for distinguishing the five cultivars (highlighted in bold in the heatmap, Fig. 3). These significant 15 variables remained mostly consistent across the PLS-DA components (Supplementary Figure 4). Five key metabolites emerged as the most important for the model: ferulic acid, theobromine, and octopamine, followed by rosmarinic acid and gibberellin A4 (VIP scores > 1.5) (Fig. 3 and Supplementary Figure 4, B to F). The Catiguá MG2 cultivar exhibited elevated concentrations of four key metabolites, with the exception of octopamine, which was present in higher concentrations in the Paraíso 2 and Catuaí cultivars (Fig. 3 and Supplementary Figure 4, B to F, VIP Scores and heatmap).

4. DISCUSSION

In field-based plant metabolomics studies, maintaining the viability of plant material from collection to analysis remainss a major challenge. The metabolic extraction protocol outlined by Salem et al. (2020), adapted in this study for freeze-dried leaves, successfully facilitated the non-targeted metabolomics analysis of *C. arabica* cultivars. This approach allowed the identification of a substantial number of metabolic features, highlighting the most significant metabolic classes, such as sugars, amino acids, phenols, and lipids, in *C. arabica* cultivars grown in Cerrado Mineiro, one of Brazil's most important coffee-producing regions. Although the lipid content identified was lower than typically reported for coffee plants (Bastian et al. 2021), this finding aligns with the method employed, which is designed to extract polar metabolites (Salem et al. 2020). Additionally, several of the metabolites identified in the leaves of the five *C. arabica* cultivars have already been described in coffee leaves by other studies (Chen 2018, Cangeloni et al. 2022, Montis et al. 2022, Castro-Moretti et al. 2023), providing consistency and validation of the findings.

To distinguish the five *C. arabica* cultivars, sugars and amino acids accounted for nearly 50% of the most significant metabolites, but other important compound classes, such as phenylpropanoids, alkaloids, and lipids, also played key roles. Notably, two phenolic compounds, ferulic acid and rosmarinic acid, were crucial for differentiating the cultivars. Ferulic acid, synthesized via the shikimate pathway (Manivel et al. 2020), plays essential roles in structural and physiological processes. Its contrasting accumulation patterns, along with

shikimate, suggest a bottleneck in its conversion in the Catuaí, Paraíso 2, and Sarchimor cultivars. Ferulic acid contributes to cell wall stability (Kumar and Pruthi, 2014), potentially enhancing tolerance to environmental stresses. This compound is a precursor of chlorogenic acids (CGA) (Marques and Farah, 2009), a major class of phenolic acids known to accumulate in young coffee leaves (Monteiro et al., 2020). Previous studies have shown its presence in leaves (Chen 2018) and fruits (Kumar and Pruthi, 2014), but it has not been detected in stems (Carréra et al. 2023), highlighting its tissue-specific roles. Beyond its structural functions, ferulic acid may also exhibit antioxidant properties against biotic stressors (Ahlawat et al. 2024). Similarly, rosmarinic acid, another phenylpropanoid, was consistently detected in the cultivars studied here, as well as in *C. arabica* Obatã IAC 1669-20 (Castro-Moretti et al. 2023). However, it has been reported as absent in *C. arabica* Mundo Novo 376/4 and Iapar-59 (Da Silva et al. 2021). These results suggest that rosmarinic acid could serve as a cultivar-specific biomarker. Additionally, its antioxidant and anti-stress properties further support its role in enhancing cultivar resilience.

Alkaloids, a well-known and functionally significant group of coffee metabolites, accounted for 12% of the significant features identified. Among these, theobromine, a methylxanthine involved in caffeine biosynthesis (Jin et al. 2016, Montis et al. 2024), was particularly noteworthy. A recent study demonstrated a biochemical connection between leaf and bean metabolites, showing that caffeine concentrations in young leaves (Monteiro et al. 2020) were comparable to those found in green beans (Zheng et al. 2004). This relationship underscores the importance of leaf metabolomics not only for understanding plant physiological responses but also for providing insights into traits associated with coffee bean quality. Theobromine concentrations were significantly higher in the Catiguá MG2 cultivar compared to the others, reflecting its unique metabolic profile (Figure 3A). Dopamine, while classified as a biogenic amine, exhibits alkaloid-like properties, and was also identified as one of the significant metabolites. Derived from tyrosine, through the action of polyphenol oxidases and DOPA decarboxylase, dopamine plays a crucial role in plant growth, development, and in mitigating stress responses to both biotic and abiotic factors (Cherubino et al. 2023, Liu et al. 2020). Similarly, octopamine, another tyrosine-derived amine (Lee et al. 2009, Kanehisa et al. 2023), also plays a key role in distinguishing the five cultivars (Figure 3C). Together, these findings underscore the role of alkaloids and biogenic amines in not only distinguishing cultivars but also enhancing stress resilience and potentially influencing cup quality.

Beyond alkaloids, metabolites derived from the shikimate pathway, such as tryptophan, also emerged as key distinguishing features. Tryptophan, which accumulated more in the

Catiguá MG2 cultivar, is synthesized after anthranilate conversion (Desmet et al. 2021) and plays a pivotal role in stress response pathways. Along with dopamine and shikimate (Desmet et al. 2021, Cherubino et al. 2023), tryptophan highlighted the critical role of alkaloids in distinguishing the coffee cultivars. Gibberellin A4, a bioactive gibberellin and growth-promoting hormone, also varied significantly among cultivars. This compound is known to enhance drought stress resilience (Shohat et al. 2021), exhibit antioxidant activity (Nani et al. 2024) and, when combined with gibberellin A7 (GA4+7), mitigates the severe reduction of photosynthetic pigments, proteolysis, and lipid peroxidation (Ritonga et al. 2023). These findings are particularly relevant given that leaf samples were collected during the dry season, suggesting a link between gibberellin A4 concentrations and the varying drought tolerances of the studied cultivars. Such differences provide valuable insights into how specific metabolites contribute to cultivar adaptability under challenging environmental conditions.

The identification of these metabolic signatures in coffee leaves not only provides a biochemical basis for distinguishing *C. arabica* cultivars but also highlights their broader ecological and agronomic significance. Metabolites like theobromine, rosmarinic acid, and ferulic acid are closely tied to physiological and stress-adaptive traits, suggesting that targeted metabolomics could play a vital role in addressing challenges posed by climate change and increasing pest pressures. For instance, phenolic compounds could enhance pathogen resistance, while methylxanthines and gibberellins could support drought tolerance, ensuring the resilience of coffee plants in diverse growing conditions.

Future research should focus on targeted metabolomics and enzyme activity assays to better understand the regulatory mechanisms underlying these metabolic pathways. By studying how metabolites like alkaloids, phenylpropanoids, and gibberellins respond to specific stressors—such as drought, pests, or temperature fluctuations—researchers could identify key markers for breeding resilient cultivars. Multi-omics approaches integrating metabolomics with transcriptomics and proteomics under controlled stress conditions would provide a deeper understanding of the dynamic interplay between genotype, environment, and metabolism. For example, exploring caffeine metabolism or the biosynthesis of phenolic compounds in stressed plants could yield actionable insights for breeding and agricultural optimization.

Finally, the potential applications of leaf metabolomics extend beyond differentiation to breeding programs aimed at improving both agronomic and sensory attributes of coffee (Gamboa-Becerra et al. 2019). Metabolites like theobromine, ferulic acid, and gibberellin A4 not only contribute to stress tolerance but also indirectly influence traits associated with flavor and cup quality. Breeders could leverage these biomarkers to select cultivars that combine

superior resilience with desirable sensory profiles. For instance, Catiguá MG2's distinct metabolic profile—marked by elevated concentrations of theobromine, rosmarinic acid, and gibberellin A4—reflects its adaptability to drought and its superior flavor characteristics, as confirmed in previous studies (Reichel et al. 2023).

By establishing the metabolic underpinnings of *C. arabica* cultivar differentiation, this study demonstrates how leaf metabolomics can support coffee breeding programs, improve agricultural practices, and ensure the future of high-quality coffee production. Moving forward, combining metabolomics with agronomic and ecological data will be essential for developing strategies to enhance coffee sustainability and productivity in the face of evolving industry demands and environmental challenges.

STATEMENTS AND DECLARATIONS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

Financial support was provided by the Brazilian agencies: National Counsel of Technological and Scientific Development- CNPq, the Brazilian Consortium Coffee Research and Development, the National Institute of Coffee Science and Technology (INCT Café/CNPq), Coordination for the Improvement of Higher Education Personnel (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES) (Funding code: 001), and the Foundation for Research Support of the State of Minas Gerais (FAPEMIG). Financial support by the Access to Research Infrastructures, Horizon2020 Programme of the European Union (EPPN2020 Grant Agreement 731013) and Foundation for Science and Technology (FCT) from Portugal, Research Unit LEAF (UID/AGR/04129/2020) and UCIBIO (FCT UIDP/04378/2020; FCT UIDB/04378/2020).

REFERENCES

Ahlawat YK, Singh M, Manorama K, Nita L, Zaid A, Zulfiqar F (2024) Plant phenolics: neglected secondary metabolites in plant stress tolerance. Braz. J. Bot 47: 703–721. https://doi.org/10.1007/s40415-023-00949-x.

Amalia F, Irifune T, Takegami T, Yusianto, Sumirat U, Putri SP, Fukusaki E (2023) Identification of potential quality markers in Indonesia's Arabica specialty coffee using

GC/MS-based metabolomics approach. Metabolomics 19: 1-11. https://doi.org/10.1007/s11306-023-02051-5.

Bastian F, Hutabarat OS, Dirpan A, Nainu F, Harapan H, Emran TB, Simal-Gandara J (2021) From Plantation to Cup: Changes in Bioactive Compounds during Coffee Processing. Foods. https://doi.org/10.3390/ foods10112827.

Bollen R, Rojo-Poveda O, Verleysen L, Ndezu R, Tshimi EA, Mavar H, Ruttink T, Honnay O, Stoffelen P, Stévigny C, Souard F, Delporte C (2024). Metabolite profiles of green leaves and coffee beans as predictors of coffee sensory quality in Robusta (*Coffea canephora*) germplasm from the Democratic Republic of the Congo. Applied Food Research, 4(2): 100560. https://doi.org/10.1016/j.afres.2024.100560.

Caixeta ET, Resende MDV, Alkimim ER, Sousa TV, de Oliveira ACB, Pereira AA, Alves RS (2022) Aceleração do melhoramento do cafeeiro via seleção genômica: agilidade e eficácia no lançamento de novas cultivares, Documentos 17. Embrapa Café, Brasília.

Cangeloni L, Bonechi C, Leone G, Consumi M, Andreassi M, Magnani A, Rossi C, Tamasi G (2022) Characterization of Extracts of Coffee Leaves (*Coffea arabica* L.) by Spectroscopic and Chromatographic/Spectrometric Techniques. Foods. https://doi.org/10.3390/ foods11162495.

Carréra JC, De Souza RR, Gama Batista AC, Campolina GA, Da Silva Júnior FG, Gavilanes ML, Guimarães RJ, Cardoso M das G, Mori FA (2024). Using underutilized residues of coffee to obtain valuable dietary and antioxidant bioactive compounds. Journal of the Science of Food and Agriculture 104(5): 2660-2668. https://doi.org/10.1002/jsfa.13151.

Castro-Moretti FR, Cocuron J-C, Castillo-Gonzalez H, Escudero-Leyva E, Chaverri P, Guerreiro-Filho O, Slot JC, Alonso AP (2023) A metabolomic platform to identify and quantify polyphenols in coffee and related species using liquid chromatography-mass spectrometry. Front. Plant Sci. https://doi.org/10.3389/fpls.2022.1057645.

Chekol H, Warkineh B, Shimber T, Mierek-Adamska A, Dabrowska, GB, Degu A (2024) Drought Stress Responses in Arabica Coffee Genotypes: Physiological and Metabolic Insights. Plants 13: 828. https://doi.org/10.3390/plants13060828.

Chen X (2018) A review on coffee leaves: Phytochemicals, bioactivities, and applications, Critical Reviews. Food Science and Nutrition 59(6): 1008-1025. https://doi.org/10.1080/10408398.2018.1546667.

Cherubino Ribeiro TH, de Oliveira RR, das Neves TT, Santiago WD, Mansur BL, Saczk AA, Vilela de Resende ML, Chalfun-Junior A (2023) Metabolic Pathway Reconstruction Indicates the Presence of Important Medicinal Compounds in *Coffea* Such as L-DOPA. International Journal of Molecular Sciences. https://doi.org/10.3390/ ijms241512466.

Cuadros-Inostroza Á, Caldana C, Redestig H, Kusano M, Lisec J, Peña-Cortés H, Willmitzer L, Hannah MA (2009) TargetSearch - a Bioconductor package for the efficient preprocessing of GC-MS metabolite profiling data. BMC Bioinformatics 10: 1-12. https://doi.org/10.1186/1471-2105-10-428.

Da Silva JAG, de Resende MLV, Ribeiro IS, Lima AR, Albuquerque LRM, Monteiro ACA, Pereira MHB, Botelho DMdS (2021) Chemical Composition, Production of Secondary Metabolites and Antioxidant Activity in Coffee Cultivars Susceptible and Partially Resistant to Bacterial Halo Blight. Plants. https:// doi.org/10.3390/plants10091915.

De Carvalho CHS, Bartelega L, Sera GH, Matiello JB, Santinato SRAF, Hotz AL (2022). Catálogo de cultivares de café arábica. Documentos 16. Embrapa Café, Brasília.

Desmet S, Morreel K, Dauwe R (2021) Origin and Function of Structural Diversity in the Plant Specialized Metabolome. Plants. https://doi.org/10.3390/ plants10112393.

Dussarrat T, Schweiger R, Ziaja D, Nguyen TT, Krause L, Jakobs R, Eilers E J, Müller C (2023) Influences of chemotype and parental genotype on metabolic fingerprints of tansy plants uncovered by predictive metabolomics. Scientific Reports. https://doi.org/10.1038/s41598-023-38790-7.

Empresa de Pesquisa Agropecuária de Minas Gerais – EPAMIG (2019) Cultivares de café – Programa de Melhoramento Genético EPAMIG, EMBRAPA Café, UFV e UFLA. Epamig, Minas Gerais.

Erban A, Schauer N, Fernie AR, Kopka J (2007) Nonsupervised Construction and Application of Mass Spectral and Retention Time Index Libraries From Time-of-Flight Gas Chromatography-Mass Spectrometry Metabolite Profiles. In: Weckwerth W. (ed)

Metabolomics - Methods in Molecular Biology™, v. 358. Humana Press, pp 19-38. https://doi.org/10.1007/978-1-59745-244-1_2.

Gamboa-Becerra R, Hernández-Hernández MC, González-Ríos Ó, Suárez-Quiroz ML, Gálvez-Ponce E, Ordaz-Ortiz JJ, Winkler R (2019) Metabolomic Markers for the Early Selection of *Coffea canephora* Plants with Desirable Cup Quality Traits. Metabolites 9: 214. https://doi.org/10.3390/metabo9100214.

Guimarães PS, Schenk JCM, Cintra LC, Giachetto PF, Silvarolla MB, Padilha L, Maluf MP (2021) Large-scale prospection of genes on caffeine-free Coffea arabica plants – Discovery of novel markers associated with development and secondary metabolism. Plant Gene, 27: 100314. https://doi.org/10.1016/j.plgene.2021.100314.

Fazuoli LC, de Carvalho HS, Carvalho GR et al (2007) CULTIVARES DE CAFÉ ARÁBICA (*Coffea arabica* L.),125-198. In: Carvalho CHS (ed) Cultivares de café. 2007. Embrapa, Brasília, pp. 125-198.

Hummel J, Strehmel N, Bölling C, Schmidt S, Walther D, Kopka J (2013) Mass spectral search and analysis using the golm metabolome database. In: Weckwerth W, Kahl G (eds). The Handbook of Plant Metabolomics. Wiley-VCH Verlag GmbH & Co. KGaA, pp 321–343. https://doi.org/10.1002/9783527669882.ch18.

International Coffee Organization (2024) Trade Statistic Tables, Production, Cofffee Trade Status. http://www.ico.org/ trade_statistics.asp. Acessed 20 June 2024.

Jin JQ, Yao MZ, Ma CL, Ma JQ, Chen L (2016) Natural allelic variations of TCS1 play a crucial role in caffeine biosynthesis of tea plants and their related species. Plant Physiology and Biochemistry 100: 18-26. https://doi.org/10.1016/j.plaphy.2015.12.020.

Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M (2023) KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Research, v. 51 (D1): D587-D592.

Kind T, Wohlgemuth G, Lee DY, Lu Y, Palazoglu M, Shahbaz S, Fiehn O (2009) FiehnLib: Mass Spectral and Retention Index Libraries for Metabolomics Based on Quadrupole and Time-of-Flight Gas Chromatography/Mass Spectrometry. Analytical Chemistry. 81 (24): 10038-10048. https://doi.org/10.1021/ac9019522.

Kitashova A, Brodsky V, Chaturvedi P, Pierides I, Ghatak A, Weckwerth W, Nägele T (2023) Quantifying the impact of dynamic plant-environment interactions on metabolic regulation. Journal of Plant Physiology 290: 154116. https://doi.org/10.1016/j.jplph.2023.154116.

Kumar N, Pruthi V (2014) Potential applications of ferulic acid from natural sources. Biotechnology Reports 4: 86-93. https://doi.org/10.1016/j.btre.2014.09.002.

Lee K, Kang K, Park M, Park S, Back K (2009) Enhanced octopamine synthesis through the ectopic expression of tyrosine decarboxylase in rice plants. Plant Science. 176: 46-50. https://doi.org/10.1016/j.plantsci.2008.09.006.

Liu Q, Gao T, Liu W, Liu Y, Zhao Y, Liu Y, Li W, Ding K, Ma F, Li C (2020) Functions of dopamine in plants: a review. Plant Signal Behav. https://doi.org/10.1080/15592324.2020.1827782.

Montis A, Souard F, Delporte C, Stoffelen P, Stévigny C, Van Antwerpen P (2022) Targeted and Untargeted Mass Spectrometry-Based Metabolomics for Chemical Profiling of Three Coffee Species. Molecules. https://doi.org/ 10.3390/molecules27103152.

Montis A, Delporte C, Noda Y, Stoffelen P, Stévigny C, Hermans C, Van Antwerpen P, Souard Florence (2024) Targeted metabolomics and transcript profiling of methyltransferases in three coffee species. Plant Science. https://doi.org/10.1016/j.plantsci.2024.112117.

Marques V, Farah A (2009) Chlorogenic acids and related compounds in medicinal plants and infusions. Food Chemistry 113(4): 1370-1376. https://doi.org/10.1016/j.foodchem.2008.08.086.

Martins SCV, Araújo WL, Tohge T, Fernie AR, DaMatta FM (2014) In High-Light-Acclimated Coffee Plants the Metabolic Machinery Is Adjusted to Avoid Oxidative Stress Rather than to Benefit from Extra Light Enhancement in Photosynthetic Yield. PLOS ONE 9(4): e94862. https://doi.org/10.1371/journal.pone.0094862.

Mihai RA, Ortiz-Pillajo DC, Iturralde-Proaño KM, Vinueza-Pullotasig MY, Sisa-Tolagasí LA, Villares-Ledesma ML, Melo-Heras EJ, Cubi-Insuaste NS, Catana RD (2024) Comprehensive

Assessment of Coffee Varieties (*Coffea arabica* L.*; Coffea canephora* L.) from Coastal, Andean, and Amazonian Regions of Ecuador; A Holistic Evaluation of Metabolism, Antioxidant Capacity and Sensory Attributes. Horticulturae 10: 200. https://doi.org/10.3390/horticulturae10030200.

Monteiro Â, Colomban S, Azinheira HG, Guerra-Guimarães L, Silva M do C, Navarini L, Resmini M. Dietary Antioxidants in Coffee Leaves: Impact of Botanical Origin and Maturity on Chlorogenic Acids and Xanthones. Antioxidants 9(1): 6. https://doi.org/10.3390/antiox9010006.

Nagay JHC (1999) Café do Brasil: dois séculos de história. Formação econômica 3:17-23.
Oliveira ACB, Pereira AA, Caixeta ET, Resende MDV, Ribeiro MF (2021) Cultivares de café resistentes à ferrugem: alternativa viável para a cafeicultura das Matas de Minas. Embrapa Café, Brasília.

Nani BD, Rosalen PL, Lazarini JG, De Araújo LP, Dos Reis MS, Breseghello I, Cunha TM, De Alencar, SM, Da Silveira NJ, Franchin M (2022) A Study on the Anti-NF-κB, Anti-Candida, and Antioxidant Activities of Two Natural Plant Hormones: Gibberellin A4 and A7. Pharmaceutics 14(7): 1347. https://doi.org/10.3390/pharmaceutics14071347.

Pang Z, Lu Y, Zhou G, Hui F, Xu L, Viau C, Spigelman AF, MacDonald PE, Wishart DS, Li S, Xia J (2024) MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation. Nucleic Acids Research 52 (W1): W398–W406 https://doi.org/10.1093/nar/gkae253.

RCoreTeam (2023) R: a language and environment for statistical computing. version 4.3.1. R Foundation for statistical computing, Vienna, Austria. https://www.R-project.org/. Acessed 27 October 2023.

Reichel T, De Resende MLV, Nadaleti DHS, Santos FO, Botelho CE (2023) Potential of rust-resistant arabica coffee cultivars for specialty coffee production. Bioscience Journal. https://doi.org/10.14393/BJ-v39n0a2023-66.

Ritonga FN, Zhou D, Zhang Y, Song R, Li C, Li, J., & Gao, J. The Roles of Gibberellins in Regulating Leaf Development. Plants, 12(6): 1243. https://doi.org/10.3390/plants12061243.

Salem MA, Jüppner J, Bajdzienko K, Giavalisco P (2016) Protocol: a fast, comprehensive and reproducible one-step extraction method for the rapid preparation of polar and semi-polar metabolites, lipids, proteins, starch and cell wall polymers from a single sample. Plant Methods 12: 45. https://doi.org/10.1186/s13007-016-0146-2.

Salem MA, Perez de Souza L, Serag A, Fernie AR, Farag MA, Ezzat SM, Alseekh S (2020) Metabolomics in the Context of Plant Natural Products Research: From Sample Preparation to Metabolite Analysis. Metabolites. https://doi.org/10.3390/metabo10010037.

Sardans J, Peñuelas J, Rivas-Ubach A (2011) Ecological metabolomics: overview of current developments and future challenges. Chemoecology 21:191–225. https://doi.org/10.1007/s00049-011-0083-5.

Shen S, Zhan C, Yang C, Fernie AR, Luo J (2023) Metabolomics-centered mining of plant metabolic diversity and function: Past decade and future perspectives. Molecular Plant.16: 43–63. https://doi.org/10.1016/j.molp.2022.09.007.

Shohat H, Cheriker H, Kilambi HV, Illouz Eliaz N, Blum S, Amsellem Z, Tarkowská D, Aharoni A, Eshed Y, Weiss D (2021) Inhibition of gibberellin accumulation by water deficiency promotes fast and long-term 'drought avoidance' responses in tomato. New Phytol 232:1985-1998. https://doi.org/10.1111/nph.17709.

Souard F, Delporte C, Stoffelen P, Thévenot EA, Noret N, Dauvergne B, Kauffmann J-M, Van Antwerpen P, Stévigny C (2017) Metabolomics fingerprint of coffee species determined by untargeted-profiling study using LC-HRMS, Food Chemistry, 245: 603-612. https://doi.org/10.1016/j.foodchem. 2017.10.022.

Sumner LW, Amberg A, Barrett D. *et al* (2007). Proposed minimum reporting standards for chemical analysis. Metabolomics 3: 211–221. https://doi.org/10.1007/s11306-007-0082-2.

Wishart DS, Guo A, Oler E et al (2022) HMDB 5.0: the Human Metabolome Database for 2022. Nucleic Acids Research 50 (D1): D622–D631. https://doi.org/10.1093/nar/gkab1062.

Zhai H, Dong W, Fu X, Li G, Hu F (2024) Integration of widely targeted metabolomics and the e-tongue reveals the chemical variation and taste quality of Yunnan Arabica coffee prepared

using different primary processing methods, Food Chemistry. https://doi.org/10.1016/j.fochx.2024.101286.

Zheng X, Ashihara H (2004) Distribution, biosynthesis and function of purine and pyridine alkaloids in *Coffea arabica* seedlings. Plant Sci. 166: 807–813.

Table 1 Classification of the five *Coffea arabica* cultivars studied: genetic origin, resistance to coffee leaf rust (CLR) and nematode, cup quality and coffee production.

Cultivar	**Genetic origin** (1)	**CLR resistance**	**Nematode resistance**	**Cup quality**	**Coffee production**
Catuaí Vermelho IAC 144	Caturra amarelo IAC 476-11 x Mundo Novo IAC 374-19	Susceptible (2,3)	Susceptible (2)	Good (2)	High (2)
Catiguá MG2	Catuaí amarelo IAC 86 x Híbrido de Timor UFV 440-10	Resistant (2,3)	Susceptible (2,4)	Excellent (5)	High (4)
Catiguá MG3	Catuaí amarelo IAC 86 x Híbrido de Timor UFV 440-10	Resistant (2,3)	Resistant (2,4)	Good (4)	High (2,4)
Paraíso 2	Catuaí amarelo IAC 30 x Híbrido de Timor UFV 445-46	Resistant (3)	Susceptible (2)	Excellent (5)	High (2)

Sarchimor MG 8840	Villa Sarchi CIFC 971/10 x Híbrido de Timor CIFC 832/2	Resistant (2)	Susceptible (2)	Regular (2)	High (2)

(1) IAC - Instituto Agronômico de Campinas (Brazil); MG - Minas Gerais (Brazil); CIFC - Centro de Investigação das Ferrugens do Cafeeiro (Portugal); UFV - Universidade Federal de Viçosa (Brazil). (2) De Carvalho et al. (2022); (3) Oliveira et al. (2021); (4) Fazuoli et al. (2007); (5) Empresa de Pesquisa Agropecuária de Minas Gerais (2019).

Table 2 Classification of the statistically significant metabolic features distinguishing the five *C. arabica* cultivars**.**

Metabolic Feature[(a)]	**p-value[(b)]**	**RI interval[(c)]**	**Linked metabolite**	**Metabolic class**
1-Pyrroline-2-carboxylate (1TMS)	< 0.01	374286 - 374526	1-Pyrroline-2-carboxylate	Amino acids
Butanoic acid, 4-amino-3-hydroxy- (3TMS)	< 0.001	425294 - 426562	4-Amino-3-hydroxybutanoate	Amino acids
Cystathionine (4TMS)	< 0.001	753045 - 754956	L-Cystathionine	Amino acids
Tryptophan (2TMS)_1	< 0.001	790065 - 790537	Tryptophan	Amino acids
Tryptophan (2TMS)_2	< 0.01	814376 - 814681	Tryptophan	Amino acids
Glutamic acid, N-acetyl- (3TMS)	< 0.01	579209 - 581005	N-Acetyl-L-glutamate	Amino acids
Hippuric acid (1TMS)	< 0.01	692005 - 693045	Hippuric acid	Amino acids
Leucine (2TMS)	< 0.01	307233 - 307462	L-Leucine	Amino acids
Orotic acid, 4,5-dihydro- (3TMS)	< 0.001	615528 - 616572	L-Dihydroorotate	Amino acids
Tyrosine, 3-nitro- (3TMS)	< 0.01	795791 - 796300	3-Nitrotyrosine	Amino acids

4-Hydroxyphenyl-beta-glucopyranoside (5TMS)	< 0.001	871143 - 871361	Arbutin	Sugar
Arabitol (5TMS)	< 0.001	501789 - 503019	D-Arabitol	Sugar
beta-D-Fructofuranosyl-(2,1)-beta-D-Fructofuranose (1MEOX) (8TMS) MP	< 0.01	855615 - 856221	Inulobiose	Sugar
Gentiobiose (1MEOX) (8TMS) MP	< 0.001	925546 - 926040	Gentiobiose	Sugar
Gluconic acid-1,5-lactone (4TMS)	< 0.001	624583 - 625875	D-Glucono-1,5-lactone	Sugar
Digitoxose (1MEOX) (3TMS) MP	< 0.01	479899 - 480784	D-Digitoxose	Sugar
Glyceric acid (3TMS)	< 0.01	346482 - 346736	L-Glycerate	Sugar
Digitoxose (1MEOX) (3TMS) BP	< 0.01	479892 - 481076	D-Digitoxose	Sugar
Glycerol (3TMS)	< 0.01	295565 - 295805	Glycerol	Sugar
2-Piperidinecarboxylic acid (2TMS)	< 0.01	374241 - 374353	Pipecolic acid	Alkaloids
Theobromine (1TMS)	< 0.001	785361 - 787089	Theobromine	Alkaloids
Theophylline (1TMS)	< 0.001	737150 - 737800	Theophylline	Alkaloids
Thiazole, 4-methyl-5-hydroxyethyl- (1TMS)	< 0.001	426506 - 426715	4-Methyl-5-(2'-hydroxyethyl)-thiazole	Alkaloids
Tryptamine, 1-methyl- (2TMS)	< 0.001	785205 - 786414	N-Methyltryptamine	Alkaloids
Ferulic acid, trans- (2TMS)	< 0.001	754284 - 755047	Ferulic acid	Phenylpropanoids or phenolic compounds

Hydrocaffeic acid (3TMS)	< 0.001	666766 - 667192	3,4-Dihydroxyphenylpropanoate	Phenylpropanoids or phenolic compounds
Rosmarinic acid (5TMS)	< 0.01	1133997 - 1135298	Rosmarinic acid	Phenylpropanoids or phenolic compounds
Aconitic acid, cis- (3TMS)	< 0.01	581718 - 583537	cis-Aconitate	Other organic acids
Hippuric acid, 2-hydroxy- (3TMS)	< 0.01	685475 - 686300	Salicyluric acid	Other organic acids
Malic acid (3TMS)	< 0.01	442458 - 443060	L-Malate	Other organic acids
Shikimic acid (4TMS)	< 0.001	584778 - 586715	Shikimate	Other organic acids
Phenethylamine (2TMS)	< 0.001	489914 - 490638	Phenethylamine	Amines
Dopamine (4TMS)	< 0.01	692779 - 693045	Dopamine	Amines
Octopamine (4TMS)	< 0.001	666161 - 666481	Octopamine	Amines
Ergosterol (1TMS)	< 0.01	1052029 - 1053269	Ergosterol	Lipids
Gibberellin A4 (2TMS)	< 0.001	943755 - 944850	Gibberellin A4	Lipids
Ascorbic acid (4TMS)	< 0.01	650395 - 651185	L-Ascorbic acid	Others
Dehydroascorbic acid dimer (2MEOX) MP	< 0.01	624221 - 625314	Dehydroascorbic acid	Others
Galactonic acid-1,4-lactone (4TMS)	< 0.001	617495 - 618546	D-Galactono-1,4-lactone	Others
Kynurenine (3TMS)	< 0.01	779498 - 780164	L-Kynurenine	Others

Orotic acid (3TMS)	< 0.01	584880 - 586808	Orotic acid	Others

[(a)] Products of the derivatization process, annotated according to the GOLM specific library; [(b)]ANOVA and Tukey HSD test ($p<0.05$); [(c)] RI (retention index) interval of the metabolic features found in the samples.

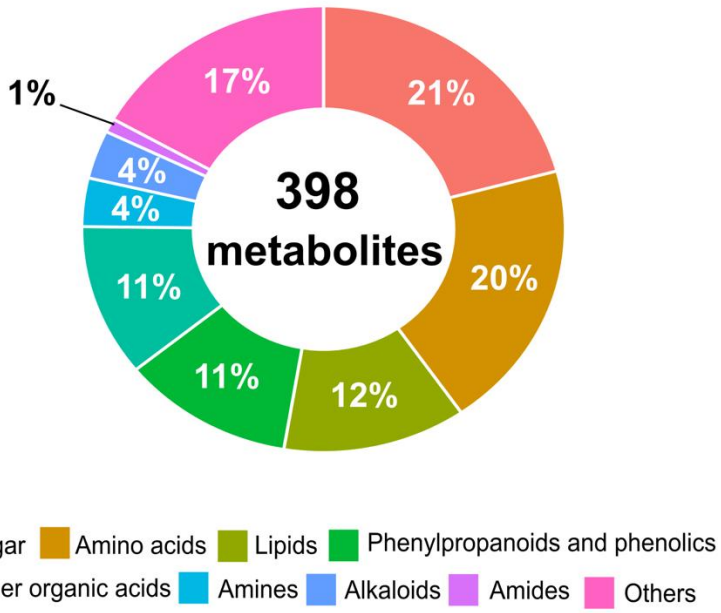


Fig. 1 Distribution of 398 unique metabolites detected in *C. arabica* leaves, categorized into various metabolic classes.

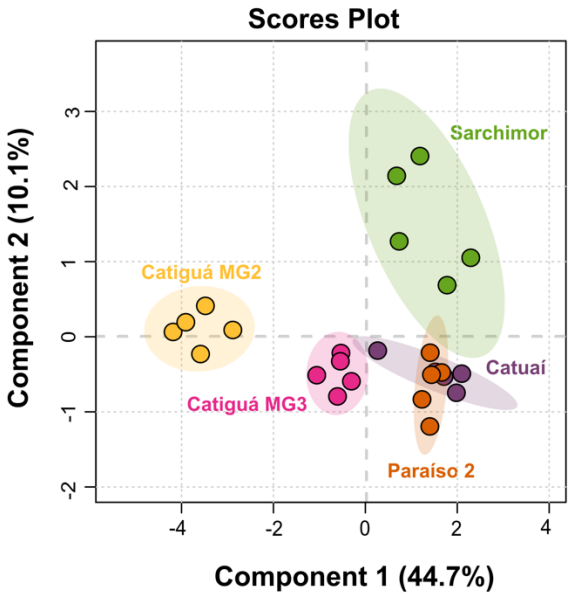


Fig. 2 Leaf non-targeted metabolomic analysis of the five *C. arabica* cultivars. A: Partial Least Squares Discriminant Analysis (PLS-DA) performed for the significant metabolites that change between cultivars, according to the ANOVA and Tukey HSD test ($p < 0.05$).

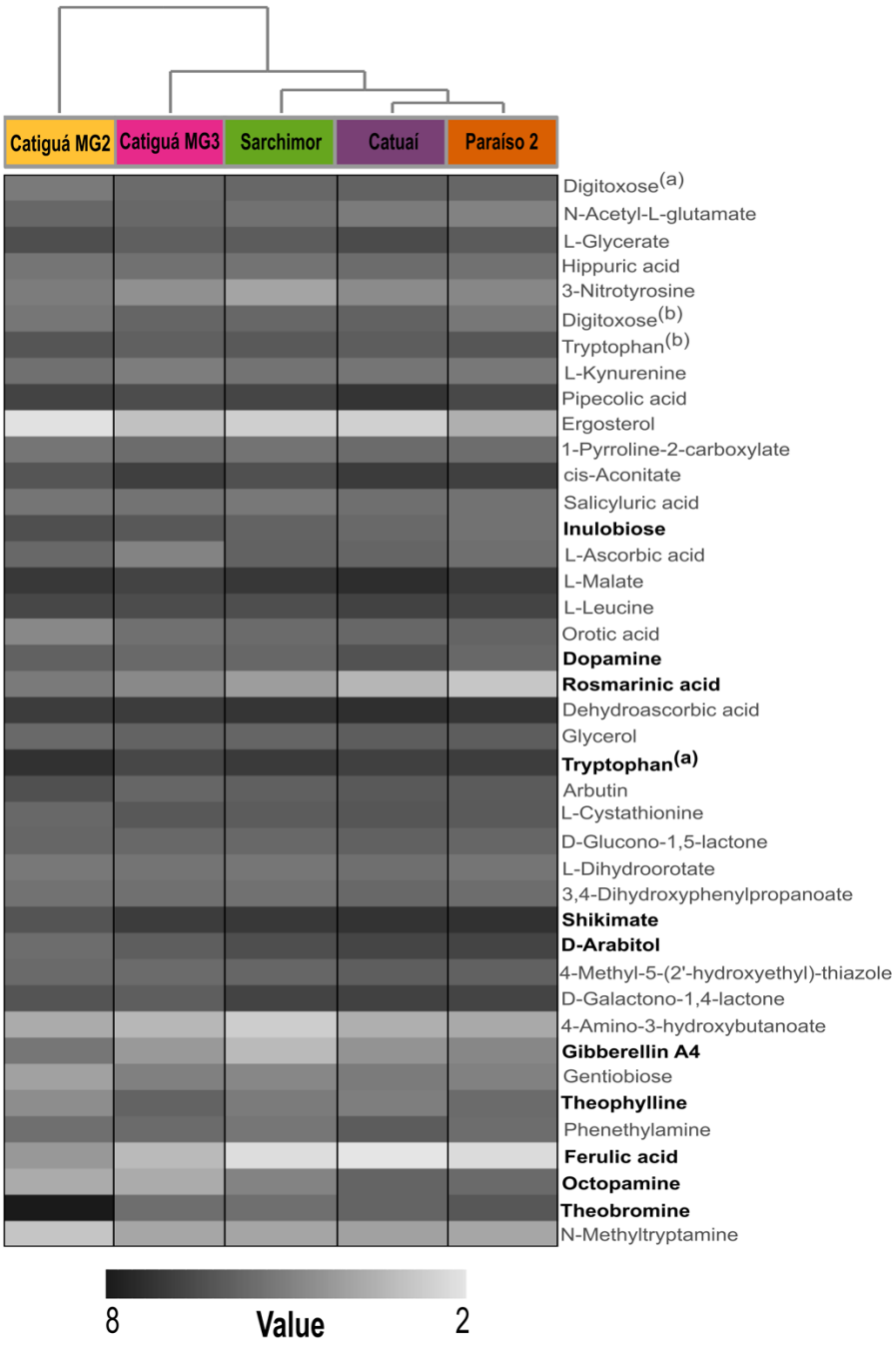


Fig. 3 Heatmap analysis of the metabolites that significantly change between cultivars, according to the ANOVA and Tukey HSD test ($p < 0.05$). The color scale represents metabolic feature relative value, shown in a gradient from black (higher value) to white (lower value). Metabolic features highlighted in bold represent key discriminants for cultivar differentiation based on the PLS-DA model.

SUPPLEMENTARY FIGURE 1

Theoretical and Experimental Plant Physiology (Thematic Collection: Metabolomics and plant metabolism regulation and integration)

NON-TARGETED METABOLOMIC ANALYSIS OF FIELD-GROWN *Coffea arabica* CULTIVARS REVEALS DISTINCT LEAF METABOLIC SIGNATURES

Authors: Jéfyne Campos Carréra[1], Leonor Guerra-Guimarães[2,3], John Charles D'Auria[4], Luana de Jesus Sartori[1], Carla Pinheiro[5,6], Vânia Aparecida Silva[7], Margarete Lordelo Volpato[7], Gladyston Rodrigues Carvalho[7], Fabio Akira Mori[1]

[1]UFLA - Universidade Federal de Lavras, 37200-000 Lavras, Minas Gerais, Brasil.

[2]CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal.

[3]LEAF - Linking Landscape, Environment, Agriculture and Food Research Centre and TERRA Associated Laboratory, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017, Lisboa, Portugal.

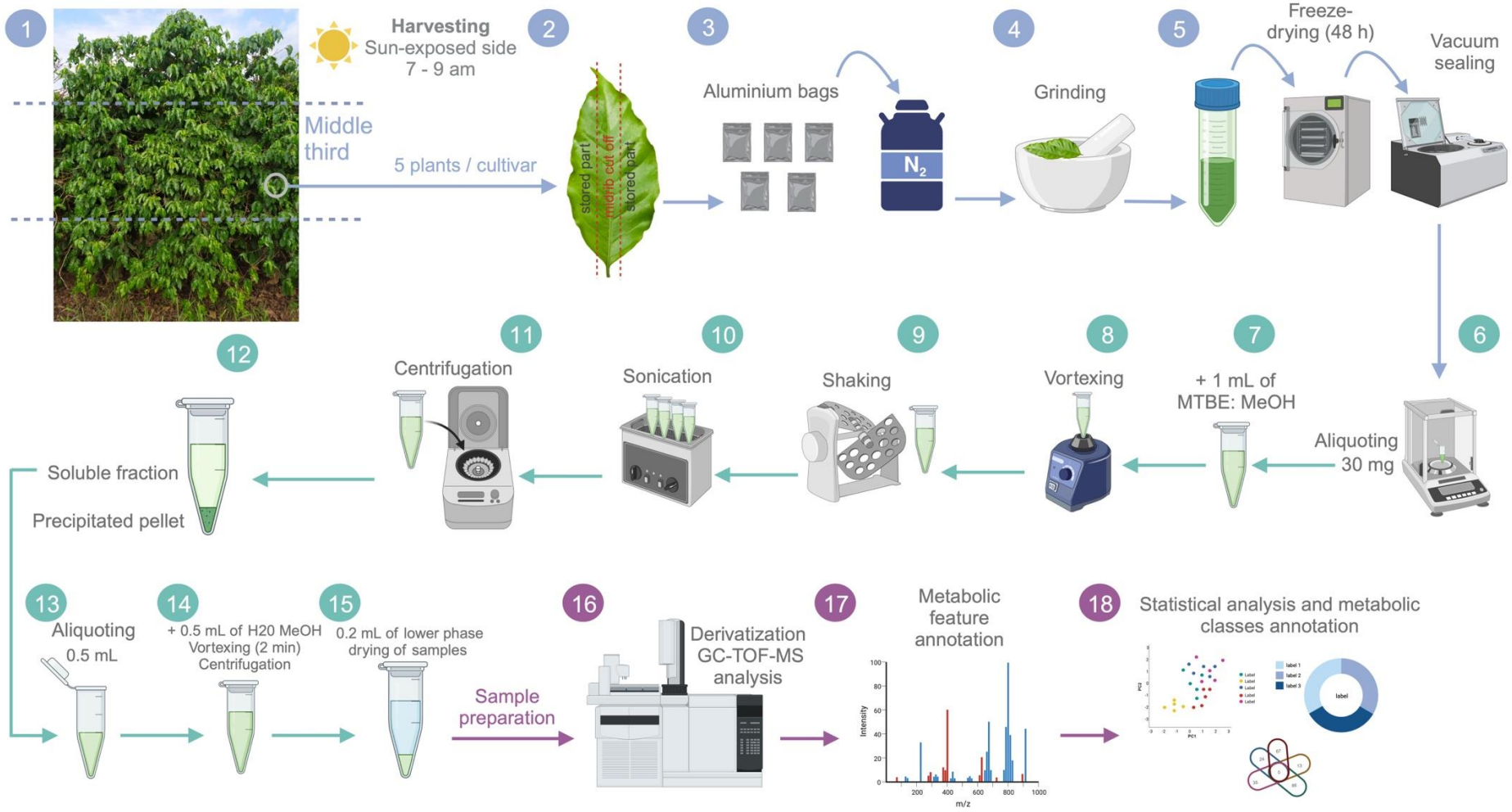
[4]Research Group Metabolic Diversity, Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK Gatersleben), OT Gatersleben, Corrensstraße 3, 06466 Seeland, Germany

[5]UCIBIO Applied Molecular Biosciences Unit, Department of Life Sciences, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.

[6]Associate Laboratory i4HB Institute for Health and Bioeconomy, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.

[7]EPAMIG – Empresa de Pesquisa Agropecuária de Minas Gerais, 37200-000 Lavras, Minas Gerais, Brasil.

Correspondence author: Leonor Guerra-Guimarães, CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal. E-mail: leonorguimaraes@isa.ulisboa.pt. https://orcid.org/0000-0002-9676-1036.



Supplementary Figure 1 Representation of the leaf collection, extraction and metabolomic process. Steps 1 to 5: Leaf collection and sampling. Steps 6 to 15: Preparation of leaf extracts. Steps 16 to 18: Untargeted metabolomics analysis and data processing.

SUPPLEMENTARY FIGURE 2

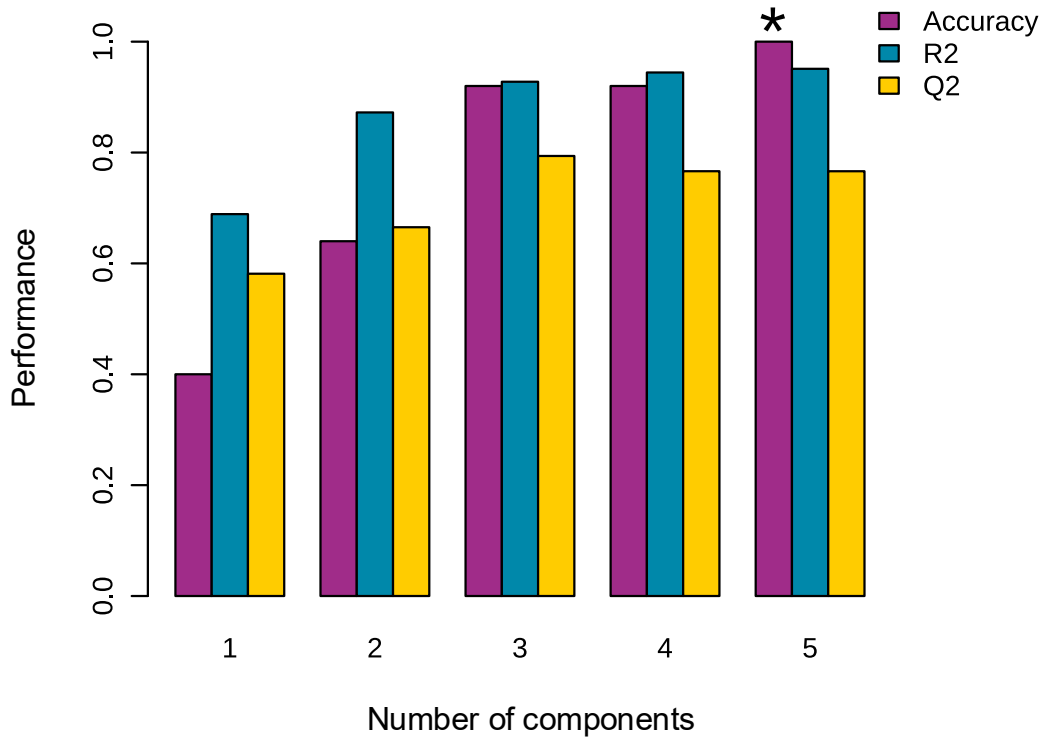
Theoretical and Experimental Plant Physiology (Thematic Collection: Metabolomics and plant metabolism regulation and integration)

NON-TARGETED METABOLOMIC ANALYSIS OF FIELD-GROWN *Coffea arabica* CULTIVARS REVEALS DISTINCT LEAF METABOLIC SIGNATURES

Authors: Jéfyne Campos Carréra[1], Leonor Guerra-Guimarães[2,3], John Charles D'Auria[4], Luana de Jesus Sartori[1], Carla Pinheiro[5,6], Vânia Aparecida Silva[7], Margarete Lordelo Volpato[7], Gladyston Rodrigues Carvalho[7], Fabio Akira Mori[1]

[1]UFLA - Universidade Federal de Lavras, 37200-000 Lavras, Minas Gerais, Brasil.
[2]CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal.
[3]LEAF - Linking Landscape, Environment, Agriculture and Food Research Centre and TERRA Associated Laboratory, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017, Lisboa, Portugal.
[4]Research Group Metabolic Diversity, Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK Gatersleben), OT Gatersleben, Corrensstraße 3, 06466 Seeland, Germany
[5]UCIBIO Applied Molecular Biosciences Unit, Department of Life Sciences, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.
[6]Associate Laboratory i4HB Institute for Health and Bioeconomy, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.
[7]EPAMIG – Empresa de Pesquisa Agropecuária de Minas Gerais, 37200-000 Lavras, Minas Gerais, Brasil.

Correspondence author: Leonor Guerra-Guimarães, CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal. E-mail: leonorguimaraes@isa.ulisboa.pt. https://orcid.org/0000-0002-9676-1036.



Supplementary Figure 2 Cross-validation (5-fold CV) results considering five components of the PLS-DA Model analysis.

SUPPLEMENTARY FIGURE 3

Theoretical and Experimental Plant Physiology (Thematic Collection: Metabolomics and plant metabolism regulation and integration)

NON-TARGETED METABOLOMIC ANALYSIS OF FIELD-GROWN *Coffea arabica* CULTIVARS REVEALS DISTINCT LEAF METABOLIC SIGNATURES

Authors: Jéfyne Campos Carréra[1], Leonor Guerra-Guimarães[2,3], John Charles D'Auria[4], Luana de Jesus Sartori[1], Carla Pinheiro[5,6], Vânia Aparecida Silva[7], Margarete Lordelo Volpato[7], Gladyston Rodrigues Carvalho[7], Fabio Akira Mori[1]

[1]UFLA - Universidade Federal de Lavras, 37200-000 Lavras, Minas Gerais, Brasil.

[2]CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal.

[3]LEAF - Linking Landscape, Environment, Agriculture and Food Research Centre and TERRA Associated Laboratory, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017, Lisboa, Portugal.

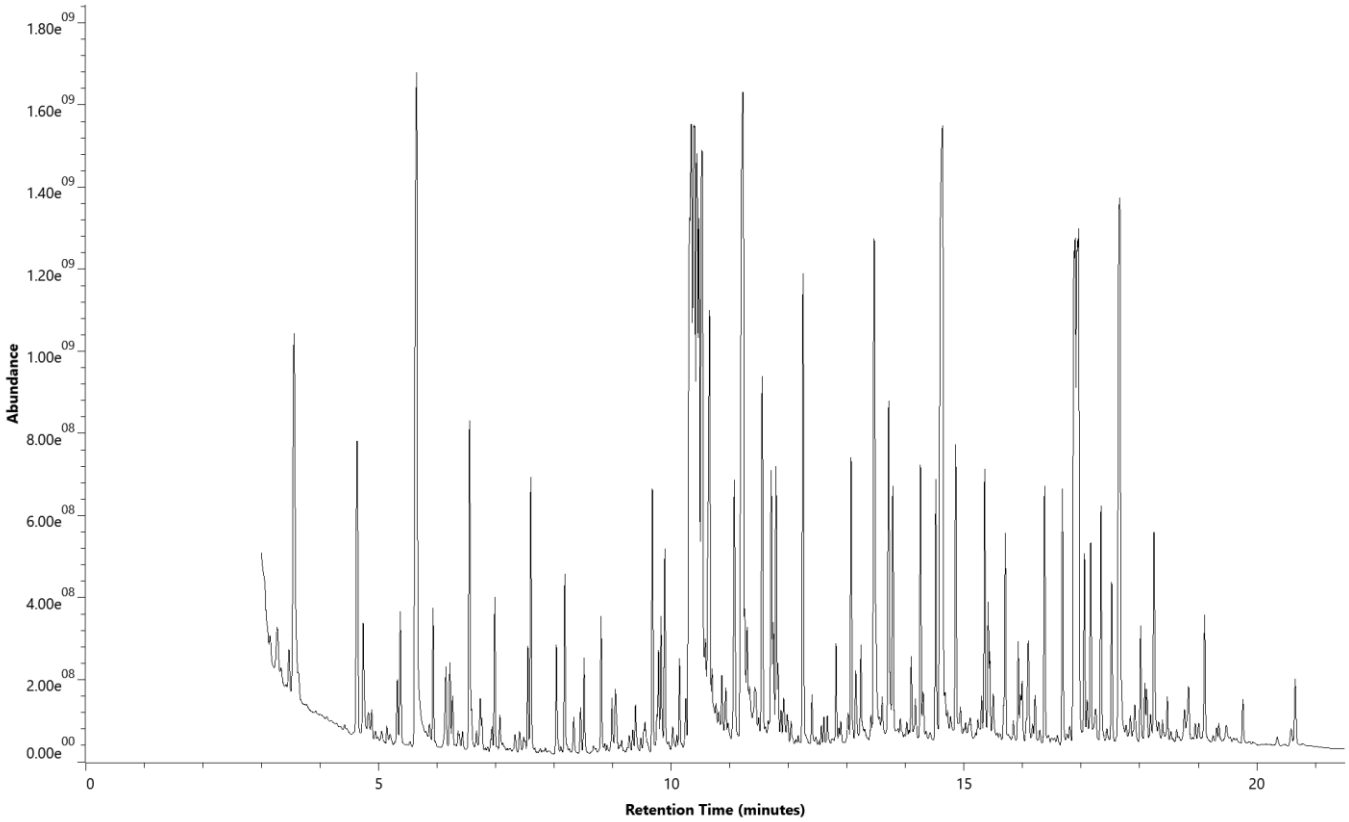
[4]Research Group Metabolic Diversity, Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK Gatersleben), OT Gatersleben, Corrensstraße 3, 06466 Seeland, Germany

[5]UCIBIO Applied Molecular Biosciences Unit, Department of Life Sciences, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.

[6]Associate Laboratory i4HB Institute for Health and Bioeconomy, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.

[7]EPAMIG – Empresa de Pesquisa Agropecuária de Minas Gerais, 37200-000 Lavras, Minas Gerais, Brasil.

Correspondence author: Leonor Guerra-Guimarães, CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal. E-mail: leonorguimaraes@isa.ulisboa.pt. https://orcid.org/0000-0002-9676-1036.



Supplementary Figure 3 Total Ion Chromatogram of *Coffea arabica* leaf in Patrocínio (Cerrado Mineiro), Minas Gerais, Brazil.

SUPPLEMENTARY FIGURE 4

Theoretical and Experimental Plant Physiology (Thematic Collection: Metabolomics and plant metabolism regulation and integration)

NON-TARGETED METABOLOMIC ANALYSIS OF FIELD-GROWN *Coffea arabica* CULTIVARS REVEALS DISTINCT LEAF METABOLIC SIGNATURES

Authors: Jéfyne Campos Carréra[1], Leonor Guerra-Guimarães[2,3], John Charles D'Auria[4], Luana de Jesus Sartori[1], Carla Pinheiro[5,6], Vânia Aparecida Silva[7], Margarete Lordelo Volpato[7], Gladyston Rodrigues Carvalho[7], Fabio Akira Mori[1]

[1]UFLA - Universidade Federal de Lavras, 37200-000 Lavras, Minas Gerais, Brasil.

[2]CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal.

[3]LEAF - Linking Landscape, Environment, Agriculture and Food Research Centre and TERRA Associated Laboratory, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017, Lisboa, Portugal.

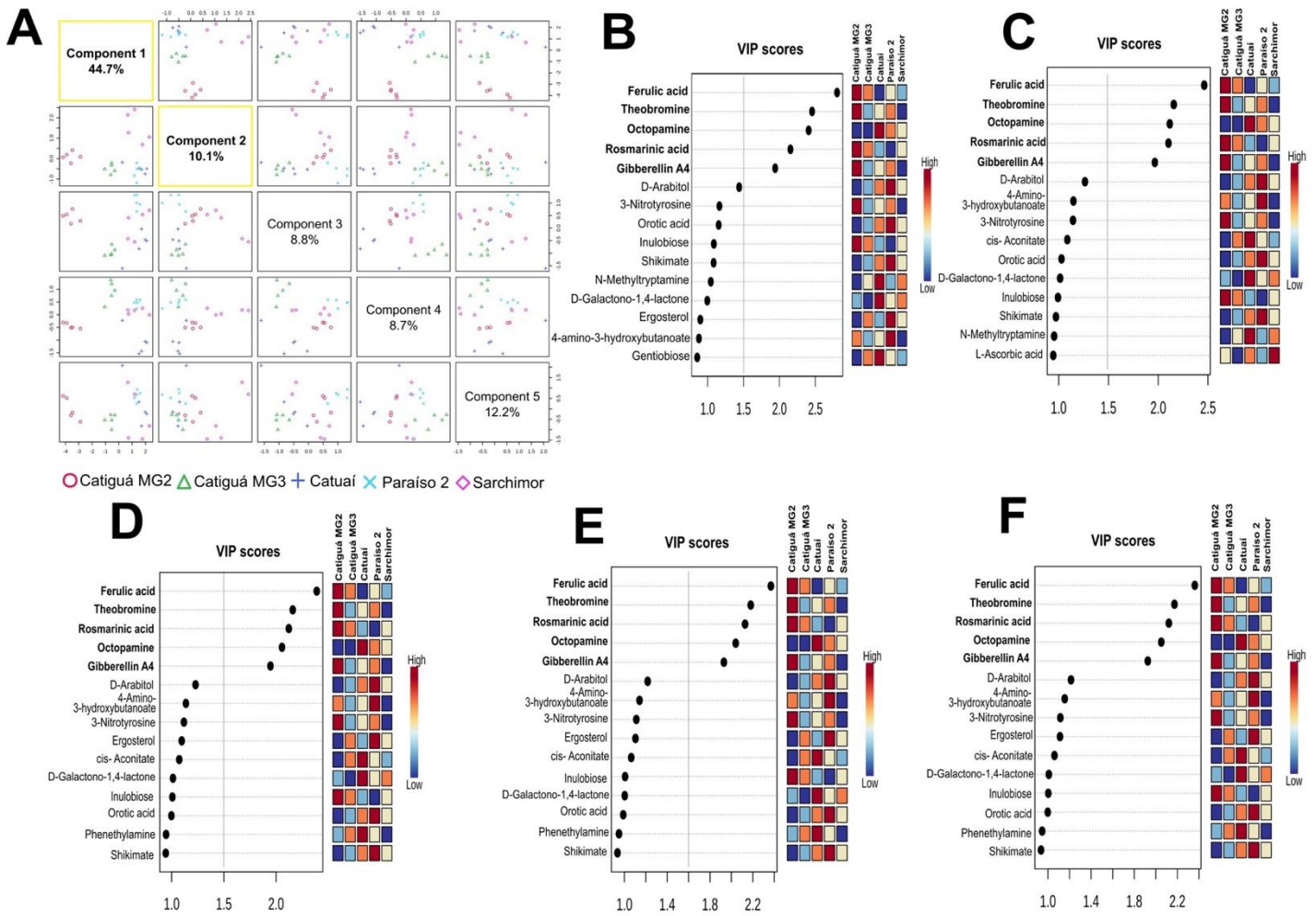
[4]Research Group Metabolic Diversity, Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK Gatersleben), OT Gatersleben, Corrensstraße 3, 06466 Seeland, Germany

[5]UCIBIO Applied Molecular Biosciences Unit, Department of Life Sciences, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.

[6]Associate Laboratory i4HB Institute for Health and Bioeconomy, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.

[7]EPAMIG – Empresa de Pesquisa Agropecuária de Minas Gerais, 37200-000 Lavras, Minas Gerais, Brasil.

Correspondence author: Leonor Guerra-Guimarães, CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal. E-mail: leonorguimaraes@isa.ulisboa.pt. https://orcid.org/0000-0002-9676-1036.



Supplementary Figure 4 PLS-DA Model analysis. A: Percentual values of variance explanation of the five components. B: Variable Importance Plot of Component 1. C: Variable Importance Plot of Component 2. D: Variable Importance Plot of Component 3. E: Variable Importance Plot of Component 4. F: Variable Importance Plot of Component 5. (*) Indication of the highest value of accuracy. R2:

Coefficient of determination. Q2: Predictive relevance, based on sum of squared erros (Szymańska et al. 2012). Metabolic features highlighted in bold were identified as key discriminants for distinguishing between cultivars based on the PLS-DA model. Reference: Szymańska E, Saccenti E, Smilde AK, Westerhuis JA (2012) Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. Metabolomics 8(1):3-16. doi: 10.1007/s11306-011-0330-3

SUPPLEMENTARY TABLE 1

Theoretical and Experimental Plant Physiology (Thematic Collection: Metabolomics and plant metabolism regulation and integration)

NON-TARGETED METABOLOMIC ANALYSIS OF FIELD-GROWN *Coffea arabica* CULTIVARS REVEALS DISTINCT LEAF METABOLIC SIGNATURES

Authors: Jéfyne Campos Carréra[1], Leonor Guerra-Guimarães[2,3], John Charles D'Auria[4], Luana de Jesus Sartori[1], Carla Pinheiro[5,6], Vânia Aparecida Silva[7], Margarete Lordelo Volpato[7], Gladyston Rodrigues Carvalho[7], Fabio Akira Mori[1]

[1]UFLA - Universidade Federal de Lavras, 37200-000 Lavras, Minas Gerais, Brasil.

[2]CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal.

[3]LEAF - Linking Landscape, Environment, Agriculture and Food Research Centre and TERRA Associated Laboratory, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017, Lisboa, Portugal.

[4]Research Group Metabolic Diversity, Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK Gatersleben), OT Gatersleben, Corrensstraße 3, 06466 Seeland, Germany

[5]UCIBIO Applied Molecular Biosciences Unit, Department of Life Sciences, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.

[6]Associate Laboratory i4HB Institute for Health and Bioeconomy, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.

[7]EPAMIG – Empresa de Pesquisa Agropecuária de Minas Gerais, 37200-000 Lavras, Minas Gerais, Brasil.

Correspondence author: Leonor Guerra-Guimarães, CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal. E-mail: leonorguimaraes@isa.ulisboa.pt. https://orcid.org/0000-0002-9676-1036.

Supplementary Table 1. Soil chemical characterization of the experimental locations made at depth of 0–20 cm.

pH	**N**	**K**	**P**	**Na**	**Ca**	**Mg**	**Al**	**H+Al**	**SB**	**e**	**CEC**
(H_2O)	gKg^{-1}	------ mg dm^{-3}------			---------- $cmol_c$ dm^{-3}----------						
5.1	3.08	111.9	13.7	3.0	3.9	1.1	0.10	3.9	5.3	5.4	9.2
BS	**m**	**OM**		**P-rem**	**Zn**	**Fe**	**Mn**	**Cu**	**B**	**S**	
- (%) -		dag kg^{-1}		mg L^{-1}	---------- mg dm^{-3}----------						
57.8	1.86	3.78		11.75	6.7	17.9	26.7	5.9	0.18	36.9	
Soil texture (dag kg^{-1})											
Clay			**Sand**				**Silt**				
41.5			20.5				38.0				

pH in water (1:2.5); Ca^{2+} – Mg^{2+} – Al^{3+}– extractor: KCl, 1mol L^{-1}; P- Na - K- Fe - Zn- Mn- Cu – Mehlich-1 extractor; B: Hot water extraction; S - Extractor - Monocalcium phosphate in acetic acid; H + Al- Extrator: SMP; SB: sum of bases; e: effective cation exchange capacity; CEC: Cation Exchange Capacity at pH 7.0; BS: base saturation index; m: aluminum saturation index; OM: organic matter (Oxidation: Na2Cr2O7 4N+ H2SO4 10N); P-rem: remaining phosphorus.

SUPPLEMENTARY TABLE 2

Theoretical and Experimental Plant Physiology (Thematic Collection: Metabolomics and plant metabolism regulation and integration)

NON-TARGETED METABOLOMIC ANALYSIS OF FIELD-GROWN *Coffea arabica* CULTIVARS REVEALS DISTINCT LEAF METABOLIC SIGNATURES

Authors: Jéfyne Campos Carréra1, Leonor Guerra-Guimarães2,3, John Charles D'Auria4, Luana de Jesus Sartori1, Carla Pinheiro5,6, Vânia Aparecida Silva7, Margarete Lordelo Volpato7, Gladyston Rodrigues Carvalho7, Fabio Akira Mori1

1UFLA - Universidade Federal de Lavras, 37200-000 Lavras, Minas Gerais, Brasil.
2CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal.
3LEAF - Linking Landscape, Environment, Agriculture and Food Research Centre and TERRA Associated Laboratory, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017, Lisboa, Portugal.
4Research Group Metabolic Diversity, Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK Gatersleben), OT Gatersleben, Corrensstraße 3, 06466 Seeland, Germany
5UCIBIO Applied Molecular Biosciences Unit, Department of Life Sciences, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.
6Associate Laboratory i4HB Institute for Health and Bioeconomy, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.
7EPAMIG – Empresa de Pesquisa Agropecuária de Minas Gerais, 37200-000 Lavras, Minas Gerais, Brasil.

Correspondence author: Leonor Guerra-Guimarães, CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal. E-mail: leonorguimaraes@isa.ulisboa.pt. https://orcid.org/0000-0002-9676-1036.

Supplementary Table 2 All 398 Metabolic features, corresponding metabolites and metabolic classes present in coffee leaves (*Coffea arabica* L.) field-grown in Cerrado Mineiro, Minas Gerais, Brasil.

Feature	Enrichment_name	Suggested_class
Spermine (5TMS) MP	Spermine	Amines
Acetic acid, 3-hydroxyphenyl- (3TMS)	3-Hydroxyphenylacetate	Other organic acids
Ribulose (1MEOX) (4TMS) BP1	D-Ribulose	Sugar
Abscisic acid (1MEOX) (1TMS)_1	(+)-Abscisic acid	Lipids
Histidine, N-tau-methyl- (3TMS)	1-Methylhistidine	Amino acids
Glucose, 2-amino-2-deoxy- (1MEOX) (5TMS) MP	2-Amino-2-deoxy-D-glucose	Sugar
Propane-1,3-diol, 2-amino-2-methyl- (3TMS)	2-Amino-2-methyl-1,3-propandiol	Others
Norbornane-2-carboxylic acid, 2-amino- (2TMS) MP	2-Aminobicyclo[2.2.1]heptane-2-carboxylic acid	Lipids
Propanoic acid, 2-amino-2-methyl-3-hydroxy- (2TMS)	2-Methylserine	Amino acids
Adenine (3TMS)	Adenine	Others
Pyruvic acid, 3-hydroxy- (1MEOX) (2TMS) BP	3-Hydroxypyruvic acid	Other organic acids
Malic acid, 3-oxalo- (TMS)	3-Oxalomalate	Other organic acids
Pyruvic acid, 4-hydroxyphenyl- (1MEOX) (2TMS) BP	4-Hydroxyphenylpyruvate	Phenylpropanoids or phenolic compounds
Isocaproic acid, 2-oxo- (1MEOX) (1TMS) BP	4-Methyl-2-oxopentanoate	Other organic acids
Tryptamine, 5-methoxy- (2TMS) BP	5-Methoxytryptamine	Amines
Asparagine (4TMS) BP2	L-Asparagine	Amino acids
Purine, 6-benzylamino-, 9-beta-D-glucopyranosyl- (5TMS)	6-Benzylaminopurine 9-(beta-D-glucoside)	Others
Digitoxose (1MEOX) (3TMS) MP	D-Digitoxose	Sugar
Erythrose (1MEOX) (3TMS) MP	D-Erythrose	Sugar
Fructose-1-phosphate (6TMS) MP	D-Fructose 1-phosphate	Sugar
Glucose (1MEOX) (5TMS) BP	D-Glucose	Sugar
Glyceraldehyde (1MEOX) (2TMS) BP	D-Glyceraldehyde	Sugar
Psicose (1MEOX) (5TMS) BP	D-Psicose	Sugar
Butanoic acid, 4-amino- (3TMS)	gamma-Aminobutyric acid	Amino acids
Glycine (3TMS)	Glycine	Amino acids
Hippuric acid (1TMS)	Hippuric acid	Amino acids
Indole-3-acetamide (3TMS)	Indole-3-acetamide	Others
Butyric acid, 2,4-diamino-, DL- (3TMS)	2,4-Diaminobutyric acid	Amino acids
Cystathionine (4TMS)	L-Cystathionine	Amino acids
Orotic acid, 4,5-dihydro- (4TMS)	L-Dihydroorotate	others
Galactose (1MEOX) (5TMS) MP	L-Galactose	Sugar
Gulose (1MEOX) (5TMS) BP	D-Gulose	Sugar

Glutamine [-H2O] (3TMS) MP	L-Glutamine	Amino acids
Histidine (4TMS)	L-Histidine	Amino acids
Homoserine (3TMS)	L-Homoserine	Amino acids
Kynurenine (3TMS)	L-Kynurenine	Others
Lipoamide, alpha- (1TMS)	Lipoamide	Lipids
Malonic acid (2TMS)	Malonate	Other organic acids
Melibiose (1MEOX) (8TMS) MP	Melibiose	Sugar
Glutamic acid, N-acetyl- (3TMS)	N-Acetyl-L-glutamate	Amino acids
Aspartic acid, N-(aminocarbonyl)- (4TMS)	N-Carbamoyl-L-aspartate	Amino acids
Oxaloacetate (1MEOX) (3TMS) MP	Oxaloacetate	Other organic acids
Pantothenic acid, D- (4TMS)	Pantothenate	others
Phylloquinone	Phylloquinone	Lipids
Piceatannol (4TMS) MP	Piceatannol	Phenylpropanoids or phenolic compounds
Putrescine (4TMS)	Putrescine	Amines
Pyruvic acid (2TMS)	Pyruvate	Other organic acids
Senecionine (1TMS) BP2	Senecionine	Alkaloids
Sophorose (1MEOX) (8TMS) MP	Sophorose	Sugar
Spermidine [?] (5TMS)	Spermidine	Amines
Tartronic acid (4TMS)	Tartronic acid	Other organic acids
Taxifolin (1MEOX) (5TMS) MP	Taxifolin	Phenylpropanoids or phenolic compounds
Tetrahydroalstonine (1TMS)	Tetrahydroalstonine	Alkaloids
Octopamine (4TMS)	Octopamine	Phenylpropanoids or phenolic compounds
Saccharin (1TMS)_2	Saccharin	Others
Tryptophan (2TMS)_1	Tryptophan	Amino acids
Tartronic acid, 2-(methylaminomethyl)- (1MEOX) (3TMS)	2-(Methylaminomethyl)tartronic acid	Others
Acetic acid, 2-hydroxyphenyl- (2TMS)	2-Hydroxyphenylacetate	Other organic acids
Guanosine-2*,3*-cyclic-monophosphate (4TMS)	2',3'-Cyclic GMP	Others
Guanosine-3*,5*-cyclic-monophosphate (4TMS)	3',5'-Cyclic GMP	Others
alpha-D-Glucopyranosyl-(1,6)-D-glucitol (9TMS)	Isomaltitol	Sugar
Arabinonic acid-1,4-lactone (3TMS)	D-Arabinono-1,4-lactone	Sugar
Ribonic acid-1,4-lactone (3TMS)	Ribonolactone	Sugar
beta-D-Fructofuranosyl-(2,1)-beta-D-Fructofuranose (1MEOX) (8TMS) MP	Inulobiose	Sugar
Epicatechin (5TMS)	Epicatechin	Phenylpropanoids or phenolic compounds
Cellobiitol (9TMS)	Cellobiotol	Sugar
Inositol, allo- (6TMS)	allo-Inositol	Sugar
Inositol, myo- (6TMS)	Inositol	Sugar

Butanoic acid, 2,4-diamino- (4TMS)	L-2,4-Diaminobutanoate	Amino acids
Cinnamic acid, 3-methoxy-, trans- (1TMS)	3-Methoxycinnamic acid	Phenylpropanoids or phenolic compounds
Mannose (1MEOX) (5TMS) BP	D-Mannose	Sugar
Ornithine (3TMS)	DL-Ornithine	Amino acids
Ornithine (4TMS)	L-Ornithine	Amino acids
Diethanolamine (3TMS)	Diethanolamine	Amines
Dopamine (4TMS)	Dopamine	Amines
Ethanol, 2-(4-hydroxyphenyl)- (2TMS)	4-Hydroxyphenylethanol	Phenylpropanoids or phenolic compounds
Fucose (1MEOX) (4TMS) MP	2-Deoxyglucose	Sugar
Glycylglycine, DL- (4TMS)	Glycylglycine	Amino acids
Glucopyranoside, 1-O-methyl-, alpha- (4TMS)	Methyl alpha-D-galactopyranoside	Sugar
Prolyl-glycine (2TMS)	Prolylglycine	Amino acids
Propane-1,2-diol, 3-amino (4TMS)	3-Aminopropane-1,2-diol	Others
Imidazole-4-acetic acid, 1-methyl- (1TMS)	Methylimidazoleacetic acid	Others
Histamine, 1-methyl- (2TMS)	1-Methylhistamine	Amines
1-Pyrroline-2-carboxylate (1TMS)	1-Pyrroline-2-carboxylate	Amino acids
Cyclohexane-1,2-dicarboxylic acid, trans- (2TMS)	1,2-Cyclohexanedicarboxylic acid	Other organic acids
1,3-Dihydroxyaceton (1MEOX) (2TMS) MP	Dihydroxyacetone	Sugar
1,4-Naphthoquinone, 2-methyl-	Menadione	Others
Cadaverine (4TMS)	Cadaverine	Amines
Butanoic acid, 2-amino- (2TMS)	D-2-Aminobutyrate	Amino acids
Galactose, 2-deoxy- (1MEOX) (4TMS) MP	2-Deoxy-D-galactose	Sugar
Glucosamine, N-acetyl- (1MEOX) (4TMS)	2-Acetamido-2-deoxy-D-glucose	Sugar
Ribose, 2-deoxy- (1MEOX) (3TMS) MP	2-Deoxy-D-ribose	Sugar
Butanoic acid, 2-hydroxy- (2TMS)	2-Hydroxybutyric acid	Lipids
Pyridine, 2-hydroxy- (1TMS)	2-Hydroxypyridine	Others
Fumaric acid, 2-methyl- (2TMS)	Mesaconate	Lipids
Butanoic acid, 2-oxo- (1MEOX) (1TMS) MP	2-Oxobutanoate	Lipids
Picolinic acid (1TMS)	Picolinic acid	Alkaloids
Succinic acid, 2,3-dimethyl- (2TMS)	2,3-Dimethylsuccinic acid	Lipids
Cinnamic acid, 2,5-dimethoxy-, trans- (1TMS)	2,5-Dimethoxycinnamic acid	Phenylpropanoids or phenolic compounds
Propanoic acid, 3-amino-3-(4-hydroxyphenyl)- (3TMS)	beta-Tyrosine	Amino acids
Shikimic acid, 3-dehydro- (1MEOX) (3TMS) MP	3-Dehydroshikimate	Phenylpropanoids or phenolic compounds
Anthranilic acid, 3-hydroxy- (3TMS)	3-Hydroxyanthranilate	Other organic acids
Benzoic acid, 3-hydroxy- (2TMS)	3-Hydroxybenzoic acid	Other organic acids
Pyridine, 3-hydroxy- (1TMS)	3-Hydroxypyridine	Others
Pentanedioic acid, 3-methyl- (2TMS)	3-Methylglutaric acid	Lipids

Mandelic acid, 3,4-dihydroxy- (4TMS)	3,4-Dihydroxymandelic acid	Phenylpropanoids or phenolic compounds
Acetic acid, 3,4-dihydroxyphenyl- (3TMS)	3,4-Dihydroxyphenylacetate	Phenylpropanoids or phenolic compounds
Cinnamic acid, 3,4,5-trimethoxy-, trans- (1TMS)	3,4,5-Trimethoxycinnamic acid	Phenylpropanoids or phenolic compounds
Cinnamic acid, 3,5-dimethoxy-, trans- (1TMS)	3,5-Dimethoxycinnamic acid	Phenylpropanoids or phenolic compounds
4-(Methylamino)benzoic acid (2TMS)	N-Methyl-4-aminobenzoate	Other organic acids
Acetic acid, 4-chlorophenyl- (1TMS)	4-Chlorophenylacetic acid	Others
4-Hydroxyphenyl-beta-glucopyranoside (5TMS)	Arbutin	Sugar
Cinnamic acid, 4-methoxy-, trans- (1TMS)	4-Methoxycinnamic acid	Phenylpropanoids or phenolic compounds
4-Pyridoxic acid (3TMS)	4-Pyridoxate	Others
Glyceric acid (3TMS)	L-Glycerate	Sugar
Nicotinic acid, 6-hydroxy- (2TMS)	6-Hydroxynicotinate	Other organic acids
Octanoic acid, 8-amino- (4TMS)	8-Aminooctanoic acid	Lipids
Acetamide, 2-phenyl- (2TMS)	2-Phenylacetamide	Others
Acetoacetic acid (1MEOX) (1TMS) BP	Acetoacetate	Lipids
Adenosine, 2*-deoxy- (3TMS)	Deoxyadenosine	Others
Ajmaline (2TMS)	Ajmaline	Alkaloids
Alanine (2TMS)	L-Alanine	Amino acids
Cysteinesulfinic acid (3TMS)	L-Cysteinesulfinic acid	Amino acids
Alanine, beta-, N-(2-hydroxyethyl)- (3TMS)	3-(2-hydroxyethylamino)propanoic acid	Amino acids
Adenosine, alpha- (4TMS) BP	Adenosine	Others
Tocopherol, alpha- (1TMS)	alpha-Tocopherol	Lipids
Anthranilic acid (2TMS)	Anthranilate	Other organic acids
Arabitol (5TMS)	D-Arabitol	Sugar
Ascorbic acid (4TMS)	L-Ascorbic acid	Others
Asparagine (3TMS)	Asparagine	Amino acids
Aspartic acid (3TMS)	L-Aspartate	Amino acids
Benzimidazole, 5,6-dimethyl- (1TMS)	Dimethylbenzimidazole	Others
Benzoic acid, 2-amino-3-methoxy- (2TMS)	3-Methoxyanthranilate	Other organic acids
Benzoic acid, 4-hydroxy- (2TMS)	4-Hydroxybenzoate	Phenylpropanoids or phenolic compounds
Benzylalcohol (1TMS)	Benzyl alcohol	Others
Benzylamine (2TMS)	Benzylamine	Amines
Alanine, beta- (3TMS)	beta-Alanine	Amino acids
beta-Alanyl-lysine (6TMS)	beta-Alanyl-L-lysine	Amino acids
Glycerol-2-phosphate (4TMS)	Glycerol 2-phosphate	Lipids
Sitosterol, beta- (1TMS)	Beta-Sitosterol	Lipids
Biotin, dethio- (3TMS)	Dethiobiotin	Lipids

Butanoic acid, 3-hydroxy- (2TMS)	D-beta-Hydroxybutyric acid	Lipids
Butanoic acid, 4-acetamido- (1TMS)	4-Acetamidobutanoate	Amino acids
Butanoic acid, 4-amino-3-hydroxy- (3TMS)	4-Amino-3-hydroxybutanoate	Amino acids
Butanoic acid, 4-hydroxy- (2TMS)	(R)-3-Hydroxybutanoate	Other organic acids
Butyro-1,4-lactam, 2-amino- (2TMS)	3-Aminopyrrolidin-2-one	Others
Caffeic acid, cis- (3TMS)	Caffeate	Phenylpropanoids or phenolic compounds
Caffeic acid, trans- (3TMS)	trans-Caffeate	Phenylpropanoids or phenolic compounds
Caffeine	Caffeine	Alkaloids
Homoserine lactone, N-heptanoyl-	N-Heptanoylhomoserine lactone	Lipids
Homoserine lactone, N-octanoyl-	N-Octanoyl-L-homoserine lactone	Lipids
Calystegine B2 (1MEOX) (4TMS)	N-Methylcalystegine B2	Alkaloids
Catechin (5TMS)	(+)-Catechin	Phenylpropanoids or phenolic compounds
Hydantoin, 1-methyl- (2TMS)	N-Methylhydantoin	Others
Cellobiose, D- (1MEOX) (8TMS) MP	Cellobiose	Sugar
Cholecalciferol, 25-hydroxy- (2TMS) MP	Calcidiol	Lipids
Cholestane, 3beta-hydroxy-, 5alpha- (1TMS)	5alpha-Cholestan-3alpha-ol	Lipids
Cinnamic acid, 4-hydroxy-, trans- (2TMS)	p-Coumaric acid	Phenylpropanoids or phenolic compounds
Hypoxanthine (2TMS)	Hypoxanthine	Alkaloids
Aconitic acid, cis- (3TMS)	cis-Aconitate	Other organic acids
Iminodiacetic acid (3TMS)	Iminodiacetic acid	Amino acids
Citric acid (4TMS)	Citrate	Other organic acids
Citric acid, 2-methyl- (4TMS) BP	2-Methylcitrate	Other organic acids
Citrulline (3TMS)	L-Citrulline	Amino acids
Coniferylalcohol, trans- (2TMS)	Coniferyl alcohol	Phenylpropanoids or phenolic compounds
Coniferylaldehyde, trans- (1MEOX) (1TMS) BP	Coniferyl aldehyde	Phenylpropanoids or phenolic compounds
Indole-3-acetonitrile (1TMS)	3-Indoleacetonitrile	Others
Coumarine	Coumarin	Phenylpropanoids or phenolic compounds
Cysteine (3TMS)	L-Cysteine	Amino acids
Cysteine, S-methyl-, DL- (2TMS)	S-Methyl-L-cysteine	Amino acids
Cysteinyl-glycine (3TMS)	L-Cysteinylglycine	Amino acids
Cystine (4TMS)	Cystine	Amino acids
Isobutanoic acid, 3-amino- (3TMS)	(R)-3-Amino-2-methylpropanoate	Amino acids
Cystineamine (4TMS)	Cystamine	Others
Cytidine (4TMS)	Cytidine	Others
Cytosine (2TMS)	Cytosine	Others
Cytosine, 5-methyl- (2TMS)	5-Methylcytosine	Others
Allose (1MEOX) (5TMS) MP	D-Allose	Sugar

Dehydroascorbic acid dimer (2MEOX) MP	Dehydroascorbic acid	Others
Dihydrozeatin riboside (4TMS)	Dihydrozeatin riboside	Others
Epigallocatechin (6TMS)	(-)-Epigallocatechin	Phenylpropanoids or phenolic compounds
Ergosterol (1TMS)	Ergosterol	Lipids
Eriodictyol (5TMS)	Eriodictyol	Phenylpropanoids or phenolic compounds
Erythronic acid (4TMS)	D-Erythronate	Sugar
Erythronic acid-1,4-lactone (2TMS)	Erythrono-1,4-lactone	Others
Lactic acid, DL- (2TMS)	L-Lactic acid	Other organic acids
Erythrose-4-phosphate (1MEOX) (4TMS) MP	D-Erythrose 4-phosphate	Sugar
Farnesol (1TMS) BP2	2-trans,6-trans-Farnesol	Others
Farnesol (1TMS) MP	trans-Farnesol	Lipids
Levodopa (4TMS)	L-Dopa	Amino acids
Ferulic acid, trans- (2TMS)	Ferulate	Phenylpropanoids or phenolic compounds
Flavone, 3-hydroxy- (1TMS)	3-Hydroxyflavone	Phenylpropanoids or phenolic compounds
Fructose (1MEOX) (5TMS) BP	D-Fructose	Sugar
Fructose-1,6-diphosphate (1MEOX) (7TMS) BP	D-Fructose 1,6-bisphosphate	Sugar
Fucosterol (1TMS)	Fucosterol	Lipids
Fumaric acid (2TMS)	Fumarate	Other organic acids
Lysine, N-epsilon-acetyl- (3TMS)	N6-Acetyl-L-lysine	Amino acids
Galactaric acid (6TMS)	D-Galactarate	Sugar
Galactinol (9TMS)	Galactinol	Sugar
Galactonic acid (6TMS)	D-Galactonate	Other organic acids
Galactonic acid-1,4-lactone (4TMS)	D-Galactono-1,4-lactone	Others
Maleimide, N-ethyl-	N-Ethylmaleimide	Amides
Galactose, 3,6-anhydro- (1MEOX) (3TMS)	3,6-Anhydrogalactose	Sugar
Malic acid, 2-methyl- (3TMS)	D-Citramalate	Lipids
Valero-1,5-lactam (1TMS)	N-Methyl-2-pyrrolidinone	Others
Gentiobiose (1MEOX) (8TMS) MP	Gentiobiose	Sugar
Gibberellin A3 (3TMS)	Gibberellin A3	Lipids
Gibberellin A4 (2TMS)	Gibberellin A4	Lipids
Glucaric acid-1,4-lactone (4TMS)	D-Glucaro-1,4-lactone	Others
Glucoheptonic acid (7TMS)	Glucoheptonic acid	Sugar
Glucoheptose (1MEOX) (6TMS) MP	d-Glycero-d-galacto-heptose	Sugar
Gluconic acid (6TMS)	D-Gluconic acid	Sugar
Gluconic acid-1,5-lactone (4TMS)	D-Glucono-1,5-lactone	Sugar
Gluconic acid, 2-amino-2-deoxy- (7TMS)	2-Amino-2-deoxy-D-gluconate	Sugar
Gluconic acid, 2-oxo- (1MEOX) (5TMS) BP	2-Dehydro-D-gluconate	Sugar
Gluconic acid, 2-oxo- (1MEOX) (5TMS) MP	2-Keto-D-gluconic acid	Sugar
Glucose-6-phosphate (1MEOX) (6TMS) BP	D-Glucose 6-phosphate	Sugar

Glucose-6-phosphate (1MEOX) (6TMS) MP	Glucose 6-phosphate	Sugar
Glucose-6-phosphate, 2-amino-2-deoxy- (6TMS)	D-Glucosamine 6-phosphate	Sugar
Glucose, 1,6-anhydro, beta- (3TMS)	1,6-Anhydro-beta-l-idopyranose	Sugar
Glucuronic acid (1MEOX) (5TMS) BP	Glucuronic acid	Sugar
Glutamic acid (3TMS)	L-Glutamate	Amino acids
Glutamic acid, N-methyl- (3TMS)	N-Methyl-L-glutamate	Amino acids
Glutaric acid (2TMS)	Glutaric acid	Lipids
Glutaric acid, 2-hydroxy- (3TMS)	(R)-2-Hydroxyglutarate	Other organic acids
myo-Inositol-1-phosphate (7TMS)	myo-Inositol 1-phosphate	Others
Myricetin (6TMS)	Myricetin	Phenylpropanoids or phenolic compounds
Glutaric acid, 2-oxo- (1MEOX) (2TMS) MP	2-Oxoglutarate	Other organic acids
Glutaric acid, 3-oxo- (1MEOX) (3TMS)	3-Oxoglutaric acid	Other organic acids
Glyceric acid-2-phosphate (4TMS)	2-Phospho-D-glycerate	Sugar
Glycerol (3TMS)	Glycerol	Sugar
Nonacosanoic acid (1TMS)	Nonacosanoic acid	Lipids
Glycerol-3-phosphate (4TMS)	Glycerol-3-phosphate	Lipids
Glycine, 2-phenyl- (2TMS)	Phenylacetic acid	Amino acids
Glycolic acid (2TMS)	Glycolic acid	Other organic acids
Octadecenoic acid, 6-(Z)- (1TMS)	Petroselinic acid	Lipids
Octanoic acid, 3-hydroxy- (2TMS)	3-Hydroxyoctanoate	Lipids
Guanidine (3TMS)	Guanidine	Amides
Guanine (3TMS)	Guanine	Others
Guanosine, 2*-deoxy- (5TMS)	Deoxyguanosine	Others
Gulonic acid, 2-oxo-, DL- (1MEOX) (5TMS) MP	2-Keto-L-gluconate	Sugar
Hexadecan-1-ol, n- (1TMS)	1-Hexadecanol	Lipids
Hippuric acid, 2-hydroxy- (3TMS)	Salicyluric acid	Other organic acids
Histidinol (4TMS)	L-Histidinol	Amines
Homocystine (4TMS)	L-Homocystine	Amino acids
Homoglutamine (3TMS)	Homoglutamine	Amino acids
Homoserine lactone, N-tetradecanoyl- (1TMS)	N-(2-oxooxolan-3-yl)tetradecanamide	Amino acids
Homoserine, O-succinyl- (3TMS)	O-Succinyl-L-homoserine	Amino acids
Hydantoin, 5-propionate- (3TMS)	Hydantoin-5-propionate	Others
Hydrocaffeic acid (3TMS)	3,4-Dihydroxyphenylpropanoate	Phenylpropanoids or phenolic compounds
Hydrocinnamic acid (1TMS)	3-Phenylpropanoate	Phenylpropanoids or phenolic compounds
Hydroquinone (2TMS)	Hydroquinone	Phenylpropanoids or phenolic compounds
Indol-3-yl-glucoside, beta- (5TMS)	Indican	Sugar
Indole-3-acetic acid (2TMS)	Indole-3-acetate	Alkaloids
Phenylalanine (2TMS)	L-Phenylalanine	Amino acids

3-Indoleacetic acid, 5-methoxy- (2TMS)	5-Methoxyindoleacetate	Others
Indole-3-acetic acid, 5-hydroxy-, 1H- (2TMS)	5-Hydroxyindoleacetate	Others
Indole-3-pyruvic acid (1MEOX) (2TMS) MP	Indolepyruvate	Others
Inosine (4TMS)	Inosine	Others
Inosine-5*-monophosphate (5TMS)	Inosinic acid	Others
Isoflavone, 5,7-dihydroxy-4*-methoxy- (2TMS)	Biochanin A	Phenylpropanoids or phenolic compounds
Isoleucine (2TMS)	L-Isoleucine	Amino acids
Isomaltose (1MEOX) (8TMS) MP	Isomaltose	Sugar
Isovaleric acid, 2-oxo- (1MEOX) (1TMS) BP	3-Methyl-2-oxobutanoic acid	Other organic acids
Isovaleric acid, 2-oxo- (1MEOX) (1TMS) MP	2-Keto-3-methylbutyric acid	Other organic acids
Kestose, 1- (11TMS)	1-Kestose	Sugar
Kynurenic acid (2TMS)	Kynurenic acid	Other organic acids
Menthol, L-(-)- (1TMS)	L-Menthol	Lipids
Lactic acid, 3-(4-hydroxyphenyl)- (3TMS)	(±)-2-Hydroxy-3-(2-hydroxyphenyl)propanoic acid	Phenylpropanoids or phenolic compounds
Lanosterol (1TMS)	Lanosterol	Lipids
Leucine (2TMS)	L-Leucine	Amino acids
Leucrose (1MEOX) (8TMS) MP	Leucrose	Sugar
Leucyl-glycyl-glycine (4TMS)	Leucyl-glycyl-glycine	Amino acids
Levulinic acid, 5-amino- (1MEOX) (3TMS) MP	5-Aminolevulinate	Lipids
Limonene, ()-	Limonene	Lipids
Loganin (5TMS)	Loganin	Lipids
Lysine (4TMS)	L-Lysine	Amino acids
Lyxose (1MEOX) (4TMS) BP	D-Lyxose	Sugar
Propaonic acid, 2-amino-3-ureido- (2TMS)	L-2-Amino-3-ureidopropionic acid	Amino acids
Lyxose (1MEOX) (4TMS) MP	D-Arabinose	Sugar
Maleamic acid (2TMS) BP	Maleamic acid	Lipids
Malic acid (3TMS)	L-Malate	Other organic acids
Malic acid, 2-isopropyl- (3TMS)	2-Isopropylmalic acid	Lipids
Malic acid, 3-isopropyl-, threo- (3TMS)	3-Isopropylmalate	Lipids
Maltose (1MEOX) (8TMS) BP	Maltose	Sugar
Maltotriitol (12TMS)	1,4-b-D-Cellotriitol	Sugar
Maltotriose (1MEOX) (11TMS) MP	Maltotriose	Sugar
Mannitol (6TMS)	Mannitol	Sugar
Mannopyranoside, 1-O-methyl-, alpha- (4TMS)	Methyl beta-D-glucopyranoside	Sugar
Methionine (2TMS)	L-Methionine	Amino acids
Methionine, N-formyl- (2TMS)	N-Formylmethionine	Amino acids
Galactopyranoside, 1-O-methyl-, beta- (4TMS)	Methyl beta-D-galactopyranoside	Sugar
Aspartic acid, N-(3-indolylacetyl)- (4TMS)	(Indol-3-yl)acetyl-L-aspartate	Amino acids
Aspartic acid, N-acetyl- (3TMS)	N-Acetyl-L-aspartate	Amino acids
Homoserine lactone, N-dodecanoyl- (1TMS)	N-Dodecanoyl-DL-homoserine lactone	Amino acids

N,N*-Diacetylchitobiose (1MEOX) (6TMS) MP	Chitobiose	Sugar
Lysine, N6-biotinyl- (5TMS)	N6-D-Biotinyl-L-lysine	Amino acids
Neopterin, 7,8-dihydro- (6TMS)	7,8-Dihydroneopterin	Others
Nicotinic acid (1TMS)	Nicotinate	Other organic acids
Threonine, O-methyl-, L- (2TMS)	O-Methyl-L-threonine	Amino acids
Octacosane, n-	Octacosane	Others
Quinic acid, 3-caffeoyl-, cis- (6TMS)	Cis-5-Caffeoylquinic acid	Phenylpropanoids or phenolic compounds
Octanoic acid, n- (1TMS)	Octanoic acid	Lipids
Ornithine, N2-acetyl- (3TMS)	N-Acetylornithine	Amino acids
Orotic acid (3TMS)	Orotic acid	others
Oxalic acid (2TMS)	Oxalate	Other organic acids
Oxaloacetate (1MEOX) (2TMS) BP	Oxalacetic acid	Lipids
Oxamic acid (2TMS)	Oxamate	Other organic acids
Oxamide (2TMS)	Oxamide	Amide
Perillic acid (1TMS)	Perillic acid	Lipids
Phenethylamine (2TMS)	Phenethylamine	Amines
Phenol, 2-amino- (2TMS)	2-Aminophenol	Others
Phenylalanine, N-acetyl- (2TMS)	N-Acetyl-L-phenylalanine	Amino acids
Phenylpyruvic acid (2TMS)	Phenylpyruvate	Phenylpropanoids or phenolic compounds
Phosphoenolpyruvic acid (3TMS)	Phosphoenolpyruvate	Other organic acids
Phytol (1TMS) MP	Phytol	Lipids
2-Piperidinecarboxylic acid (2TMS)	Pipecolic acid	Alkaloids
Porphobilinogen (4TMS)	Porphobilinogen	Amines
Prephenic acid (1MEOX) (3TMS)	Prephenate	Other organic acids
Proline (2TMS)	L-Proline	Amino acids
Proline, 3-hydroxy-, trans- (3TMS)	trans-3-Hydroxy-L-proline	Amino acids
Proline, 4-hydroxy-, trans- (3TMS)	Procollagen trans-3-hydroxy-L-proline	Amino acids
Serine, cyclo- (2TMS)	Cycloserine	Others
Propanoic acid, 3-(2-hydroxyphenyl)- (2TMS)	3-(2-Hydroxyphenyl)propanoic acid	Phenylpropanoids or phenolic compounds
3-(2,4-Dihydroxyphenyl)-propanoic acid (3TMS)	3-(2,4-Dihydroxyphenyl)propanoic acid	Phenylpropanoids or phenolic compounds
Putrescine, N-acetyl- (2TMS)	N-Acetylputrescine	Amines
Pyridoxal (1MEOX) (2TMS) MP	Pyridoxal	Others
Pyridoxine (3TMS)	Pyridoxine	Others
Pyrrole-2-carboxylic acid (2TMS)	Pyrrole-2-carboxylate	Other organic acids
Quercetin (5TMS)	Quercetin	Phenylpropanoids or phenolic compounds
Quinic acid (5TMS)	Quinic acid	Others
Quinic acid, 1-caffeoyl-, trans- (6TMS)	1-Caffeoylquinic acid	Others

Quinic acid, 3-caffeoyl-, trans- (6TMS)	Chlorogenic acid	Phenylpropanoids or phenolic compounds
Sorbose (1MEOX) (5TMS) BP	L-Sorbose	Sugar
Quinic acid, 4-caffeoyl-, trans- (6TMS)	Cryptochlorogenic acid	Phenylpropanoids or phenolic compounds
Quinic acid, 5-caffeoyl-, trans- (6TMS)	Chlorogenate	Phenylpropanoids or phenolic compounds
Quinoline-2-carboxylic acid, 4,8-dihydroxy- (3TMS)	Xanthurenic acid	Others
Resveratrol, trans- (3TMS)	Resveratrol	Phenylpropanoids or phenolic compounds
Rhamnose (1MEOX) (4TMS) MP	L-Rhamnose	Sugar
Ribonic acid (5TMS)	D-Ribonate	Sugar
Ribose-5-phosphate (1 MEOX) (5TMS) BP	D-Ribose 5-phosphate	Sugar
Ribose, 2-deoxy- (1MEOX) (3TMS) BP	Deoxyribose	Sugar
Rosmarinic acid (5TMS)	Rosmarinic acid	Phenylpropanoids or phenolic compounds
Saccharopine (5TMS)	L-Saccharopine	Amino acids
Salicylic acid (2TMS)	Salicylic acid	Phenylpropanoids or phenolic compounds
Scopolamine (1TMS)	Scopolamine	Alkaloids
Sebacic acid (2TMS)	Sebacic acid	Lipids
Secologanin (TMS) MP	Secologanin	Lipids
Methionine, seleno- (1TMS)	L-Selenomethionine	Amino acids
Serine (3TMS)	L-Serine	Amino acids
Serine, N-acetyl- (2TMS)	N-Acetylserine	Amino acids
Serine, O-phospho- (4TMS)	O-Phospho-L-serine	Amino acids
Shikimic acid (4TMS)	Shikimate	Other organic acids
Shikimic acid-3-phosphate (5TMS)	Shikimate 3-phosphate	Other organic acids
Sinapic acid, trans- (2TMS)	Sinapate	Phenylpropanoids or phenolic compounds
Cinnamaldehyde, 3,5-dimethoxy-4-hydroxy-, trans- (1MEOX) (1TMS) BP	Sinapaldehyde	Phenylpropanoids or phenolic compounds
Sorbitol, 1,4:3,6-dianhydro- (2TMS)	Isosorbide	Others
Squalene, all-trans-	Squalene	Lipids
Stigmasterol (1TMS)	Stigmasterol	Lipids
Stilbene, 3,5-dihydroxy-, trans- (2TMS)	Pinosylvin	Phenylpropanoids or phenolic compounds
Succinic acid (2TMS)	Succinate	Other organic acids
Sucrose (8TMS)	Sucrose	Sugar
Tagatose (1MEOX) (5TMS) BP	D-Tagatose	Sugar
Talose (1MEOX) (5TMS) BP	D-Talose	Sugar
Tetracosane, n-	Lignocerane	Others
Tetratriacontane	Tetratriacontane	Others
Theobromine (1TMS)	Theobromine	Alkaloids

Theophylline (1TMS)	Theophylline	Alkaloids
Thiazole, 4-methyl-5-hydroxyethyl- (1TMS)	4-Methyl-5-(2'-hydroxyethyl)-thiazole	Alkaloids
Threitol (4TMS)	D-Threitol	Sugar
Threonic acid (4TMS)	Threonate	Sugar
Threonine, allo- (3TMS)	L-Allothreonine	Amino acids
Tocopherol, beta- (1TMS)	beta-Tocopherol	Lipids
Trehalose, alpha,beta- (8TMS)	Neotrehalose	Sugar
Tropic acid (2TMS)	Tropic acid	Other organic acids
Tryptamine, 1-methyl- (2TMS)	N-Methyltryptamine	Alkaloids
Tryptophan, 5-hydroxy- (4TMS)	5-Hydroxy-L-tryptophan	Others
Turanose (1MEOX) (8TMS) MP	Turanose	Sugar
Tyramine (3TMS)	Tyramine	Amines
Tyrosine (3TMS)	L-Tyrosine	Amino acids
Tyrosine, N-acetyl- (3TMS)	N-Acetyl-L-tyrosine	Amino acids
Tyrosine, 3-iodo- (3TMS)	3-Iodo-L-tyrosine	Amino acids
Tyrosine, 3-nitro- (3TMS)	3-Nitrotyrosine	Amino acids
Tyrosine, meta- (3TMS)	Meta-Tyrosine	Amino acids
Uracil (2TMS)	Uracil	others
Uracil, dihydro- (2TMS)	Dihydrouracil	Others
Urea (2TMS)	Urea	Amides
Uridine (3TMS)	Uridine	Others
Uridine, 2*-deoxy- (3TMS) MP	2'-Deoxyuridine	Others
Acrylic acid, 4-imidazole- (2TMS)	Urocanate	Other organic acids
Valeric acid, 5-amino- (3TMS)	5-Aminovaleric acid	Amino acids
Valeric acid, 3-hydroxy-3-methyl- (1TMS)	3-Hydroxy-3-methylpentanoic acid	Others
Valine (2TMS)	L-Valine	Amino acids
Vanillic acid (2TMS)	Vanillate	Phenylpropanoids or phenolic compounds
Xylose, D- (1MEOX) (4TMS)	alpha-D-Xylose	Sugar
Xylulose (1MEOX) (4TMS) MP	D-Xylulose	Sugar
Xylulose-5-phosphate (1MEOX) (5TMS) BP	D-Xylulose 5-phosphate	Sugar
Zeatin, 9-beta-ribofuranosyl-, trans- (4TMS)	trans-Zeatin riboside	Others

ARTIGO 2: MACHINE LEARNING-ENHANCED ECOMETABOLOMICS REVEALS CULTIVAR-SPECIFIC METABOLIC RESPONSES AND AGRONOMIC PERFORMANCE IN FIELD-GROWN *Coffea arabica* L.

Artigo redigido de acordo com as normas da revista *Information Processing in Agriculture*

Machine learning-enhanced ecometabolomics reveals cultivar-specific metabolic responses and agronomic performance in field-grown *Coffea arabica* L.

Jéfyne Campos Carréra[1], Leonor Guerra-Guimarães[2,3], Paulo Henrique Sales Guimarães[1], John Charles D'Auria[4], Carla Pinheiro[5,6], Luana de Jesus Sartori[1], Vânia Aparecida Silva[7], Margarete Lordelo Volpato[7], Christiano de Sousa Machado de Matos[7], Gladyston Rodrigues Carvalho[7], Fabio Akira Mori[1]

[1]UFLA - Universidade Federal de Lavras, 37203-202 Lavras, Minas Gerais, Brasil.
[2]CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal.
[3]LEAF - Linking Landscape, Environment, Agriculture and Food Research Centre and TERRA Associated Laboratory, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017, Lisboa, Portugal.
[4]Research Group Metabolic Diversity, Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK Gatersleben), OT Gatersleben, Corrensstraße 3, 06466 Seeland, Germany
[5]UCIBIO Applied Molecular Biosciences Unit, Department of Life Sciences, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.
[6]Associate Laboratory i4HB Institute for Health and Bioeconomy, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.
[7]EPAMIG – Empresa de Pesquisa Agropecuária de Minas Gerais, 37200-000 Lavras, Minas Gerais, Brasil.

Correspondence author: Jéfyne Campos Carréra, Universidade Federal de Lavras, 37203-202 Lavras, Minas Gerais, Brasil. E-mail: jefynecarrera@gmail.com. https://orcid.org/0000-0001-6597-7152.

ORCID List

Leonor Guerra-Guimarães: https://orcid.org/0000-0002-9676-1036
Paulo Henrique Sales Guimarães: https://orcid.org/0000-0001-9158-1688
John Charles D'Auria: https://orcid.org/0000-0002-4865-3938
Luana de Jesus Sartori: https://orcid.org/0000-0003-0764-977X

Carla Pinheiro: https://orcid.org/0000-0002-7799-0183
Vânia Aparecida Silva: https://orcid.org/0000-0001-6640-3644
Margarete Lordelo Volpato: https://orcid.org/0000-0002-6905-6876
Christiano de Sousa Machado de Matos: https://orcid.org/0000-0001-7844-3352
Gladyston Rodrigues Carvalho: https://orcid.org/0000-0002-4466-674X
Fabio Akira Mori: https://orcid.org/0000-0002-7468-018X

ABSTRACT

Coffee production is significantly influenced by environmental factors. To characterize how different *Coffea arabica* cultivars adapt to field conditions, an ecometabolomic analysis was conducted. This study assessed 97 variables related to leaf biochemical and physiological characteristics and agronomic traits across five cultivars, along with climate and soil characteristics, from two experimental sites in the Cerrado Mineiro region of Brazil. To handle this complex dataset, models using six machine-learning algorithms were evaluated. According to our data, the Random Forest model showed the best performance in identifying the key variables for cultivar classification. The modeling indicated that the cultivars Sarchimor, Catiguá MG2, and Catuaí were more sensitive to environmental changes, whereas Paraíso 2 and Catiguá MG3 were less affected. Notably, Paraíso 2 consistently displayed stable agronomic traits, highlighting its superior adaptability and resilience under field conditions. Additionally, pathway analysis revealed specific metabolic pathways: the metabolic pathways of alanine, aspartate, and glutamate; β-alanine; purine; and cysteine and methionine, as well as, biosynthetic pathways of arginine and pantothenate and coenzyme A (CoA), that are potentially linked to agronomic characteristics, stress tolerance, and disease resistance in coffee plants under field conditions.

Keywords: Data modeling, Untargeted metabolomics, *Coffea arabica*, Cerrado Mineiro, Brazilian coffee.

1. Introduction

Plants exhibit dynamic metabolic capacity that enables them to adapt to a wide range of environmental conditions [1]. However, in recent years, cultivated plants have faced accelerated environmental shifts that may affect their development, yield, and susceptibility to pathogens and pests [2,3]. Coffee, one of the most economically significant crops globally, is

predominantly derived from the species *Coffea arabica* L. and *C. canephora* Pierre ex A.Froehner [4,5]. Brazil, the world's leading coffee producer, cultivates a wide variety of cultivars, each exhibiting distinctive traits, such as cup quality, drought tolerance, and disease resistance [5,6]. In recent years, the selection and development of coffee cultivars with enhanced adaptability to environmental conditions have emerged as a critical research priority [7]. This effort is essential not only to ensure the sustainable production of coffee but also to maintain a consistent supply of high-quality coffee that contributes to market value and enhances economic returns for producers.

The evaluation of coffee plant development under field conditions typically relies on the analysis of leaf morphology, photosynthetic activity, water content, and vegetative índices [8, 9, 10, 11]. Additionally, genomic, proteomic, and transcriptomic analyses of *Coffea* species have been conducted for various research purposes [12, 13, 14, 15, 16]. Nonetheless, studies of *Coffea* metabolism have largely focused on identifying interspecies differences [17, 18], or by assessing the factors related to beverage quality in *C. arabica* [19, 20]. Although physiological experiments and biochemical techniques are valuable for understanding coffee plant traits [21], they rarely consider how these findings are influenced by environmental factors, such as climate and soil characteristics.

Considering the impact of environmental changes on cultivar development, ecometabolomics has emerged as a new frontier in plant physiological research. This multidisciplinary approach enables the identification of biomarkers and physiological mechanisms associated with crop adaptation and resistance to various stressors [22]. Ecometabolomic studies, which also focus on plant-environment interactions, typically employ non-targeted metabolomic analysis alongside biochemical and morphological data from the species under investigation, as well as environmental characteristics [23].

Over the past decade, several ecometabolomics studies examining plant-environment interactions have incorporated machine-learning techniques [23]. Moreover, machine learning algorithms have been extensively applied in agriculture, aiding in the identification of ideotypes [24], the detection of field diseases [25], and the analysis of production efficiency [26]. Various machine learning methodologies have been applied to cultivar classification models based on morphological traits and spectroscopic parameters [27, 28], as well as metabolic signatures derived from multi-omics approaches [29].

The application of machine learning techniques to the ecometabolomic analysis of *C. arabica* cultivars offers several advantages. These techniques help identify underlying patterns within the data and improve both the accuracy and precision of cultivar classification [30]. The

primary objective of this study was to identify the key variables for classifying five *C. arabica* cultivars based on their environmental context. Using ecometabolomic models generated by machine learning algorithms, we aimed to understand how these variables interact to influence coffee yield and cup quality in the field.

2. Material and Methods

2.1 Field Experiments

The experimental sites were strategically chosen upon the establishment of Coffee Demonstration units, a collaborative endeavor between EPAMIG and local farming communities. In this project, a standardized set of cultivars was planted simultaneously in multiple municipalities within the Cerrado Mineiro region. The primary objective of this initiative was to evaluate the field performance of *Coffea arabica* plants, which have been grown since 2016 in the experimental germplasm fields of EPAMIG, located in the municipality of Patrocínio (18° 57’ S, 46°54’ W) and private farm in the city of Monte Carmelo (18°56’10” S, 47°22’36” W) (Fig. 1). In the experimental site in Patrocínio, the plant cultivation followed dry farming practices, while in Monte Carmelo, irrigation was employed. The irrigation system consists of surface-level Ram drippers spaced at intervals of 0.7 meters, each delivering a flow rate of 2.3 L h^{-1}. The plants were irrigated every 48 h over a three-hour period, whereby received approximately 9 L of water. The experimental sites in Patrocínio and Monte Carmelo are approximately 1,000 meters above sea level, with an elevation difference of up to 100 m between them. Monte Carmelo occupies the higher elevation.

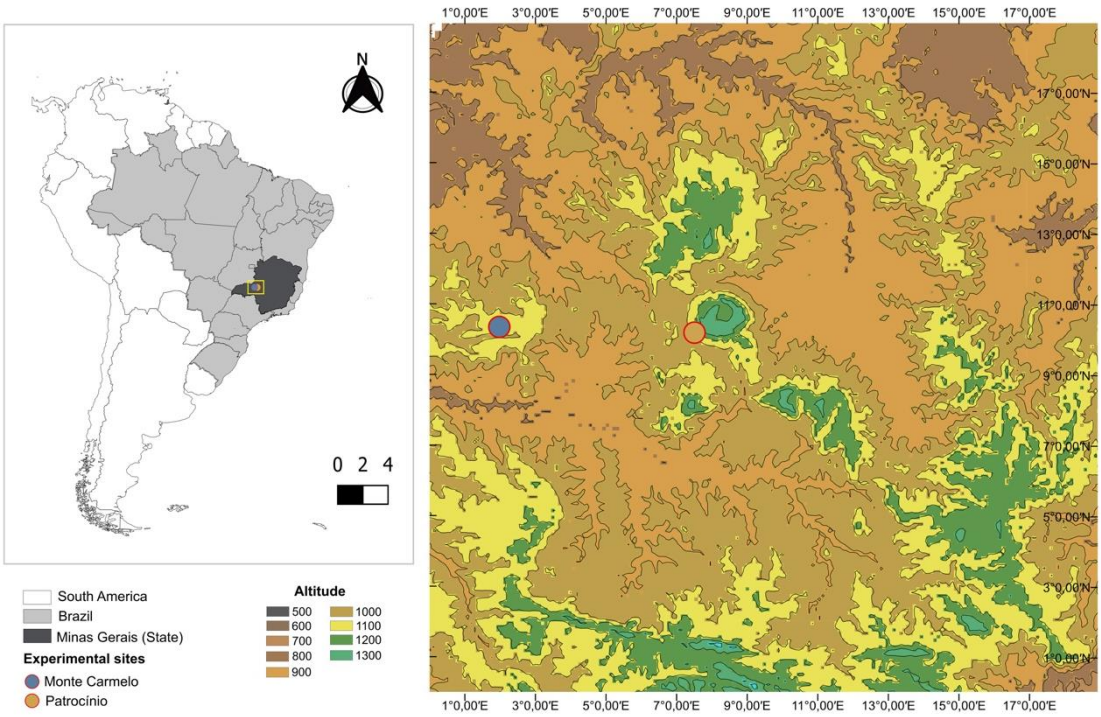


Fig. 1. Location of the experimental sites of Monte Carmelo and Patrocínio, situated in the Cerrado Mineiro region of Minas Gerais, Brazil.

2.2. Climatic Data Retrieval

Climate data were obtained from the agroclimatology database of the Prediction of Worldwide Energy Resources (POWER project, https://power.larc.nasa.gov) of the National Aeronautics and Space Administration [31]. Information was collected hourly from July 24 to August 20 (2022), using the coordinates of the two experimental areas. This evaluation period was selected to assess potential climatic variations between the two experimental sites, specifically considering the dry season in Cerrado Mineiro and the timing of sample collection (August 17-18, 2022). Eighteen variables were considered, corresponding to the following categories: solar flux and related temperatures, humidity and precipitation, and wind and pressure. Based on the average values recorded during the sampling period (7–9 am), the experimental locations exhibited a degree of similarity with the 18 variables analyzed (Supplementary Fig. S1 to S3). Average values for each sampling day were incorporated into the model.

2.3 Soil Sampling and Analysis

At each experimental site (Patrocinio and Monte Carmelo), two soil samples were collected at 20 cm depth. Soil samples were analyzed for the following elements and properties: nitrogen (N), potassium (K), phosphorus (P), calcium (Ca), magnesium (Mg), sodium (Na), aluminum (Al), zinc (Zn), iron (Fe), manganese (Mn), copper (Cu), boron (B), sulfur (S), phosphorus (P-rem), organic matter, pH (in water), potential acidity (H+Al), cation exchange capacity (CEC), effective cation exchange capacity (ECEC), base saturation (V%), aluminum saturation index (m), sum of bases (SB), and sand, clay, and silt content. The values of the 25 soil variables associated with each experimental site used in the ecometabolomic models are presented in Supplementary Table S1.

2.4 Plant Material

The five *C. arabica* cultivars selected for this study were chosen based on specific characteristics, including resistance to *Hemileia vastatrix* (coffee leaf rust, CLR) and *Meloidogyne exigua* (nematode), as well as their yield and cup quality (Table 1). The plant spacing is 3.8 m x 0.6 m in the same crop row (plot).

At the two experimental sites (Monte Carmelo and Patrocinio), a pair of fully expanded leaves from the middle third of five individual plants per cultivar was collected. The leaf pre-dawn water potential was assed at 5 a.m. and leaf collection for metabolomic assays took place between 7 and 9 a.m.. After collection the middle veins were excised, and the remaining leaf material was preserved using liquid nitrogen. The order of cultivars leaf collection remained consistent across both experimental sites, ensuring uniform sampling from the sun-exposed side of the plant.

Table 1. Classification of the five *Coffea arabica* cultivars studied: genetic origin, resistance to coffee leaf rust (CLR) and nematodes, cup quality and coffee production.

Cultivar	Genetic origin[1]	CLR resistance	Nematode resistance	Cup quality	Coffee production
Catuaí Vermelho IAC 144	Caturra amarelo IAC 476-11 x Mundo	Susceptible[2, 3]	Susceptible[2]	Regular[2]	High[2]

	Novo IAC 374-19				
Catiguá MG2	Catuaí amarelo IAC 86 x Híbrido de Timor UFV 440-10	Resistant[2,3]	Susceptible[2,4]	Distinct[2]	High[4]
Catiguá MG3	Catuaí amarelo IAC 86 x Híbrido de Timor UFV 440-10	Resistant[2,3]	Resistant[2,4]	Good[4]	High[2,4]
Paraíso 2	Catuaí amarelo IAC 30 x Híbrido de Timor UFV 445-46	Resistant[3]	Susceptible[2]	Distinct[2]	High[2]
Sarchimor MG 8840	Villa Sarchi CIFC 971/10 x Híbrido de Timor CIFC 832/2	Resistant[2]	Susceptible[2]	Regular[2]	High[2]

[1] IAC - Instituto Agronômico de Campinas (Brazil); MG - Minas Gerais (Brazil); CIFC - Centro de Investigação das Ferrugens do Cafeeiro (Portugal); UFV - Universidade Federal de Viçosa (Brazil). [2] [6]. [3] [32]. [4] [33].

2.5 Water potential measurement and leaf spectral indices determination

The pre-drawn water potential and leaf spectral vegetative indices were obtained from the same leaf samples. Measurements were conducted on fully expanded leaves from the middle third of the plants (3rd or 4th pair from the tip of an actively growing plagiotropic branch). The leaf water potential (measured in MPa) was determined following the methodology of Scholander [34] using a Scholander pressure chamber (1000 PMS Instruments Plant Moisture).

The leaf vegetative indices were obtained using a mini spectrometer (CID Bioscience Model: CI 610) [11]. The ten indices utilized were the Anthocyanin Reflectance Index (ARI1), cumulative values of NDVI (CNDVI), Greenness Index (GI), Gitelson and Merzlyak Index 1 (GM1), Normalized Difference Vegetation Index (NDVI), normalized pigment chlorophyll ratio index (NPCI), Photochemical Reflectance Index (PRI), Red-edge Normalized Difference Vegetation Index (ReNDVI), Structure Insensitive Pigment Index (SIPI), and scaled Photochemical Reflectance Index (sPRI), as specified in Supplementary Table S2.

2.6 Specific leaf area

The leaves used for assessing leaf water potential and spectral indices were stored in Kraft paper bags to facilitate drying and subsequent determination of specific leaf area (SLA; $cm^2\ g^{-1}$). The leaf area was quantified by scanning the leaves using a flatbed scanner, followed by image analysis using FIJI (ImageJ) software [35]. The collected material was oven-dried at 60 °C for 48 h and weighed on a precision scale to determine dry biomass. Specific leaf area was calculated as the ratio of leaf area to dry biomass.

2.7 Leaf metabolic analysis

The leaf samples (without the middle vein) were ground with liquid nitrogen, followed by lyophilization and storage in vacuum-sealed conical centrifuge tubes. The extraction procedure was performed in accordance with the methodology described by Salem and coworkers [36] with modifications for coffee leaf matrices [37]. This involved transferring 30 mg of lyophilized material into microtubes containing 1.5 mL of the extraction solvent mixture, Methyl-tert-butyl-ether: methanol (MTBE: MeOH, 3:1, v:v), along with internal standard L-Alanine-d4 (50 µL/100 mL). The microtubes were vortexed and incubated on an orbital shaker (40 rpm) for 45 min at 4 °C followed by sonication for 15 min in a water bath at 4 °C. The microtubes were then centrifuged (10,000 × g) for 10 min at 4 °C. From the soluble fractions, aliquots (0.5 mL) were transferred to new microtubes and dried in a speed vacuum for 45 min for metabolomic analyses. The remaining solid pellets, composed of proteins and starch, were washed with 500 µL methanol using a vortex mixer for 1 min. The pellets were then centrifuged at 10,000 × g for 5 min at 4°C. This process was repeated three times [38]. The protein and starch contents were determined in the pellets. Metabolic analysis, as well as protein and starch content measurements, was performed on five biological replicates for each coffee cultivar at both sites, Monte Carmelo and Patrocínio.

2.8 Metabolite profiling

Dried metabolic extracts in sealed, crimped vials were derivatized using 10 µL of a 20 mg mL -1 solution of methoxyamine hydrochloride in pyridine and 10 µL of *N*-trimethylsilyl-*N*-methyl trifluoroacetamide (MSTFA) [39] and injected using a Gerstel MPS2-XL (Gerstel, Muhlheim/Ruhr, Germany) autosampler. A LECO Pegasus BT time-of-flight mass spectrometer (LECO, St. Joseph, MI, USA) hyphenated with an Agilent 8890 gas chromatograph and helium as the carrier gas at 1.0 mL min-1 flow and linear velocity as flow control mode. The front inlet temperature was set to 230 °C, and the samples were run in splitless mode. The capillary column used was an Agilent DB-35MS (30 m × 0.25 mm × 0.25 µm). The temperature program started at 85°C in isothermal mode for 2 min. The isothermal step was followed by a 15°C/min ramp to 360°C. Fatty acid methyl esters (FAME MIX: C8-C30) were used to determine the retention time index (Kind et al. 2009). The ion-source temperature was 250 °C. The recorded mass range was m/z = 50-600 at 30 scans/s. A solvent delay of 3 min was used. The filament bias current was -70V, and the detector voltage ranged from approximately to 1900-2000 V depending on the detector age. The instrument was tuned and validated according to the EPA tune compliance.

Metabolic features were identified using LECO ChromaTOF software in conjunction with the GOLM Metabolome Database. Peak intensities were determined using the TargetSearch package [40] in R software (4.3.1 version) [41]. For normalization, the content of the internal standard L-Alanine-d4 was utilized. Univariate analysis was performed using MetaboAnalyst 6.0 [42] to identify the most significant metabolic features that distinguished the two experimental sites. Thirty-seven metabolic features were excluded from statistical analysis because they had a constant or single value across the samples. Additionally, features with more than 50% missing values were removed, and for the remaining features, missing values were imputed with 1/5 of the minimum positive value for each variable. After filtering, the remaining metabolic features were log-transformed (Base 10). As the objective was to select the most relevant metabolic features for database integration and modeling, a volcano plot was constructed with the following parameters: 1.75 of fold change and p-value < 0.05. Metabolic classes were organized in Microsoft Office Excel (Microsoft Corp., Redmond, WA, USA).

2.9 Leaf protein extraction

Protein extraction from the pellets (as previously mentioned 2.7) was conducted using a modified version [36]. The washed pellets were resuspended in 100 µL of protein extraction buffer (30 mM Tris-HCl pH 8.8, 7 M Urea, 2M thiourea, 15 mM DTT, and 0.50% CHAPS) for 10–15 min. The samples were sonicated in an ice bath (10 min) and incubated for 30 min on an

orbital shaker (40 rpm) at room temperature. The suspended proteins were centrifuged at 14,000 × g for 15 min and the protein supernatant was transferred to a new tube. The supernatant was used to determine protein concentration using a modified Bradford assay [43].

2.10 Starch accumulation analysis

Following protein extraction, pellets were used for starch quantification. The pellets were completely dried, and a 0.4 mL of a 0.7% α-amylase aqueous solution was added to each tube. To gelatinize the starch, the tubes were boiled for 3 min followed by autoclaving at 120 °C for 50 min. After cooling, 0.6 mL of an amyloglucosidase solution (1 M acetate buffer at pH 4.8, containing 1.67 mg/mL amyloglucosidase) was added to each tube. Subsequent incubation with agitation at 60 °C for 2 h facilitated the enzymatic hydrolysis. The samples were then centrifuged at 12,000 × g for 5 min and the glucose supernatant was transferred to new tubes. As a positive control, three replicates of 10 mg commercial starch were subjected to the same procedure. Starch content was quantified using the ®R-BIOPHARM Starch Kit (no. Cat. 10207748035, Boehringer Mannheim/R-Biopharm) modified for microtiter plates with visible light detection [44]. Glucose standards ranging from 0 to 1.2 g/L were used to establish a calibration curve, and absorbance readings were recorded at 490 nm. Results are expressed as micromoles of glucose equivalents per gram of fresh weight (FW).

2.11 Cultivar yield and Cup Quality

The data on coffee yield and cup quality for each cultivar at the two experimental sites were derived from the average values corresponding to four years of production (2019 - 2022). Approximately 20 coffee plants were selected to assess the yield and cup quality, and their average harvest volumes were measured. Cup quality analysis was conducted on a 10-liter sample of “cherry” coffee.

2.12 Machine learning models

2.12.1 Dataset pre-processing

All of the following methodologies were carried out using R software (version 4.3.1) [41]. The dataset included 36 metabolic features from *C. arabica* leaves (selected by volcano plot analysis as described in Section 2.8), along with additional variables related to agronomy (2), climate (18), soil (25), leaf characteristics (14), and two factors related to the experimental sites: water regime and altitude (Fig. 2). This created a model with 97 total variables, and the density (distribution) heatmaps of all leaf variables used are provided in Supplementary Fig. S4

and S5. Scaled and normalized variables were used, and the caret package [45] was used to divide the data into 60% for training and 40% for testing. Density heatmaps illustrating the data distributions of all variables incorporated in the ecometabolomic model are presented in Supplementary Fig. S4 and S5.

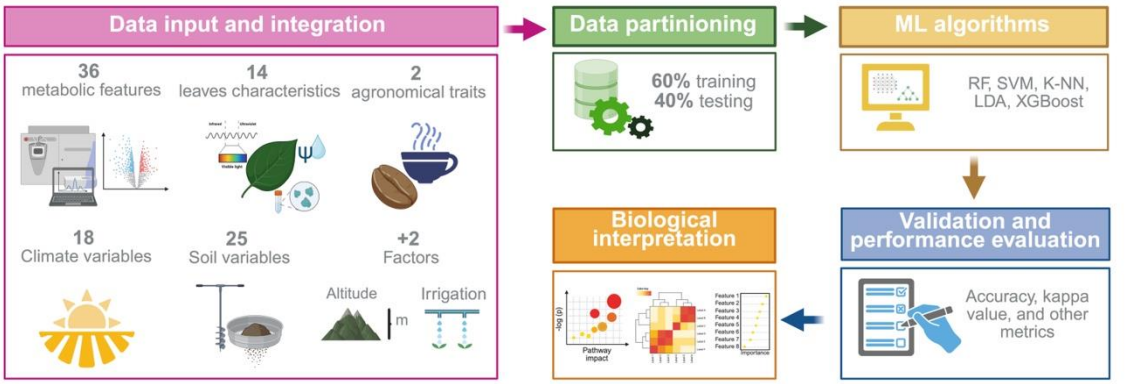


Fig. 2. The workflow of the ecometabolomic study was developed using different machine learning algorithms to obtain the most relevant variables for distinguishing *Coffea arabica* cultivars in Monte Carmelo and Patrocínio.

2.12.3 Algorithms

Machine learning methods (Table 2) were evaluated to determine their suitability for cultivar classification across the two experimental sites. The selection of methodologies was guided by their established strengths in classification modeling and proved to be effective in handling metabolomic data [46, 47, 48, 49, 50]. Random Forest, a powerful ensemble method, is known for its ability to reduce overfitting and its strong performance with large datasets [51, 52]. Extreme Gradient Boosting (XGBoost), a highly scalable and accurate algorithm, has demonstrated superior performance for various classification problems [53, 54]. The k-Nearest Neighbors (kNN) algorithm, which classifies data points based on their proximity to training samples, is well-suited for climate and soil data analysis [55]. Linear Discriminant Analysis (LDA), a dimensionality reduction technique, is effective for maximizing class separability and improving classification accuracy [56]. Finally, Support Vector Machines (SVM), a versatile algorithm, can handle high-dimensional data and nonlinear separations through kernel tricks [57]. The packages used for the analyses were as follows: “class” (k-NN), “randomForest” (Random Forest), “MASS” (Stepwise LDA), “e1071” (Support vector machine), and “XGBoost” (eXtreme Gradient Boosting) in R software (R version 4.3.1) [41].

Table 2. Machine learning methods were used to generate ecometabolomic models and their respective set parameters.

Machine Learning Method	Parameters	Reference
eXtreme Gradient Boosting (XGBoost)	Nrounds: 100 Lr: 0.3 Max_depth: 6 Min_child_weight: 1 Subsample: 1 Colsample_bytree: 1 Objective: "binary:logistic" Booster: "gbtree"	[58]
k-Nearest Neighbors (K-NN)	K: 5 Distance: 2 Algorithm: "cover_tree"	[59]
Random Forest (RF)	Mtry: $\sqrt{p}$ Nodesize: 1 Number of trees: 300	[60]
Stepwise linear discriminant analysis (Stepwise LDA)	Direction: "both" Trace: “FALSE”	[59]
Support vector machine (SVM-Lin)	Cost = 1 Gamma = 1/p Kernel: linear	[61]
Support vector machine (SVM-RBF)	Cost = 1 Gamma = 1/p Kernel: radial basis function	[61]

Algorithm (K-NN): Algorithm used to search for neighbors. Booster: Standard base model. Colsample_bytree: Fraction of resources used. Cost: Regularization parameter. Direction: Steps for adding and removing variables. Distance: Standard Euclidean distance. Gamma: Hyperparameter and the *p* in the formula represents the number of predictor variables. K: Number of neighbors considered. Lr: Learning rate. Max_depth: Maximum tree depth. Min_child_weight: Minimum sum of instances weights needed for a node splitting. Mtry: Number of variables considered at each node, where *p* is the total number of variables. Nodesize: Minimum size of end nodes. Nrounds: Number of interactions. Objective: Standard loss function for classification. Subsample: Fraction of samples used per tree. Trace: Step reports disabled

2.12.4 Performance Evaluation Metrics

To identify the algorithm with superior performance in classifying ecometabolomics models, Cohen's kappa index [62] and accuracy metrics [63] were employed. All analyses were conducted using the "caret" package [45] in R version 4.3.1 [41]. Because higher values indicate better performance, the model with the highest accuracy and Kappa index, as well as the narrowest confidence interval (Table 3), was selected for the subsequent evaluation of the most relevant variables in the classification of the ecometabolomics models. The kappa coefficient was selected as a metric for evaluating classifier performance because it is considered robust owing to its ability to normalize accuracy by the possibility of agreement by chance, which is an important consideration when assessing the performance of multiclass models [64].

Table 3. The Kappa index used in the model performance evaluation, as established by Cohen [62].

Kappa value (k)	Performance
< 0.00	Terrible
$0.00 \leq k \geq 0.20$	Bad
$0.21 \leq k \geq 0.40$	Reasonable
$0.41 \leq k \geq 0.60$	Good
$0.61 \leq k \geq 0.80$	Very Good
$0.81 \leq k \geq 1.00$	Great

An additional evaluation to determine the efficacy of the top-performing model in classifying each cultivar at each experimental site was carried out using the confusion matrices generated with the "confusionMatrix" function also from the "caret" package [45]. This analysis yielded metrics such as balanced accuracy, sensitivity, specificity, positive predictive value, and negative predictive value, providing insights into the ability of the model to correctly predict the occurrence of each cultivar relative to the reference data (Table 4).

Table 4. Metrics were used to assess cultivar classification across the two experimental sites of the optimal ecometabolomic model.

Evaluation metric	Formula
Sensitivity	$Event_{Tp} / (Event_{Tp} + Event_{Fn})$
Specificity	$Event_{Tn} / (Event_{Fp} + Event_{Tn})$
Balanced accuracy	(sensitivity + specificity) / 2
Prevalence*	$(Event_{Tp} + Event_{Fn}) / (Event_{Tp} + Event_{Fp} + Event_{Fn} + Event_{Tn})$
Positive predictive value	$(sensitivity*prevalence) / ((sensitivity*prevalence)+((1-specificity)*(1-prevalence)))$
Negative predictive value	$(specificity*(1-prevalence)) / (((1-sensitivity)*prevalence)+((specificity)*(1-prevalence)))$

Tp: True Positive, Fn: False Negative, Tn: True Negative, Fp: False Positive.

* Denotes values used solely for the calculation of other metrics.

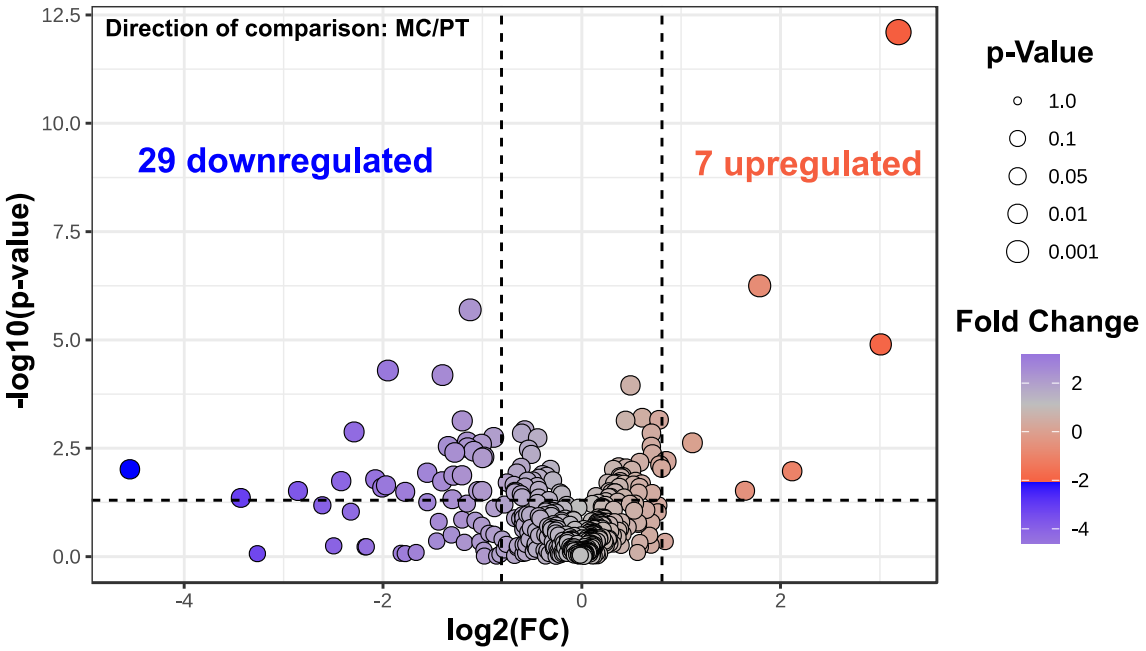
2.13 Additional data analysis

Comparative analysis of key ecometabolomic variables across cultivars and experimental sites was performed using R version 4.3.1 [41] with the following packages: “circlize” [65], “ComplexHeatmap” [66, 67], “dplyr” [68], “ggplot2” [69], “pacman” [70], and “vip” [71]. To complement the ecometabolomic model, pathway analysis was undertaken using MetaboAnalyst 6.0 [42] to identify the metabolic pathways that can significantly contribute to the observed phenotypic differences between the two experimental sites. The thirty-six most significant metabolic features that differentiated the two experimental sites were used to identify relevant metabolic pathways. This analysis incorporated 32 metabolites associated with 36 metabolic features annotated across various metabolic databases accessible through the software. The importance of each metabolic pathway was determined using Fisher's Exact Test as an enrichment method, with the *Arabidopsis thaliana* (thale cress) metabolome from the KEGG [72] database serving as the reference metabolome.

3. Results

3.1. Leaf Metabolite Profiling

Non-targeted metabolomic analysis of *C. arabica* cultivars grown in Monte Carmelo and Patrocínio revealed a metabolic profile with 320 validated features (Supplementary Table S3). However, only 36 of these metabolic features (MFs) were significantly differentially expressed between the two experimental sites, as identified by volcano plot analysis (Fig. 3). In Monte Carmelo, a predominantly downregulatory trend was observed among the 36 MFs



investigated, with 29 showing decreased expression levels and only seven exhibiting upregulation compared to Patrocínio (Fig. 3, Table 5). Notably, nearly half of these significant features belonged to amino acid and sugar categories (Table 5).

Fig. 3. Volcano plot illustrating the 36 most significant metabolic features differentiating the *C. arabica* cultivars in Monte Carmelo relative to Patrocínio. A volcano plot was generated using thresholds of fold change (FC > 1.75) and p-value (p > 0.05).

Table 5. Classification of the statistically significant metabolic features distinguishing *C. arabica* cultivars in Monte Carmelo vs. Patrocínio.

Metabolic Feature[1]	p-Value	RI interval[2]	Linked metabolite	Functional Class	Condition[3]
Alanine, beta- (3TMS)	< 0.01	394819 - 395160	beta-Alanine	Amino acids	downregulated
Asparagine (3TMS)	< 0.05	549139 - 550584	L-Asparagine (a)	Amino acids	downregulated
Asparagine (4TMS) BP2	< 0.05	479906 - 480101	L-Asparagine (b)	Amino acids	downregulated
Aspartic acid (3TMS)	< 0.05	458354 - 458549	L-Aspartate	Amino acids	downregulated
Cysteinesulfinic acid (3TMS)	< 0.05	555085 - 556703	L-Cysteinesulfinic acid/3-Sulfinoalanine	Amino acids	downregulated
Glutamic acid (3TMS)	< 0.05	508054 - 508442	L-Glutamate	Amino acids	downregulated
Glutamine, DL- (3TMS)	< 0.01	597675 - 598904	L-Glutamine (a)	Amino acids	downregulated
Glutamine [-H2O] (3TMS) MP	< 0.05	492844 - 493577	L-Glutamine (b)	Amino acids	downregulated
Homoserine (3TMS)	< 0.05	404620 - 406594	L-Homoserine	Amino acids	upregulated
Methionine (2TMS)	< 0.05	474201 - 474835	L-Methionine	Amino acids	downregulated
Phenylalanine (2TMS)	< 0.001	530493 - 531235	L-Phenylalanine	Amino acids	downregulated
Orotic acid, 4,5-dihydro- (4TMS)	< 0.01	644608 - 645312	L-Dihydroorotate	Amino acids	downregulated
Kestose, 1- (11TMS)	< 0.01	1032938 - 1033345	1-Kestose	Sugars	downregulated
Maltotriose (1MEOX) (11TMS) MP	< 0.01	1069904 - 1071158	Maltotriose	Sugars	downregulated

Melibiose (1MEOX) (8TMS) MP	< 0.05	912435 - 913451	Melibiose (a)	Sugars	downregulated
Melibiose (1MEOX) (8TMS) BP	< 0.05	921068 - 921480	Melibiose (b)	Sugars	downregulated
Sophorose (1MEOX) (8TMS) BP	< 0.05	899468 - 899912	Sophorose	Sugars	downregulated
Threonic acid (4TMS)	< 0.001	459719 - 459864	Threonate	Sugars	downregulated
Fumaric acid, 2-methyl- (2TMS)	< 0.001	403937 - 404718	Mesaconate	Lipids	upregulated
Maleamic acid (2TMS) BP	< 0.05	515650 - 516193	Maleamic acid	Lipids	downregulated
Pentanedioic acid, 3-methyl- (2TMS)	< 0.01	423383 - 423734	3-Methylglutaric acid	Lipids	upregulated
Sitosterol, beta- (1TMS)	< 0.01	1127307 - 1127615	Beta-Sitosterol	Lipids	downregulated
Caffeine	< 0.01	753318 - 755159	Caffeine	Alkaloids	downregulated
Hypoxanthine (2TMS)	< 0.01	657932 - 658494	Hypoxanthine	Alkaloids	downregulated
Tryptamine, 1-methyl- (2TMS)	< 0.05	785205 - 786414	N-Methyltryptamine	Alkaloids	upregulated
Lipoamide, alpha- (2TMS)	< 0.001	741719 - 742207	Lipoamide	Amides	downregulated
Urea (2TMS)	< 0.001	341322 - 341842	Urea	Amides	upregulated
Citric acid (4TMS)	< 0.01	592911 - 594380	Citrate	Other organic acids	downregulated
Tropic acid (2TMS)	< 0.05	519252 - 519828	Tropic acid	Other organic acids	downregulated

Cinnamic acid, 4-hydroxy-, trans- (2TMS)	< 0.05	685613 - 685840	p-Coumaric acid	Phenylpropanoids or phenolic compounds	downregulated
Piceatannol (4TMS) BP	< 0.001	879798 - 881427	Piceatannol	Phenylpropanoids or phenolic compounds	downregulated
4-Pyridoxic acid (3TMS)	< 0.05	673518 - 673970	4-Pyridoxate	Others	downregulated
Galactonic acid-1,4-lactone (4TMS)	< 0.01	617495 - 618546	D-Galactono-1,4-lactone	Others	downregulated
Guanine (3TMS)	< 0.001	778697 - 779803	Guanine	Others	upregulated
Inosine-5*-monophosphate (5TMS)	< 0.01	1037221 - 1038020	Inosinic acid	Others	upregulated
Uracil, dihydro- (2TMS)	< 0.01	494106 - 494388	5,6-Dihydrouracil	Others	downregulated

[1] Products of the derivatization process, annotated according to the GOLM-specific library; [2] retention index (RI) interval of the metabolic features found in the samples. (a) and (b) are different derivatized products of the same metabolite.

[3]According to the volcano plot analysis results (Fig. 3).

3.2 Comparative evaluation of predictive ability across Machine Learning methods

This section presents a comparative study of six machine-learning algorithms to assess their ability to classify cultivars based on ecometabolomic data. The algorithms investigated included Random Forest, Extreme Gradient Boosting, k-nearest Neighbors, Stepwise Linear Discriminant Analysis, and Support Vector Machines with both linear and radial basis function kernels. The evaluation of the best algorithm for the classification of cultivars in both experimental sites was based on 97 variables. These variables included 18 climate and 25 soil

characteristics, 2 agronomic traits, 36 metabolic features, 14 leaf characteristics, and 2 factors (altitude and water regime). Among the six algorithms evaluated, the Random Forest algorithm demonstrated the best performance in classifying cultivars in the Monte Carlo and Patrocínio coffee fields, as evidenced by its overall accuracy, kappa values, and confidence intervals (Fig. 4).

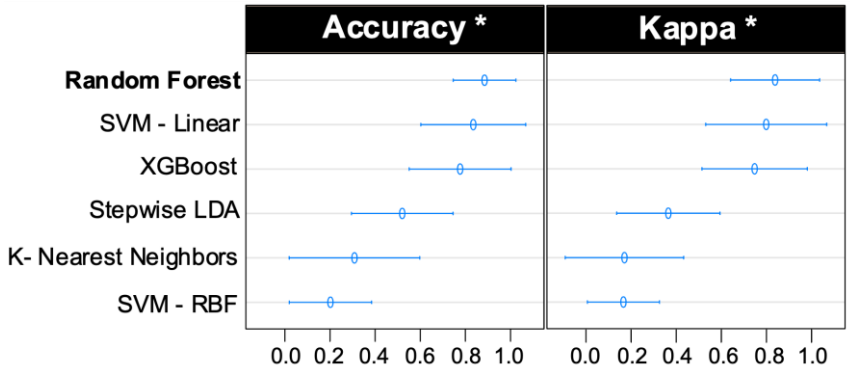


Fig. 4. Ecometabolomic Model Performance evaluation using mean accuracy and kappa values for six machine learning models in ecometabolomic data analysis. Error bars indicate confidence intervals. (*) Accuracy and Kappa values were calculated based on the classification of all ten cultivars.

3.3. Ecometabolomic analysis using the Random Forest model

Given the superior overall performance of the Random Forest model, a deeper analysis was conducted to investigate cultivar-specific patterns and potential distinctions in their behavior across experimental sites. This involved evaluating classification values from the confusion matrix using five additional performance metrics. Therefore, cultivar classification based on the confusion matrix revealed the highest evaluation metrics for Catiguá MG2, Catuaí, and Paraíso in Monte Carmelo and Sarchimor in Patrocínio (Fig. 5), indicating that the phenotypic distinctions among these cultivars were more pronounced in these specific sites, that enabled the algorithm to delineate the classification patterns for each cultivar more accurately.

In contrast, the classification performance for the Catiguá MG3 cultivar was less consistent across experimental sites, indicating greater variability in its response—particularly with respect to the Positive Predictive Value and Sensitivity metrics.

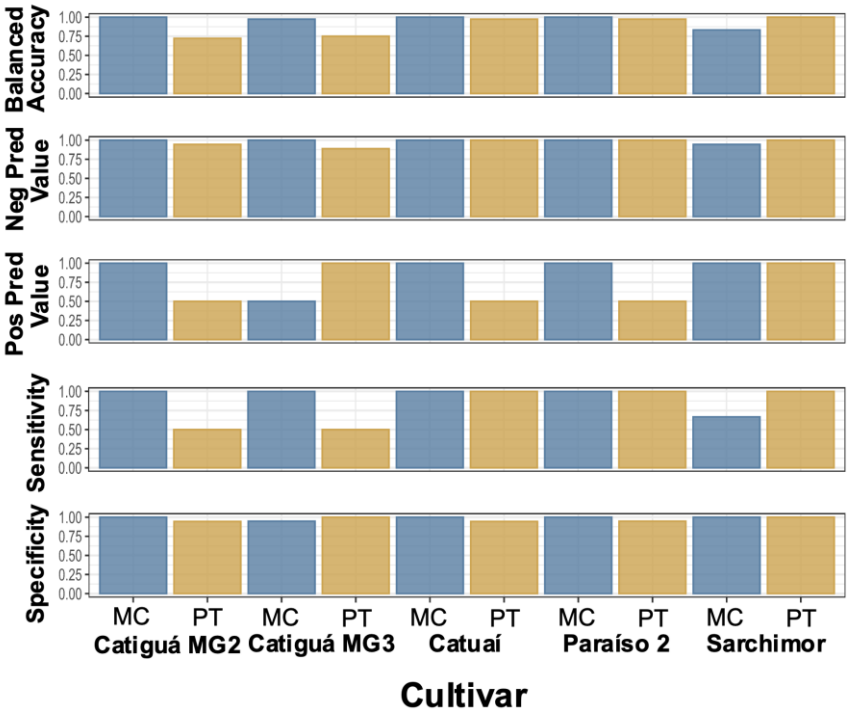


Fig. 5. Evaluation of cultivar classification in Monte Carmelo (MC) and Patrocinio (PT) using a Random Forest model based on the confusion matrix. Neg Pred Value: Negative predictive value. Pos Pred Value: Positive predictive value (Table 4 for details).

To identify the most influential variables for cultivar classification across both experimental sites, a variable importance plot was generated using the Random Forest ecometabolomics model. Applying a cut-off value of 3.0 in the variable importance plot (VIP > 3.0) of the RF model, 34 variables were identified as highly relevant for distinguishing the cultivars at the two experimental sites. These included 17 metabolic features, 13 climatic variables, two leaf characteristics (starch and specific leaf area), and two agronomic characteristics (coffee yield and cup quality) (Fig. 6).

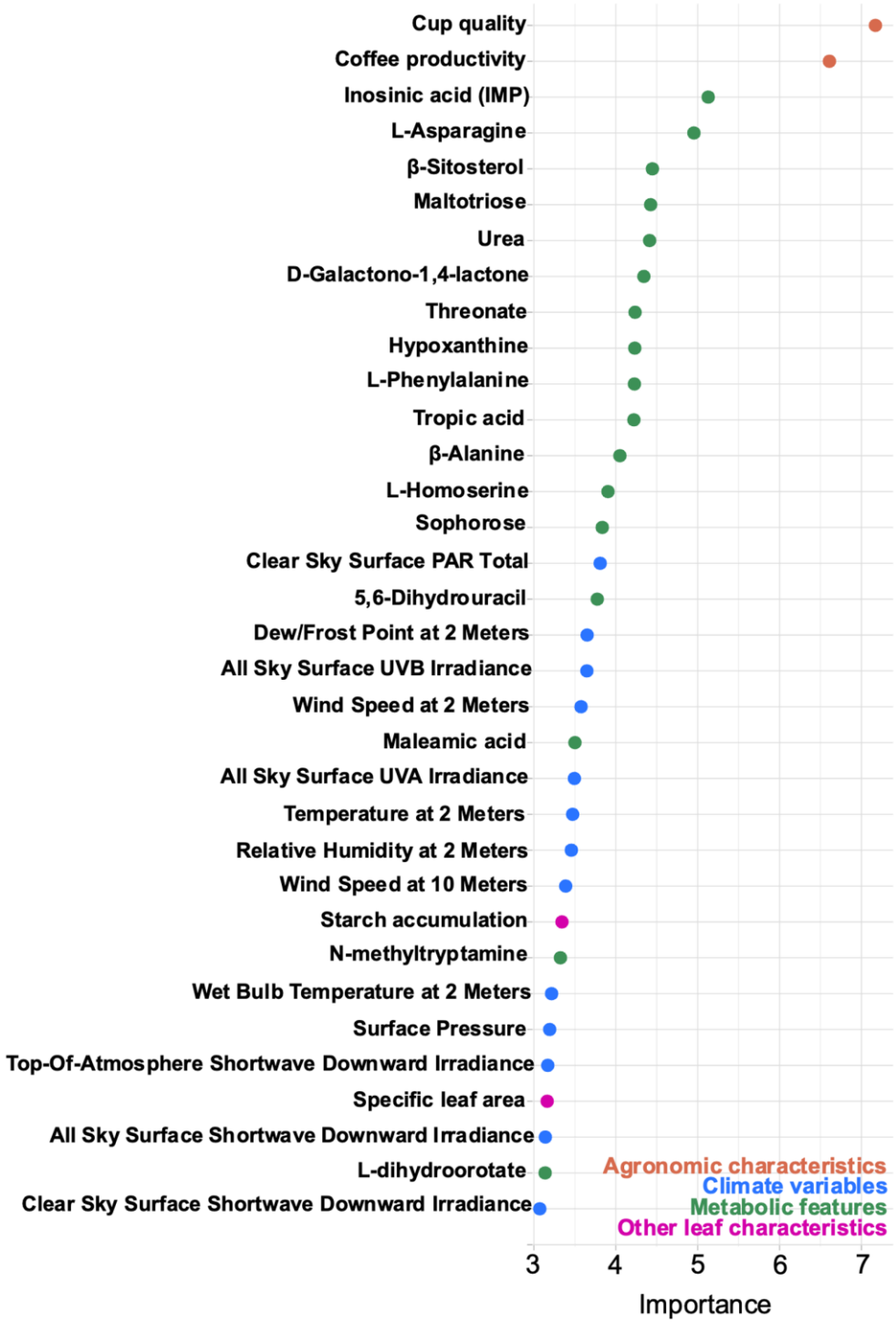


Fig. 6. Variable Importance Plot (VIP) of the Random Forest Model. The plot was generated using a cut-off value of VIP > 3.0 for variable importance.

The subsequent sections present a detailed analysis of the variations in key climate and cultivar characteristics between the two experimental sites, as identified by VIP analysis using the Random Forest ecometabolomic model.

3.4 Climatic variables

Among the relevant climate variables (Fig. 7, Supplementary Fig. S1 to S3), only three demonstrated a variation of at least 10% between Monte Carmelo and Patrocínio experimental sites. Wind-related parameters, including wind speed (WS) at 2 m and 10 m, exhibited the most substantial differences followed by relative humidity (RH) at 2 m (Fig. 7).

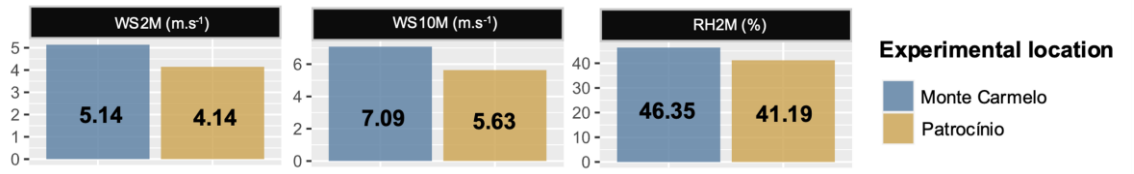


Fig. 7. The climate and soil variables of highest relevance, as identified by the ecometabolomic model generated through random forest analysis. WS2M: Wind Speed at 2 Meters; WS10M: Wind Speed at 10 Meters; RH2M: Relative Humidity at 2 Meters; 2022, July 24 to August 20.

3.5 Leaf and agronomic cultivar characteristics

The two agronomic traits, cup quality and coffee yield, were the most relevant variables for classifying cultivars within the ecometabolomic model. Among the 14 leaf traits (spectral indices, starch accumulation, soluble protein concentration, specific leaf area, and leaf water potential), only two (starch accumulation and specific leaf area) showed the highest variation across the experimental sites, thereby demonstrating greater relevance for the classification of the cultivars (Fig. 8).

Regarding agronomic characteristics, the Catiguá cultivars and Sarchimor showed the lowest values in Patrocínio, whereas in Paraíso 2, these characteristics remained consistently high at both sites (Fig. 8). Conversely, Catuaí had the lowest values among the observed parameters. The specific leaf area of the cultivars was higher in Monte Carmelo. The starch content in Sarchimor was the same in Monte Carmelo and Patrocínio, whereas the starch content of the other cultivars differed between the experimental locations.

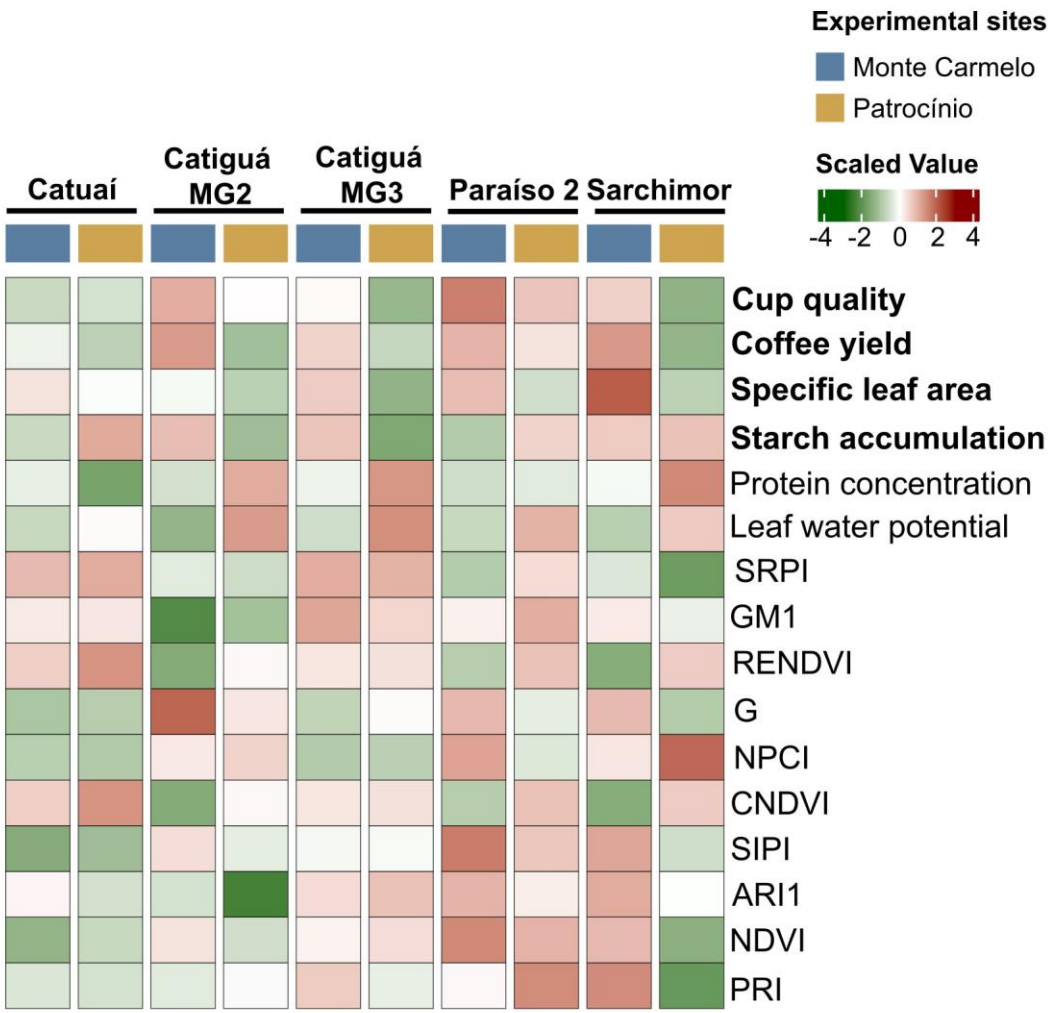


Fig. 8. Leaf and agronomic cultivar traits in Monte Carmelo and Patrocínio. Characteristics highlighted in bold were highly relevant according to the Random Forest ecometabolomic model. Spectral indices: sPRI- scaled Photochemical Reflectance Index; GM1 - Gitelson and Merzlyak Index 1; ReNDVI - Red-edge Normalized Difference Vegetation Index; GI - Greenness Index; NPCI - Normalized pigment chlorophyll ratio index; CNDVI - Cumulated values of NDVI; SIPI - Structure Insensitive Pigment Index; ARI1 - Anthocyanin Reflectance Index; NDVI - Normalized Difference Vegetation Index; PRI - Photochemical Reflectance Index.

3.6 Metabolic features and pathways

Within our dataset, the metabolic features varied among cultivars and across experimental sites (Fig. 9). Catiguá MG2 consistently exhibited contrasting levels of most metabolic features, with higher amounts observed in Patrocínio compared to Monte Carmelo,

whereas Catiguá MG3 showed only minor variation between the experimental sites. The cultivars Paraíso 2 and Sarchimor showed similarities in the expression of some metabolic features across the experimental sites. Conversely, the Catuaí cultivar exhibited a distinct metabolic profile relative to the other cultivars, marked by the accumulation of metabolites typically associated with the Patrocínio cultivar or, for some metabolites, by consistent accumulation across both experimental locations.

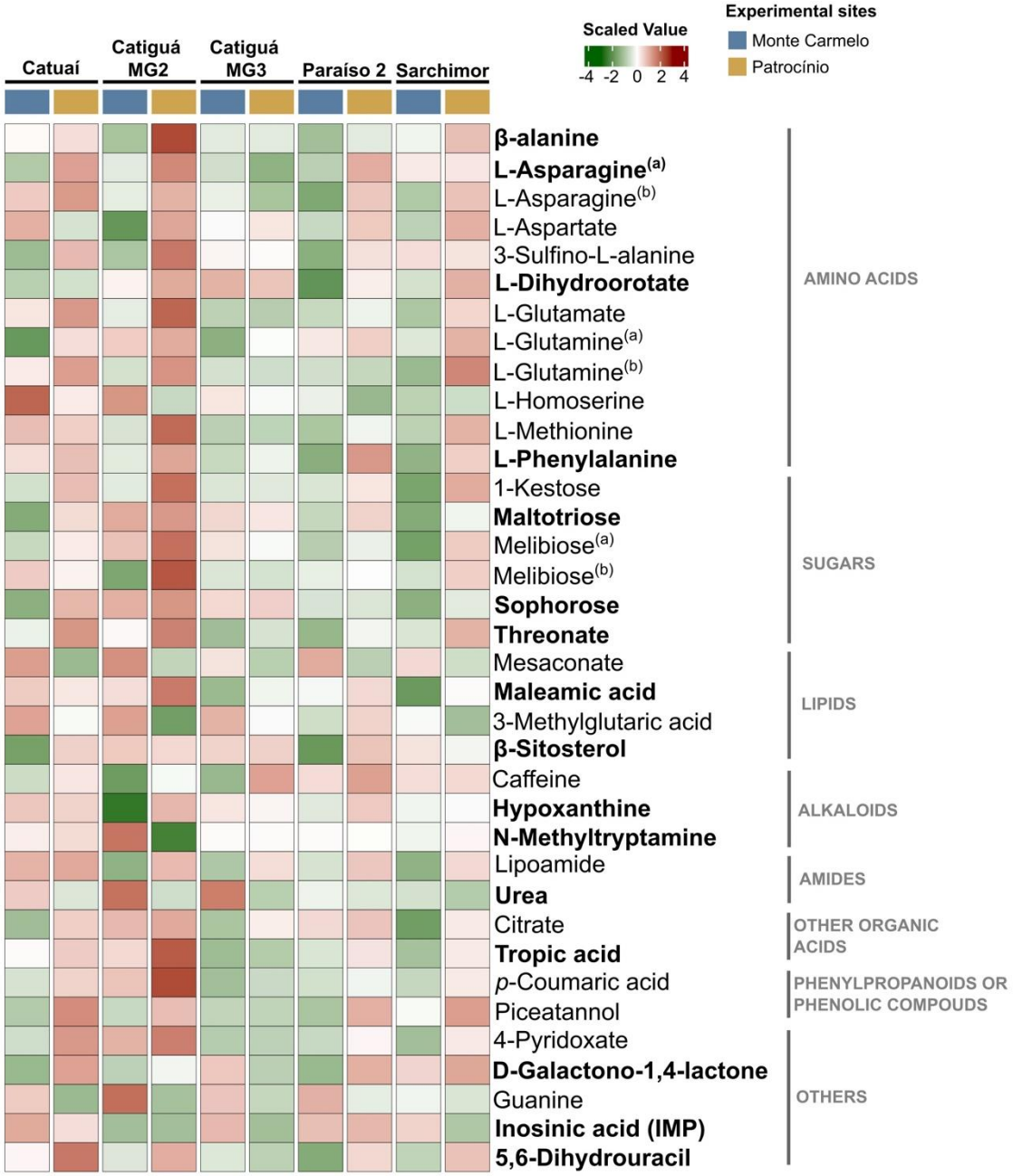


Fig. 9. Metabolic features (metabolites) identified as most significant for the cultivars grown across both experimental sites. Metabolites highlighted in bold were also highly relevant according to the Random Forest ecometabolomic model.

To identify the metabolic pathways that most significantly contributed to the differentiation between the two experimental sites, pathway enrichment analysis was conducted using the 36 identified key metabolic features (Fig. 9). The MFs participate in 31 distinct metabolic pathways, with six significant pathways demonstrating the greatest impact, predominantly associated with amino acid metabolism (Fig. 10A and B).

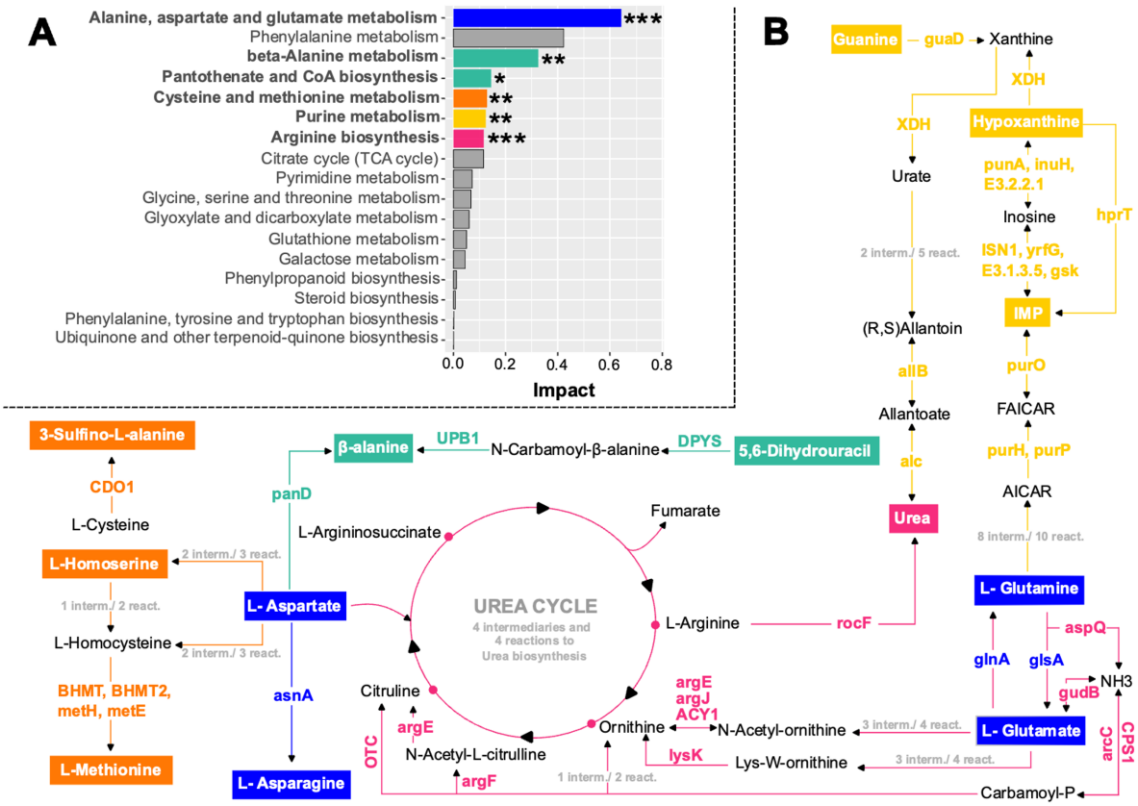


Fig. 10. Metabolic pathways of the 32 metabolites with the most significant differences between the two experimental sites (volcano plots). A: Analysis of metabolic pathways; Statistical significance is denoted as follows: (*) p-value < 0.05, (**) p-value < 0.01, and (***) p-value < 0.001. B: Integration of significant metabolic pathways with the greatest impact; metabolites detected in coffee leaves by non-targeted analysis are highlighted in colored boxes according to their respective metabolic pathways; Enzymes: ACY1 (aminoacylase); alc, ALLC (allantoicase); allB (allantoinase); arc (carbamate kinase); argE (acetylornithine deacetylase); argF (N-acetylornithine carbamoyltransferase); argJ (glutamate N-acetyltransferase); asnA (aspartate--ammonia ligase); aspQ (glutamin-(asparagin-)ase); BHMT (betaine-homocysteine S-methyltransferase); BHMT2 (homocysteine S-methyltransferase); CDO1 (cysteine dioxygenase); CPS1 (carbamoyl-phosphate synthase (ammonia)); DPYS

(dihydropyrimidinase); E3.1.3.5 (5'-nucleotidase); E3.2.2.1 (purine nucleosidase); glnA (glutamine synthetase); glsA (glutaminase); Gsk (inosine kinase); guaD (guanine deaminase); gudB (glutamate dehydrogenase); hprT (hypoxanthine phosphoribosyltransferase); inuH (inosine/uridine nucleosidase); ISN1 (IMP and pyridine-specific 5'-nucleotidase); lysK ([amino group carrier protein]-lysine/ornithine hydrolase); metE (5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase); metH (5-methyltetrahydrofolate--homocysteine methyltransferase); OTC (ornithine carbamoyltransferase); panD (aspartate 1-decarboxylase); punA (purine-nucleoside phosphorylase); purH (phosphoribosylaminoimidazolecarboxamide formyltransferase / IMP cyclohydrolase); puro (IMP cyclohydrolase); purP (5-formaminoimidazole-4-carboxamide-1-(beta)-D-ribofuranosyl 5'-monophosphate synthetase); rocF (arginase); UPB1 (beta-ureidopropionase); XDH (xanthine dehydrogenase/oxidase); yrfG (GMP/IMP 5'-nucleotidase).

4. Discussion

4.1. Machine learning-enhanced ecometabolomics for coffee cultivar discrimination

The present study used ecometabolomic analysis combined with machine learning algorithms to investigate the development of *C. arabica* cultivars at two distinct experimental sites. This approach highlighted key environmental factors distinguishing cultivars under field conditions and revealed metabolic differences between them. Among the machine learning methods evaluated, the Random Forest algorithm is a robust and widely recognized tool in metabolomic studies [73, 74, 75]- demonstrated superior performance in cultivar classification. The model identified distinct patterns exhibited by the cultivars under field conditions, as evidenced by robust performance metrics from the confusion matrix (Figure 5), underscoring its capability to analyze complex datasets.

4.2. Site-specific environmental effects on growth and development of C. arabica cultivars

The development of coffee plants under field conditions is modulated by a complex interplay between internal and external factors [7]. In the present study, Monte Carmelo and Patrocínio exhibited a high degree of similarity in key agroclimatological parameters, including solar radiation, precipitation, temperature, atmospheric pressure, and soil properties. The two sites were situated at comparable altitudes. However, discernible variations were observed in wind characteristics and relative humidity. Notably, irrigation in Monte Carmelo may have

contributed to an increase in specific leaf area during the leaf expansion phase, as this observation aligns with previous research demonstrating that irrigation can significantly enhance the leaf area index of coffee plants, irrespective of planting density, under field conditions [76].

Furthermore, drip irrigation in Monte Carmelo may have contributed to increased yield and enhanced cup quality across nearly all cultivars. However, the post-harvest period in this region, particularly in early August, is typically characterized by minimal rainfall (Supplementary Figure S2D). Precipitation in Patrocínio during the week preceding sample collection may have influenced the higher leaf water potential observed in plants at this site. A plausible explanation for this disparity lies in the different water acquisition strategies employed by irrigated and rainfed plants. Specifically, plants subjected to irrigation deficits tend to explore deeper soil horizons for water resources, whereas irrigated plants, such as those in Monte Carmelo, exhibit a greater concentration of roots in the superficial soil layers [77, 78]. This observation highlights the efficient water utilization strategies of rainfed plants in the Cerrado Mineiro region. Additionally, the elevated wind speeds recorded in Monte Carmelo during the sampling period may have influenced the disruption of boundary layer resistance, thereby affecting plant transpiration rates [79].

4.3 Cultivar-specific metabolic profiles in relation to agronomic traits

Although the environmental conditions in Monte Carmelo and Patrocínio exhibited a high degree of similarity, subtle variations were sufficient to induce significant alterations (> 10 %) in the metabolic profiles of the studied cultivars (Figure 3). The complex interactions among diverse environmental factors, coupled with the inherent dynamism of plant metabolism, can elicit nonlinear and often unpredictable physiological responses under combined environmental conditions [80]. Notably, the imposition of combined stresses can trigger adjustments in multiple primary metabolic pathways, resulting in the accumulation of sugars, amino acids, and specific metabolites associated with the tricarboxylic acid (TCA) cycle [2].

When assessing the performance of a coffee cultivar, it is essential to consider both yield and cup quality, ensuring that both parameters remain at high levels [81]. The results indicate that the five evaluated cultivars exhibited three distinct performance patterns. The Paraíso 2 cultivar exhibited consistently high average values for both yield and cup quality. Conversely, the Catiguá (MG2 and MG3) and Sarchimor cultivars displayed elevated values for these traits exclusively at the irrigated site, Monte Carmelo. Meanwhile, the Catuaí cultivar, included in this study as a disease-susceptible control, exhibited the lowest performance among all

cultivars, with the lowest average yield and cup quality values observed in Patrocínio and Monte Carmelo.

The period from August to September marks the transition from the resting phase to the preparation of buds for flowering in the phenological cycle of the coffee tree, a process triggered by the onset of rainfall or irrigation [82]. The present study revealed that, during this phenological stage, some cultivars exhibited heightened sensitivity to environmental variables, manifesting significant metabolic shifts across different locations within the Cerrado Mineiro region. These metabolic adjustments may in turn influence yield in the subsequent growing season. For instance, an analysis of metabolic differences among coffee cultivars that demonstrated high yields at least one of the experimental sites (Catiguá, Sarchimor, and Paraíso 2) revealed distinct responses. Catiguá MG2 and Sarchimor exhibited the most pronounced metabolic changes when transitioning from the irrigated site (Monte Carmelo) to the rainfed site (Patrocínio). In contrast, Paraíso 2 and Catiguá MG3 displayed the lowest metabolic variation between the two locations. Notably, the levels of specific metabolites and metabolic pathways appear to be directly associated with the average yield and cup quality, suggesting a potential link between metabolic plasticity and agronomic performance.

The superior field performance of the Paraíso 2 cultivar may be attributed to the stability of purine metabolic pathways, which play a crucial role in the biosynthesis of coffee alkaloids [83], and the efficient functioning of amino acid pathways involved in the synthesis of cysteine- and methionine-dependent compounds, as well as those related to arginine, alanine, aspartate, and glutamate metabolismo [84, 85, 86]. Additionally, the lower production of certain phenolic compounds (e.g., p-coumaric acid) and specific sugars (sophorose, melibiose, and threonate) may contribute to its agronomic advantages, as metabolites within these pathways exhibited consistent levels across both experimental sites, mirroring the stability observed in the agronomic traits of Paraíso 2. Specifically, the low sophorose levels detected in this cultivar indicate its role in anthocyanin binding [87, 88], as evidenced by the consistently high ARI1 spectral index values, an indicator of leaf anthocyanin content, at both experimental sites. This suggests a potential photoprotective strategy employed by Paraíso 2 to mitigate photoinhibition caused by excessive light exposure. As reviewed [89], anthocyanins play a crucial role in filtering excess light, thereby preventing oxidative stress, protecting the photosynthetic electron transport chain, and ultimately, reducing the severity and frequency of photoinhibition.

In general, the most pronounced differences between coffee cultivars and experimental sites were observed in amino acid metabolism, as reflected by the prominent representation of this group among differentially expressed metabolites (Table 5). Accumulation of amino acids

in plants is frequently associated with responses to adverse environmental conditions [90]. Coffee cultivars cultivated in the Cerrado Mineiro region exhibited significant activity in metabolic pathways related to amino acid biosynthesis and alkaloid production, indicating differential regulation of these pathways. Furthermore, it is noteworthy that increased expression of these six metabolic pathways (Figure 10) has been observed in studies involving plants subjected to various stress conditions [85, 91, 92, 93, 94].

Conversely, the lowest levels of the amino acid beta-alanine corresponded to the highest observed mean yield and cup quality across all cultivars and experimental sites. Beta-alanine, a constituent of two key metabolic pathways, beta-alanine metabolism, and pantothenate and CoA biosynthesis, demonstrates osmoprotective properties when converted to beta-alanine betaine [93]. Notably, in Catiguá MG3, the reduced levels of several metabolites - including amino acids - at both experimental sites might be associated with its resistance to nematodes (Table 1) and [6, 33]. In particular, beta-alanine, 5,6-dihydrouracil, and glutamate, all observed at lower concentrations in Catiguá MG3, are essential precursors for the biosynthesis of the thiol tripeptide homoglutathione (hGSH; γGlu-Cys-βAla) [93]. A correlation between hGSH depletion in roots and enhanced resistance to nematode infection in *Medicago truncatula* has been demonstrated [95].

While a definitive consensus remains elusive, several authors have argued that metabolic adjustments in crop plants can enhance field resilience [96, 97]. Based on this premise, Catiguá MG2 and Sarchimor could be considered metabolically responsive cultivars. However, it is crucial to emphasize that these metabolic shifts do not necessarily translate into improvements in the evaluated agronomic characteristics, suggesting that these cultivars altered their metabolic profiles primarily to mitigate adverse field conditions. Both Catiguá MG2 and Sarchimor exhibited an accumulation of amino acids and sugars, including melibiose, 1-kestose, and threonate, displaying a distinct metabolic profile compared with Paraíso 2. As previously noted, in complex field environments characterized by multiple interacting stressors, the overaccumulation of primary metabolites, such as sugars and amino acids, represents an early stage response to combined stress conditions [2].

Both threonate and melibiose are degradation products that are associated with distinct metabolic processes. Melibiose, a degradation product of raffinose, plays multiple roles in plants, including carbohydrate storage, responses to abiotic stress, and plant immunity [98]. Threonate, however, is a degradation product of ascorbate [99], a pathway that can also generate oxalate as a byproduct [100]. The consistently elevated levels of threonate observed in Catiguá MG2 and Sarchimor in Patrocínio suggest increased metabolic flux through this pathway at this

location. The formation of calcium oxalate crystals in leaf tissues is a mechanism for regulating excess calcium by compartmentalizing it within mesophyll cells [101]. Additionally, calcium oxalate crystals accumulated in mesophyll cells can subsequently serve as a carbon source for photosynthetic assimilation through stomatal closure, a phenomenon termed 'alarm photosynthesis', which confers an adaptive advantage under stress conditions [102].

In addition, the co-regulation of aspartate with organic acids, other amino acids such as beta-alanine, and sugars such as 1-kestose in Patrocínio for Catiguá MG2 and Sarchimor has been identified as a common acclimation response to various stress conditions [90]. Given that the accumulation of these compounds was observed in Patrocínio, the site with the lowest yield and cup quality for these two cultivars, it is likely that the metabolic pathways involved in the synthesis, consumption, or accumulation of these amino acids, and consequently the availability of pathway-dependent compounds, may be directly linked to the yield and cup quality in the Cerrado Mineiro region.

5. Conclusion

The Random Forest ecometabolomics model proposes key factors distinguishing cultivars across the two experimental sites. Despite the environmental similarities in soil, climate, and altitude between Monte Carmelo and Patrocínio, variations in leaf and agronomic characteristics were observed between the studied coffee cultivars. Sarchimor, Catiguá MG2, and Catuaí exhibited greater sensitivity to environmental fluctuations, which affected their yields and beverage quality.

Our findings highlight that Paraiso 2, recognized as a benchmark for cup quality and coffee yield, is the most stable among the studied cultivars in terms of its metabolic profile and agronomic traits. This suggests that Paraiso 2 could be a reliable option for growers seeking consistent yields and resilience under environmental fluctuations. This study further emphasizes the significance of distinct metabolic pathways associated with amino acid biosynthesis and metabolism, as well as other metabolite profiles, in coffee cultivars. It also underscores the necessity for continued investigation to elucidate the functional roles of these pathways in conferring stress tolerance and disease resistance in coffee plants.

STATEMENTS AND DECLARATIONS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

Financial support was provided by the Brazilian agencies: National Counsel of Technological and Scientific Development- CNPq, the Brazilian Consortium Coffee Research and Development, the National Institute of Coffee Science and Technology (INCT Café/CNPq), Coordination for the Improvement of Higher Education Personnel (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES) (Funding code: 001), and the Foundation for Research Support of the State of Minas Gerais (FAPEMIG). Financial support by the Access to Research Infrastructures, Horizon2020 Programme of the European Union (EPPN2020 Grant Agreement 731013) and Foundation for Science and Technology (FCT) from Portugal, Research Unit LEAF (UID/AGR/04129/2020) and UCIBIO (FCT UIDP/04378/2020; FCT UIDB/04378/2020).

References

[1] Salam U, Ullah S, Tang, Z., Elateeq, A. A., Khan, Y., Khan, J., et al. Plant Metabolomics: An Overview of the Role of Primary and Secondary Metabolites against Different Environmental Stress Factors. Life 2023; 13(3): 706. https://doi.org/10.3390/life13030706.

[2] Zandalinas SI, Balfagón D, Mittler R. Plant responses to climate change: Metabolic changes under combined abiotic stresses. Journal of Experimental Botany 2022; 73(11): 3339-3354. https://doi.org/10.1093/jxb/erac073.

[3] Singh BK, Egidi E, Guirado E, Leach JE, Liu H, Trivedi P. Climate change impacts on plant pathogens, food security and paths forward. Nature Reviews Microbiology 2023; 21(10): 640-656. https://doi.org/10.1038/s41579-023-00900-7.

[4] Davis AP, Chadburn H, Moat J, Hargreaves S, Lughadha EN. High extinction risk for wild coffee species and implications for coffee sector sustainability. Science Advances 2019; https://doi.org/10.1126/sciadv.aav3473.

[5] International Coffee Organization. Trade Statistic Tables, Production, Coffee Trade Status. [Online], 2024. Available: http://www.ico.org/ trade_statistics.asp. [20 June 2024].

[6] De Carvalho CHS, Bartelega L, Sera GH, Matiello JB, Santinato SRAF, Hotz AL. Catálogo de cultivares de café arábica, Documentos 16. Brasília, DF: Embrapa Café; 2022.

[7] Borgo L, Rabêlo FH, Marchiori PE, Guilherme LR, Guerra-Guimarães L, Resende M L. Impact of Drought, Heat, Excess Light, and Salinity on Coffee Production: Strategies for Mitigating Stress Through Plant Breeding and Nutrition. Agriculture 2025; 15(1): 9. https://doi.org/10.3390/agriculture15010009.

[8] Bote AD, Ayalew B, Ocho FL, Anten NP, Vos J. Analysis of coffee (*Coffea arabica* L.) performance in relation to radiation levels and rates of nitrogen supply I. Vegetative growth, production and distribution of biomass and radiation use efficiency. European Journal of Agronomy 2018; 92: 115-122. https://doi.org/10.1016/j.eja.2017.10.007.

[9] Tezara W, Loyaga DW, Reynel Chila VH, Herrera A. Photosynthetic Limitations and Growth Traits of Four Arabica Coffee (*Coffea arabica* L.) Genotypes under Water Deficit. Agronomy 2024; 14(8): 1713. https://doi.org/10.3390/agronomy14081713.

[10] Vilas-Boas T, Almeida HAD, Della Torre F, Modolo LV, Lovato MB, Lemos-Filho JP. Intraspecific variation in the thermal safety margin in *Coffea arabica* L. In response to leaf age, temperature, and water status. Scientia Horticulturae 2024; 337: 113455. https://doi.org/10.1016/j.scienta.2024.113455.

[11] Nunes PH, Pierangeli EV, Santos MO, Silveira HRO, de Matos CSM, Pereira AB et al. Predicting coffee water potential from spectral reflectance indices with neural networks, Smart Agricultural Technology 2023; 4: 100213. https://doi.org/10.1016/j.atech.2023.100213.

[12] Guerra-Guimarães L, Tenente R, Pinheiro C, Chaves I, Silva MC, Cardoso FMH et al. Proteomic analysis of apoplastic fluid of *Coffea arabica* leaves highlights novel biomarkers for resistance against *Hemileia vastatrix*. Front. Plant Sci. 2015; 6: 478. https://doi.org/10.3389/fpls.2015.00478.

[13] Marques I, Fernandes I, Paulo OS, Batista D, Lidon FC, Rodrigues AP et al. Transcriptomic Analyses Reveal That *Coffea arabica* and *Coffea canephora* Have More Complex Responses under Combined Heat and Drought than under Individual Stressors. Int. J. Mol. Sci. 2024; 25: 7995. https://doi.org/ 10.3390/ijms25147995.

[14] Possa KF, Silva JAG, Resende MLV, Tenente R, Pinheiro C, Chaves I et al. Primary Metabolism Is Distinctly Modulated by Plant Resistance Inducers in *Coffea arabica* Leaves Infected by *Hemileia vastatrix*. Front. Plant Sci. 2020; 11: 309. https://doi.org/10.3389/fpls.2020.00309.

[15] Salojärvi J, Rambani A, Yu Z, Guyot R, Strickler S, Lepelley M et al. The genome and population genomics of allopolyploid *Coffea arabica* reveal the diversification history of modern coffee cultivars. Nature Genetics 2024; 56(4): 721-731. https://doi.org/10.1038/s41588-024-01695-w.

[16] Zaman S, Shan Z. Literature Review of Proteomics Approach Associated with Coffee. Foods 2024; 13(11): 1670. https://doi.org/10.3390/foods13111670.

[17] Souard F, Delporte C, Stoffelen P, Thévenot EA, Noret N, Dauvergne B et al. Metabolomics fingerprint of coffee species determined by untargeted-profiling study using LC-HRMS, Food Chemistry 2017; 245: 603-612. https://doi.org/10.1016/j.foodchem. 2017.10.022.

[18] Montis A, Souard F, Delporte C, Stoffelen P, Stévigny C, Van Antwerpen P. Targeted and Untargeted Mass Spectrometry-Based Metabolomics for Chemical Profiling of Three Coffee Species. Molecules 2022; https://doi.org/ 10.3390/molecules27103152.

[19] Amalia F, Irifune T, Takegami T, Yusianto Sumirat U, Putri SP, Fukusaki E. Identification of potential quality markers in Indonesia's Arabica specialty coffee using GC/MS-based metabolomics approach. Metabolomics 2023; 19: 1-11. https://doi.org/10.1007/s11306-023-02051-5.

[20] Zhai H, Dong W, Fu X, Li G, Hu F. Integration of widely targeted metabolomics and the e-tongue reveals the chemical variation and taste quality of Yunnan Arabica coffee prepared using different primary processing methods, Food Chemistry 2024; https://doi.org/10.1016/j.fochx.2024.101286.

[21] Ge Y, Wang B, Shi X, Zhang Z, Qi M, Du H et al. Multi-Omics Analyses Unravel Genetic Relationship of Chinese Coffee Germplasm Resources. Forests 2024; 15(1): 163. https://doi.org/10.3390/f15010163.

[22] Sardans J, Gargallo-Garriga A, Urban O, Klem K, Walker TWN, Holub P et al. Ecometabolomics for a Better Understanding of Plant Responses and Acclimation to Abiotic Factors Linked to Global Change. Metabolites 2020; 10(6): 239. https://doi.org/10.3390/metabo10060239.

[23] Peters K, Worrich A, Weinhold A, Alka O, Balcke G, Birkemeyer C. Current Challenges in Plant Eco-Metabolomics. Int J Mol Sci. 2018; 19(5):1385. https://doi.org/10.3390/ijms19051385.

[24] Elbasi E, Zaki C, Topcu AE, Abdelbaki W, Zreikat AI, Cina E. et al. Crop Prediction Model Using Machine Learning Algorithms. Appl. Sci. 2023; 13: 9288. https://doi.org/10.3390/app13169288.

[25] Sajitha P, Andrushia AD, Anand N, Naser M. A review on machine learning and deep learning image-based plant disease classification for industrial farming systems. Journal of Industrial Information Integration 2024; 38: 100572. https://doi.org/10.1016/j.jii.2024.100572.

[26] Kheir AMS, Govind A, Nangia V, Devkota M, Elnashar A, Omar MED et al. Developing automated machine learning approach for fast and robust crop yield prediction using a fusion of remote sensing, soil and weather dataset. Environ. Res. Commun. 2024; 6: 041005. https://doi.org/10.1088/2515-7620/ad2d02.

[27] Zhao J, Bodner G, Rewald B. Phenotyping: Using Machine Learning for Improved Pairwise Genotype Classification Based on Root Traits. Frontiers in Plant Science 2016; 7: 221684. https://doi.org/10.3389/fpls.2016.01864.

[28] Fuentes S, Hernández-Montes E, Escalona J, Bota J, Gonzalez Viejo C, Poblete-Echeverría C. et al. Automated grapevine cultivar classification based on machine learning using leaf morpho- colorimetry, fractal dimension and near-infrared spectroscopy parameters. Computers and Electronics in Agriculture 2018; 151: 311-318. https://doi.org/10.1016/j.compag.2018.06.035.

[29] Li J, Gmitter Jr. FG, Zhang B, Wang Y. Uncovering Interactions between Plant Metabolism and Plant-Associated Bacteria in Huanglongbing-Affected Citrus Cultivars Using Multiomics Analysis and Machine Learning. Journal of Agricultural and Food Chemistry 2023; 71 (43): 16391-16401. https://doi.org/acs.jafc.3c04460.

[30] Zhong S, Zhang K, Bagheri M, Burken JG, Gu A, Li B. et al. Machine Learning: New Ideas and Tools in Environmental Science and Engineering. Environmental Science & Technology 2021; 55 (19): 12741-12754. https://doi.org/10.1021/acs.est.1c01339.

[31] NASA POWER. POWER project. Available:https://power.larc.nasa.gov. [30 November 2023].

[32] Oliveira ACB, Pereira AA, Caixeta ET, Resende MDV, Ribeiro MF. Cultivares de café resistentes à ferrugem: alternativa viável para a cafeicultura das Matas de Minas. Brasília: Embrapa Café, 2021.

[33] Fazuoli, L.C., de Carvalho, H.S., Carvalho, G.R. et al. CULTIVARES DE CAFÉ ARÁBICA (*Coffea arabica* L.). In: Carvalho, C.H.S. (ed). Cultivares de café. Brasília: Embrapa, 2007; 125-198.

[34] Scholander PF, Hammel HT, Hemmingsenm EA, Bradstreet ED. HYDROSTATIC PRESSURE AND OSMOTIC POTENTIAL IN LEAVES OF MANGROVES AND SOME OTHER PLANTS. Proc Natl Acad Sci. 2006; https://doi.org/10.1073/pnas.52.1.119.

[35] Schindelin JI, Arganda-Carreras E, Frise V, Kaynig M, Longair T, Pietzsch S et al. Fiji – an Open platform for biological image analysis. Nature Methods 2012; 9 (7): 676–82. https://doi.org/10.1038/nmeth.2019.Fiji.

[36] Salem MA, Perez de Souza L, Serag A, Fernie AR, Farag MA, Ezzat SM et al. Metabolomics in the Context of Plant Natural Products Research: From Sample Preparation to Metabolite Analysis. Metabolites 2020; https://doi.org/10.3390/metabo10010037.

[37] Carréra JC, Guerra-Guimarães L, D'Auria JC, Sartori LJ, Pinheiro C, Silva VA et al. NON TARGETED METABOLOMIC ANALYSIS OF FIELD-GROWN *Coffea arabica* CULTIVARS REVEALS DISTINCT LEAF METABOLIC SIGNATURES. Theoretical and Experimental Plant Physiology 2025; (In Press).

[38] Salem MA, Jüppner J, Bajdzienko K, Giavalisco P. Protocol: a fast, comprehensive and reproducible one-step extraction method for the rapid preparation of polar and semi-polar metabolites, lipids, proteins, starch and cell wall polymers from a single sample. Plant Methods. 2016; 10. https://doi.org/10.1186/s13007-016-0146-2.

[39] Erban A, Schauer N, Fernie AR, Kopka J. Nonsupervised Construction and Application of Mass Spectral and Retention Time Index Libraries From Time-of-Flight Gas Chromatography-Mass Spectrometry Metabolite Profiles. In: Weckwerth W. (ed.) Metabolomics - Methods in

Molecular Biology™, v. 358. Humana Press 2007; pp. 19-38. https://doi.org/10.1007/978-1-59745-244-1_2.

[40] Cuadros-Inostroza Á, Caldana C, Redestig H, Kusano M, Lisec J, Peña-Cortés H. et al. TargetSearch - a Bioconductor package for the efficient preprocessing of GC-MS metabolite profiling data. BMC Bioinformatics 2009; 10, 1-12. https://doi.org/10.1186/1471-2105-10-428.

[41] R Core Team, 2023. R: a language and environment for statistical computing. version 4.3.1. R Foundation for statistical computing, Vienna, Austria. Available: https://www.R-project.org/. [27 October 2023].

[42] Pang Z, Lu Y, Zhou G, Hui F, Xu L, Viau C. et al. MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation. Nucleic Acids Research 2024; 52 (W1): W398–W406 https://doi.org/10.1093/nar/gkae253.

[43] Ramagli LS. Quantifying protein in 2-D PAGE solubilization buffers. Methods Mol Biol. 1999; 112: 99-103. https://doi.org/10.1385/1-59259-584-7:99.

[44] Pinheiro C, Emiliani G, Marino G, Fortunato AS, Haworth M, De Carlo A. et al.. Metabolic Background, Not Photosynthetic Physiology, Determines Drought and Drought Recovery Responses in C3 and C2 Moricandias. International Journal of Molecular Sciences 2023; 24(4): 4094. https://doi.org/10.3390/ijms24044094.

[45] Kuhn M. Building Predictive Models in R Using the caret Package. Journal of Statistical Software 2008; 28(5): 1–26. https://doi.org/10.18637/jss.v028.i05.

[46] Guan X, Du Y, Ma R, Teng N, Ou S, Zhao H. et al. Construction of the XGBoost model for early lung cancer prediction based on metabolic indices. BMC Medical Informatics and Decision Making 2023; 23. https://doi.org/10.1186/s12911-023-02171-x.

[47] Mendez KM, Reinke SN, Broadhurst DI. A comparative evaluation of the generalised predictive ability of eight machine learning algorithms across ten clinical metabolomics data sets for binary classification. Metabolomics 2019; 15: 150. https://doi.org/10.1007/s11306-019-1612-4.

[48] Bifarin OO, Gaul DA, Sah S, Arnold RS, Ogan K, Master VA et al. Journal of Proteome Research 2021; 20 (7): 3629-364 https://doi.org/10.1021/acs.jproteome.1c00213.

[49] Antonelli J, Claggett BL, Henglin M, Kim A, Ovsak G, Kim N. et al. Statistical Workflow for Feature Selection in Human Metabolomics Data. Metabolites 2019; 9(7): 143. https://doi.org/10.3390/metabo9070143.

[50] Guo Y, Yu H, Chen D, Zhao YY. Machine learning distilled metabolite biomarkers for early stage renal injury. Metabolomics 2020; 16, 4. https://doi.org/10.1007/s11306-019-1624-0.

[51] Breiman L. Random Forests. Machine Learning 2001; 45: 5–32. https://doi.org/10.1023/A:1010933404324.

[52] Biau G, Scornet E. A random forest guided tour. TEST 2016; 25: 197–227. https://doi.org/10.1007/s11749-016-0481-7.

[53] Chen T, Guestrin C.. XGBoost – a scalable tree boosting system. KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 2016; 785 – 794. http://dx.doi.org/10.1145/2939672.2939785.

[54] Abdurrahman G, Sintawati M.. Implementation of xgboost for classification of parkinson's disease. J. Phys. Conf. Ser. 2020; 1538: 012024. https://doi.org/110.1088/1742-6596/1538/1/012024.

[55] Mucherino A, Papajorgji PJ, Pardalos PM. k-Nearest Neighbor Classification. In: Mucherino A, Papajorgji PJ, Pardalos PM. Mining in Agriculture. Springer Optimization and Its Applications, 34. New York, NY: Springer, 2009. https://doi.org/10.1007/978-0-387-88615-2_4.

[56] Xanthopoulos P, Pardalos PM, Trafalis TB. Linear Discriminant Analysis. In: Robust Data Mining. SpringerBriefs in Optimization. New York, NY: Springer, 2013. https://doi.org/10.1007/978-1-4419-9878-1_4.

[57] Noble WS. What is a support vector machine? Nature Biotechnology 2006; 24(12): 1565-1567. https://doi.org/10.1038/nbt1206-1565.

[58] Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H. et al. 2024. _xgboost: Extreme Gradient Boosting_. R package version 1.7.7.1, https://CRAN.R-project.org/package=xgboost.

[59] Venables WN, Ripley BD. Modern Applied Statistics with S, Fourth edition. New York: Springer, 2002. https://www.stats.ox.ac.uk/pub/MASS4/.

[60] Liaw A, Wiener M. Classification and Regression by randomForest. R News 2002; 2(3): 18-22. https://CRAN.R-project.org/doc/Rnews/.

[61] Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F, 2024. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.7-16. https://CRAN.R-project.org/package=e1071.

[62] Cohen J. A coefficient of agreement for nominal scales. Educ. Psychol. Meas 1960, 20 (1): 37-46.

[63] Dalianis H. Evaluation Metrics and Evaluation. In: Dalianis H. Clinical Text Mining. Springer, Cham. 2018; 45–53. https://doi.org/10.1007/978-3-319-78503-5_6.

[64] Nieto L, Correndo A. Classification performance metrics and indices. 2024. Available: <https://cran.r-project.org/web/packages/metrica/vignettes/available_metrics_classification.html>.[15 October 2024].

[65] Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize Implements and enhances circular visualization in R. Bioinformatics 2014; 30(19): 2811-2. https://doi.org/10.1093/bioinformatics/btu393.

[66] Gu, Z., Eils, R., Schlesner, M., 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32(18), 2847-2849. https://doi.org/10.1093/bioinformatics/btw313.

[67] Gu, Z., 2022. Complex heatmap visualization. IMeta 1(3), e43. https://doi.org/10.1002/imt2.43.

[68] Wickham H, François R, Henry L, Müller K, Vaughan D. dplyr: A Grammar of Data Manipulation. R package version 1.1.4, 2023. https://CRAN.R-project.org/package=dplyr.

[69] Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag, 2016.

[70] Rinker TW, Kurkiewicz D. pacman: Package Management for R. version 0.5.0. Buffalo, New York, 2017. http://github.com/trinker/pacman.

[71] Greenwell BM, Boehmke BC. The R Journal 2020; 12(1): 343-366.

[72] Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Research 2023; 51 (D1): D587-D592.

[73] Deklerck V, Finch K, Gasson P, Acker JV, Beeckman H, Espinoza E. Comparison of species classification models of mass spectrometry data: Kernel Discriminant Analysis vs Random Forest; A case study of Afrormosia (*Pericopsis elata* (Harms) Meeuwen). Rapid Communications in Mass Spectrometry 2017; 31(19): 1582-1588. https://doi.org/10.1002/rcm.7939.

[74] Schütz D, Achten E, Creydt M, Riedl J, Fischer M. Non-Targeted LC-MS Metabolomics Approach towards an Authentication of the Geographical Origin of Grain Maize (*Zea mays* L.) Samples. Foods 2021; 10(9): 2160. https://doi.org/10.3390/foods10092160.

[75] Cai W, Fang C, Liu L, Sun F, Xin G, Zheng J. Pseudotargeted metabolomics-based random forest model for tracking plant species from herbal products. Phytomedicine 2023; 118, 154927. https://doi.org/10.1016/j.phymed.2023.154927.

[76] Rezende FC, Dias Caldas AL, Scalco MS, Faria MA 2014. Leaf area index, plant density and water management of coffee. Coffee Science 2014; 9 (3): 374–384.

[77] Li Z, Li X, He F. Drip Irrigation Depth Alters Root Morphology and Architecture and Cold Resistance of Alfalfa. Agronomy 2022; 12, 2192. https://doi.org/10.3390/agronomy12092192.

[78] Brighenti S, Tagliavini M, Comiti F, Aguzzoni A, Giuliani N, Ben Abdelkader A. et al. Drip irrigation frequency leads to plasticity in root water uptake by apple trees. Agricultural Water Management 2024; 298: 108870. https://doi.org/10.1016/j.agwat.2024.108870.

[79] Shapira O, Hochberg U, Joseph A, McAdam S, Azoulay-Shemer T, Brodersen CR et al. Wind speed affects the rate and kinetics of stomatal conductance. The Plant Journal 2024; 120(4): 1552-1562. https://doi.org/10.1111/tpj.17066.

[80] Cappetta E, Andolfo G, Di Matteo A, Ercolano MR. Empowering crop resilience to environmental multiple stress through the modulation of key response components. J Plant Physiol. 2020; 246-247:153134. doi: 10.1016/j.jplph.2020.153134.

[81] Voltolini GB, Carvalho GR, Andrade VT, Ferreira AD, Raposo FV, Carvalho JPF et al. Agronomic Performance of Irrigated and Rainfed Arabica Coffee Cultivars in the Cerrado Mineiro Region. Agronomy 2025; 15: 222. https://doi.org/10.3390/agronomy15010222.

[82] Meireles EJL, Camargo MBP, Pezzopane JRM, Thomaziello RA, Fahl JI, Bardin L et al. Fenologia do Cafeeiro: Condições Agrometereológicas e Balanço Hídrico do Ano Agrícola 2004-2005, Documentos 5. Brasília (DF): Embrapa Informação Tecnológica, 2009.

[83] Ashihara H, Yokota T, Crozier A. Biosynthesis and Catabolism of Purine Alkaloids. Advances in Botanical Research 2013; 68: 111-138. https://doi.org/10.1016/B978-0-12-408061-4.00004-3.

[84] Winter G, Todd CD, Trovato M, Forlani G, Funck D. Physiological implications of arginine metabolism in plants. Frontiers in Plant Science 2015; 6: 150117. https://doi.org/10.3389/fpls.2015.00534.

[85] Hacham Y, Kaplan A, Cohen E, Gal M, Amir R. Sulfur metabolism under stress: Oxidized glutathione inhibits methionine biosynthesis by destabilizing the enzyme cystathionine γ-synthase. J. Integr. Plant Biol. 2024; 00: 1–14. https://doi.org/10.1111/jipb.13799.

[86] Liu Y, Zheng J, Ge L, Tang H, Hu J, Li X, et al. Integrated metabolomic and transcriptomic analyses reveal the roles of alanine, aspartate and glutamate metabolism and glutathione metabolism in response to salt stress in tomato. Scientia Horticulturae 2024; 328: 112911. https://doi.org/10.1016/j.scienta.2024.112911.

[87] Gonçalves AC, Bento C, Jesus F, Alves G, Silva LR. Sweet Cherry Phenolic Compounds: Identification, Characterization, and Health Benefits. Studies in Natural Products Chemistry 2018; 59: 31-78. https://doi.org/10.1016/B978-0-444-64179-3.00002-5.

[88] Pomilio AB, Mercader AG. Natural Acylated Anthocyanins and Other Related Flavonoids: Structure Elucidation of Ipomoea cairica Compounds and QSAR Studies Including Multidrug Resistance. Studies in Natural Products Chemistry 2018; 55: 293-322. https://doi.org/10.1016/B978-0-444-64068-0.00009-7.

[89] Gould KS. Nature's Swiss Army Knife: The Diverse Protective Roles of Anthocyanins in Leaves. Journal of Biomedicine and Biotechnology 2004; 5: 314–320.

[90] Han M, Zhang C, Suglo P, Sun S, Wang M, Su T. L-Aspartate: An Essential Metabolite for Plant Growth and Stress Acclimation. Molecules 2021; 26(7): 1887. https://doi.org/10.3390/molecules26071887.

[91] Yadav R, Saini R, Adhikary A, Kumar S. Unravelling cross priming induced heat stress, combinatorial heat and drought stress responsein contrasting chickpea varieties. Plant Physiology and Biochemistry 2022; 180: 91-105. https://doi.org/10.1016/j.plaphy.2022.03.030.

[92] Zhang Q, Tang S, Li J, Fan C, Xing L, Luo K. Integrative transcriptomic and metabolomic analyses provide insight into the long-term submergence response mechanisms of young *Salix variegata* stems. Planta 2021; 253: 88. https://doi.org/10.1007/s00425-021-03604-5.

[93] Parthasarathy A, Savka MA, Hudson AO. The Synthesis and Role of β-Alanine in Plants. Frontiers in Plant Science 2019; 10: 468525. https://doi.org/10.3389/fpls.2019.00921.

[94] Chen Q, Wang Y, Zhang Z, Liu X, Li C, Ma F, 2022. Arginine Increases Tolerance to Nitrogen Deficiency in *Malus hupehensis* via Alterations in Photosynthetic Capacity and Amino Acids Metabolism. Frontiers in Plant Science 2022; 12: 772086. https://doi.org/10.3389/fpls.2021.772086.

[95] Baldacci-Cresp F, Chang C, Maucourt M, Deborde C, Hopkins J, Lecomte P. et al. (Homo)glutathione deficiency impairs root-knot nematode development in Medicago truncatula. PLoS Pathog. 2012, 8(1): e1002471. https://doi.org/10.1371/journal.ppat.1002471.

[96] Rivero RM, Mittler R, Blumwald E, Zandalinas SI. Developing climate-resilient crops: Improving plant tolerance to stress combination. The Plant Journal 2022; 109(2): 373-389. https://doi.org/10.1111/tpj.15483.

[97] Benitez-Alfonso Y, Soanes BK, Zimba S, Sinanaj B, German L, Sharma V et al. Enhancing climate change resilience in agricultural crops. Current Biology 2023; 33(23): R1246-R1261. https://doi.org/10.1016/j.cub.2023.10.028.

[98] Sanyal R, Kumar S, Pattanayak A, Kar A, Bishi SK. Optimizing raffinose family oligosaccharides content in plants: A tightrope walk. Frontiers in Plant Science 2023; 14: 1134754. https://doi.org/10.3389/fpls.2023.1134754.

[99] Truffault V, Fry SC, Stevens RG, Gautier H. Ascorbate degradation in tomato leads to accumulation of oxalate, threonate and oxalyl threonate. The Plant Journal 2017; 89(5): 996-1008. https://doi.org/10.1111/tpj.13439.

[100] Melino VJ, Soole KL, Ford CM. Ascorbate metabolism and the developmental demand for tartaric and oxalic acids in ripening grape berries. BMC Plant Biol. 2009; 9: 145. https://doi.org/10.1186/1471-2229-9-145.

[101] Paiva EAS. Are calcium oxalate crystals a dynamic calcium store in plants? New Phytologist 2019; 223(4): 1707-1711. https://doi.org/10.1111/nph.15912.

[102] Tooulakou G, Giannopoulos A, Nikolopoulos D, Bresta P, Dotsika E, Orkoula M G. et al. Alarm Photosynthesis: Calcium Oxalate Crystals as an Internal CO2 Source in Plants. Plant Physiology 2016, 171, 2577–2585. https://doi.org/10.1104/pp.16.00111.

[103] Gitelson A, Merzlyak M, Chivkunova O. Optical Properties and Nondestructive Estimation of Anthocyanin Content in Plant Leaves. Photochemistry and Photobiology 2001; 71: 38-45.

[104] Panek E, Gozdowski D. Relationship between MODIS Derived NDVI and Yield of Cereals for Selected European Countries. Agronomy 2021; 11(2): 340. https://doi.org/10.3390/agronomy11020340.

[105] Zarco-Tejada PJ, Miller JR, Noland TL, Mohammed GH, Sampson PH. Scaling-up and Model Inversion Methods with Narrowband Optical Indices for Chlorophyll Content Estimation in Closed Forest Canopies with Hyperspectral Data. IEEE Trans. Geosci. Remote Sens. 2001; 39: 1491–1507.

[106] Gitelson AA, Merzlyak MN. Remote estimation of chlorophyll content in higher plant leaves. International Journal of Remote Sensing 1997; 18(12): 2691–2697. https://doi.org/10.1080/014311697217558.

[107] Serrano L, González-Flor C, Gorchs G. Assessment of grape yield and composition using the reflectance based Water Index in Mediterranean rainfed vineyards. Remote Sensing of Environment 2012; 118: 249-258. https://doi.org/10.1016/j.rse.2011.11.021.

[108] Merzlyak MN, Gitelson AA, Chivkunova OB, Rakitin VY. Non-destructive optical detection of leaf senescence and fruit ripening. Physiol. Plant. 1999; 106: 135–141.

[109] Gamon J, Peñuelas J, Field C. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. Remote Sensing of Environment 1992; 41(1): 35-44. https://doi.org/10.1016/0034-4257(92)90059-S.

[110] Van Beek J, Tits L, Somers B, Coppin P. Stem Water Potential Monitoring in Pear Orchards through WorldView-2 Multispectral Imagery. Remote Sensing 2013; 5(12): 6647-6666. https://doi.org/10.3390/rs5126647.

[111] Rahman AF, Cordova VD, Gamon JA, Schmid HP, Sims DA. Potential of MODIS ocean bands for estimating CO2 flux from terrestrial vegetation: A novel approach. Geophysical Research Letters 2004; 31: L10503.

Supplementary Material

Machine learning-enhanced ecometabolomics reveals cultivar-specific metabolic responses and agronomic performance in field-grown *Coffea arabica* L.

Jéfyne Campos Carréra[1*], Leonor Guerra-Guimarães[2,3], Paulo Henrique Sales Guimarães[1], John Charles D'Auria[4], Carla Pinheiro[5,6], Luana de Jesus Sartori[1], Vânia Aparecida Silva[7],

Margarete Lordelo Volpato[7], Christiano de Sousa Machado de Matos[7], Gladyston Rodrigues Carvalho[7], Fabio Akira Mori[1]

[1]UFLA - Universidade Federal de Lavras, 37203-202 Lavras, Minas Gerais, Brasil.

[2]CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal.

[3]LEAF - Linking Landscape, Environment, Agriculture and Food Research Centre and TERRA Associated Laboratory, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017, Lisboa, Portugal.

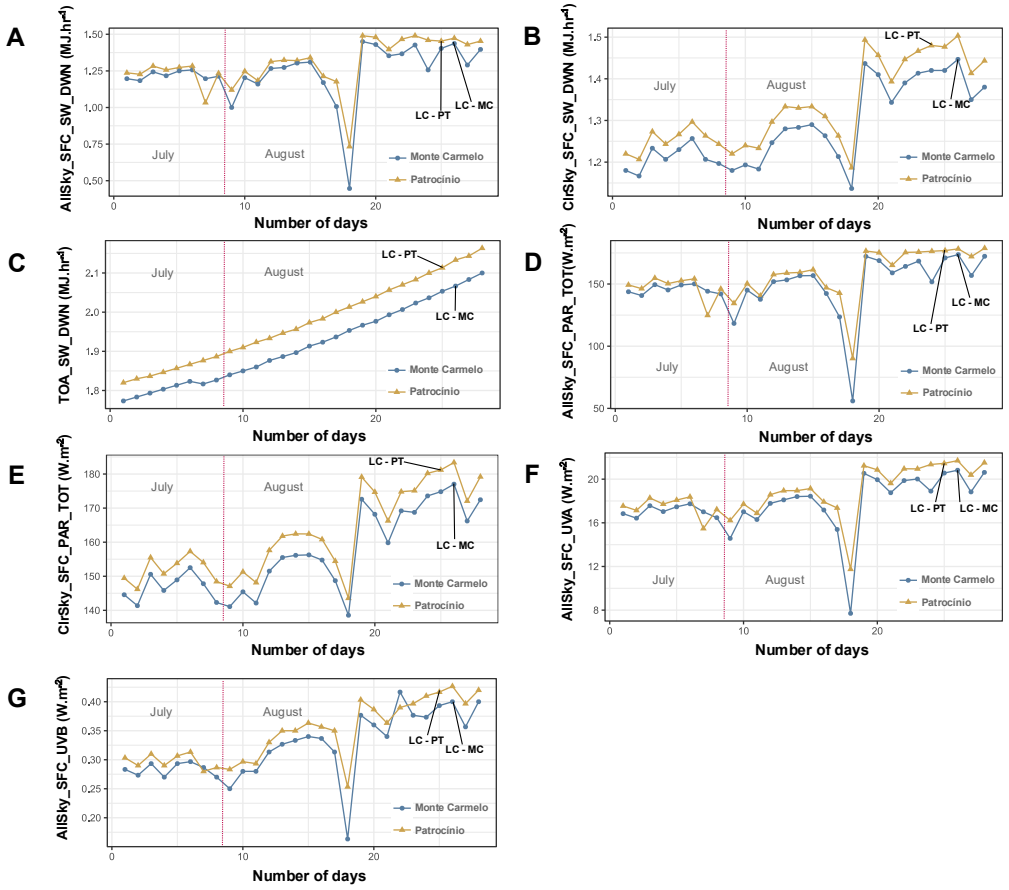
[4]Research Group Metabolic Diversity, Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK Gatersleben), OT Gatersleben, Corrensstraße 3, 06466 Seeland, Germany

[5]UCIBIO Applied Molecular Biosciences Unit, Department of Life Sciences, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.

[6]Associate Laboratory i4HB Institute for Health and Bioeconomy, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.

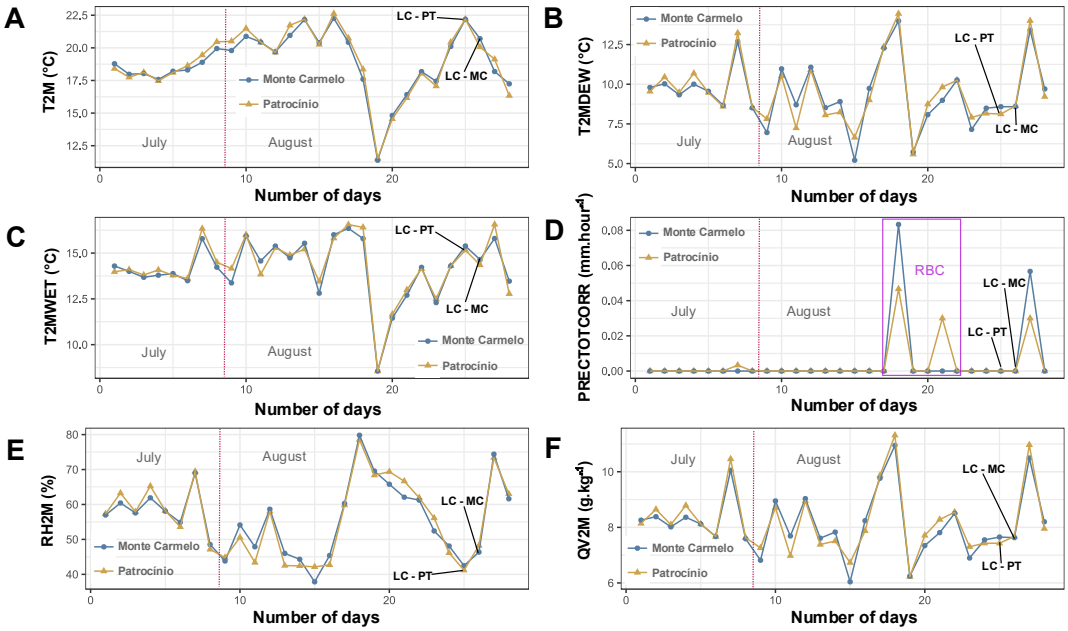
[7]EPAMIG – Empresa de Pesquisa Agropecuária de Minas Gerais, 37200-000 Lavras, Minas Gerais, Brasil.

*Correspondence author: jefynecarrera@gmail.com.



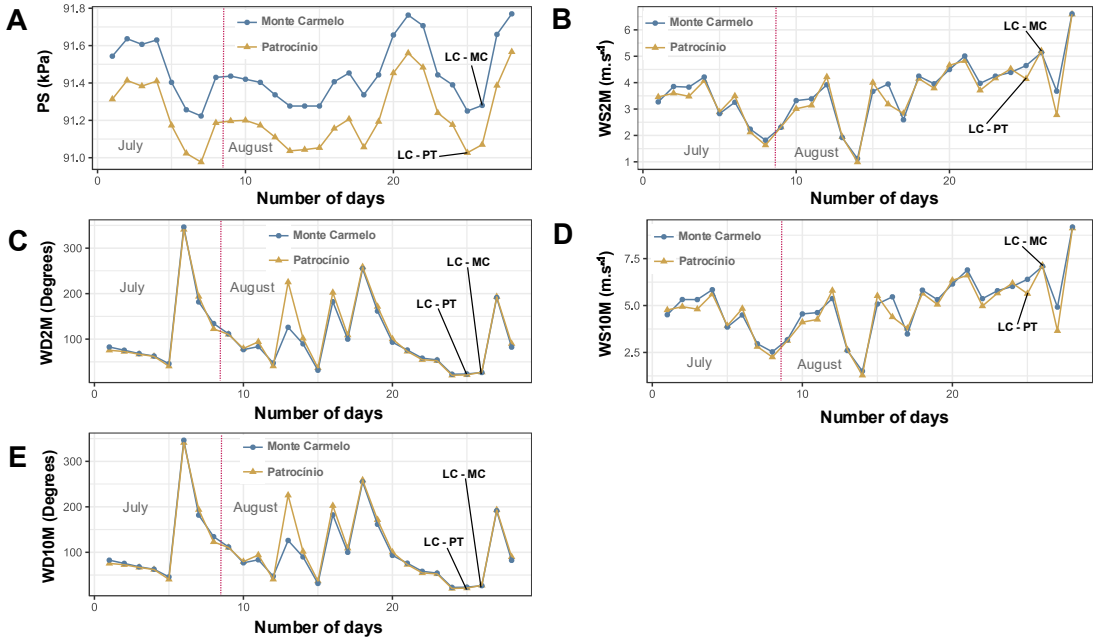
Supplementary Figure S1.

Agroclimatological variables of solar fluxes are related from July 24 to August 20, 2022. The labels (LC-PT) and (LC-MC) indicate the days of leaf collection in Patrocínio and Monte Carmelo, respectively. A: All sky surface shortwave downward irradiance. B: ClrSky_SFC_SW_DWN: Clear sky surface shortwave downward irradiance. C: Top-Of-Atmosphere shortwave downward irradiance. D: All Sky Surface Photosynthetically Active Radiation (PAR) total. E: Clear Sky Surface Photosynthetically Active Radiation (PAR) total. F: All Sky Surface UVA Irradiance (315nm-400nm). G: All Sky Surface UVB Irradiance (280nm-315nm).



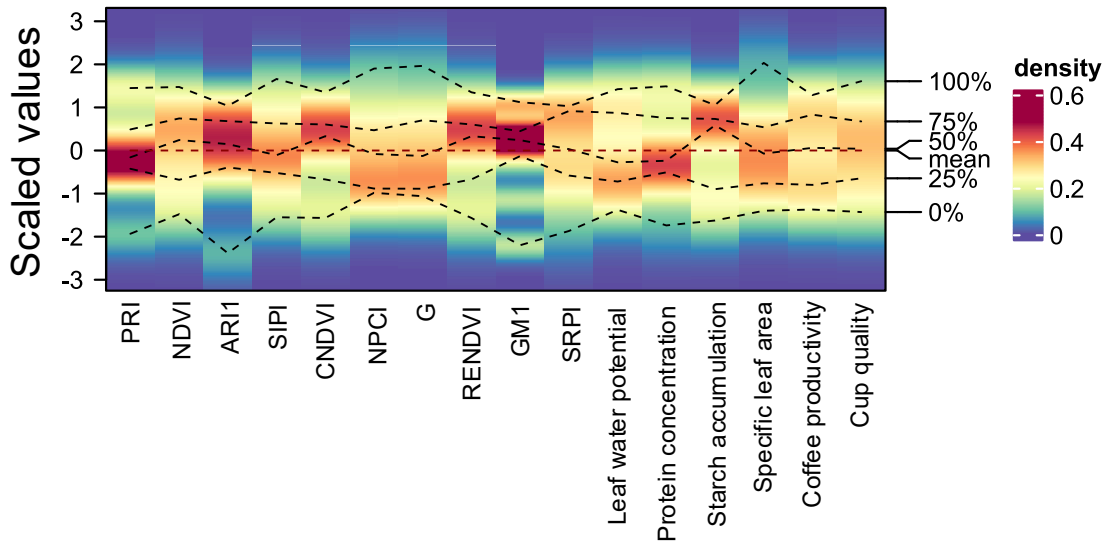
Supplementary Figure S2.

Agroclimatological variables of temperatures, humidity, and precipitation from July 24 to August 20, 2022. The labels (LC-PT) and (LC-MC) indicate the days of leaf collection in Patrocínio and Monte Carmelo, respectively. RBC: Rainfall before collection. A: Temperature at 2 Meters. B: Dew/Frost Point at 2 Meters. C: Wet Bulb Temperature at 2 Meters. D: Precipitation Corrected. E: Relative Humidity at 2 Meters. F: Specific Humidity at 2 Meters.



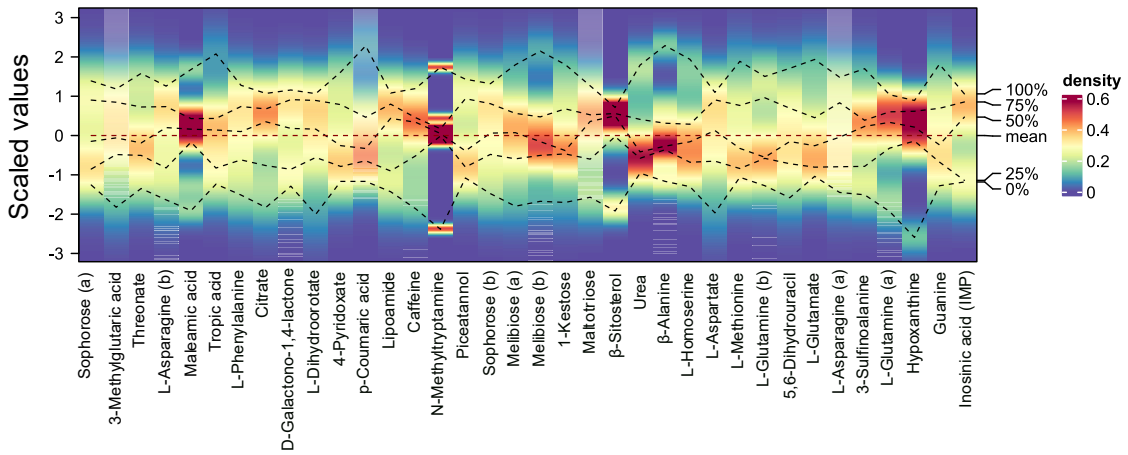
Supplementary Figure S3.

Agroclimatological variables of wind and pressure from July 24 to August 20, 2022. The labels (LC-PT) and (LC-MC) indicate the days of leaf collection in Patrocínio and Monte Carmelo, respectively. A: Surface Pressure. B: Wind Speed at 2 Meters. C: Wind Direction at 2 Meters. D: Wind Speed at 10 Meters. E: Wind Direction at 10 Meters.



Supplementary Figure S4.

Density heatmaps of the leaf characteristics and agronomical variables.



Supplementary Figure S5.

Density heatmaps of the volcano-significant metabolic features.

Supplementary Table S1.

Soil chemical characterization of the experimental locations made at depth of 0–20 cm.

Locality	**pH**	**N**	**K**	**P**	**Na**	**Ca**	**Mg**	**Al**	**H+Al**	**SB**	**e**	**CEC**
	(H_2O)	gKg^{-1}	------ mg dm^{-3}----			---------- $cmol_c$ dm^{-3}----------						
Monte Carmelo	4.8	3.12	194.1	50.6	8.0	2.5	0.9	0.35	8.4	3.9	4.2	12.2
Patrocínio	5.1	3.08	111.9	13.7	3.0	3.9	1.1	0.10	3.9	5.3	5.4	9.2
Locality	**BS**	**m**	**OM**		**P-rem**	**Zn**	**Fe**	**Mn**	**Cu**	**B**	**S**	
	- (%) -		dag kg^{-1}		mg L^{-1}	---------- mg dm^{-3}----------						
Monte Carmelo	33.4	8.95	4.29		23.65	13.3	34.1	12.9	7.3	0.19	34.0	
Patrocínio	57.8	1.86	3.78		11.75	6.7	17.9	26.7	5.9	0.18	36.9	
Locality	**Soil texture (dag kg^{-1})**											
	Clay			**Sand**				**Silt**				
Monte Carmelo	71.0			11.5				17.5				
Patrocínio	41.5			20.5				38.0				

pH in water (1:2.5); Ca^{2+} – Mg^{2+} – Al^{3+}– extractor: KCl, 1mol L^{-1}; P- Na - K- Fe - Zn- Mn- Cu – Mehlich-1 extractor; B: Hot water extraction; S - Extractor - Monocalcium phosphate in acetic acid; H + Al- Extrator: SMP; SB: sum of bases; e: effective cation exchange capacity; CEC:

Cation Exchange Capacity at pH 7.0; BS: base saturation index; m: aluminum saturation index; OM: organic matter (Oxidation: Na2Cr2O7 4N+ H2SO4 10N); P-rem: remaining phosphorus.

Supplementary Table S2.

Spectral indices utilized to evaluate the leaf health status of *Coffea arabica* cultivars in Monte Carmelo and Patrocínio.

Index	Reflectance ratio	Purpose of use	Reference
Anthocyanin Reflectance Index (ARI1)	$(1/R_{550})$ - $(1/R_{700})$	Anthocyanins content	[103]
Cumulated values of NDVI (CNDVI)	CNDVI = ∑NDVI	Plant stresses	[104]
Greenness Index (GI)	R_{554}/R_{667}	Plant stresses	[105]
Gitelson and Merzlyak Index 1 (GM1)	R_{750}/R_{550}	Chlorophyll content	[106]
Normalized Difference Vegetation Index (NDVI)	$(R_{800}-R_{680})$ / $(R_{800}+R_{680})$	Plant stresses	[107]
Normalized pigment chlorophyll ratio index (NPCI)	$(R_{660} - B_{460})$ / $(R_{660} + B_{460})$	Ratio of the total amount of photosynthetic pigments to chlorophyll	[108]
Photochemical Reflectance Index (PRI)	$(R_{531}-R_{570})$ / $(R_{531}+R_{570})$	Changes in carotenoid pigments	[109]
Red-edge Normalized Difference Vegetation Index (ReNDVI)	$(R_{NIR1}-R_{Red\ Edge})$ / $(R_{NIR1}+R_{Red\ Edge})$	Plant water stress	[110]
Structure Insensitive Pigment Index (SIPI)	$(R_{800}-R_{445})$ / $(R_{800}+R_{680})$	Maximizes sensitivity to the bulk carotenoids to chlorophyll ratio	[107]
scaled Photochemical Reflectance Index (sPRI)	(PRI+1)/2	Changes in carotenoid pigments	[111]

Supplementary Table S3.

All 320 Metabolic features, corresponding metabolites and metabolic classes present in coffee leaves (*Coffea arabica* L.) field-grown in Monte Carmelo and Patrocínio, Cerrado Mineiro, Minas Gerais, Brasil.

Feature	Linked metabolite	Metabolic group
2-Piperidinecarboxylic acid (2TMS)	Pipecolic acid	Alkaloids
Caffeine	Caffeine	Alkaloids
Calystegine B2 (1MEOX) (4TMS)	N-Methylcalystegine B2	Alkaloids
Hypoxanthine (2TMS)	Hypoxanthine	Alkaloids
Indole-3-acetic acid (2TMS)	Indole-3-acetate	Alkaloids
Scopolamine (1TMS)	Scopolamine	Alkaloids
Senecionine	Senecionine	Alkaloids
Senecionine (1TMS) BP1	Senecionine	Alkaloids
Senecionine (1TMS) BP2	Senecionine	Alkaloids
Tetrahydroalstonine	Tetrahydroalstonine	Alkaloids
Tetrahydroalstonine (1TMS)	Tetrahydroalstonine	Alkaloids
Theobromine (1TMS)	Theobromine	Alkaloids
Theophylline (1TMS)	Theophylline	Alkaloids
Thiazole, 4-methyl-5-hydroxyethyl- (1TMS)	4-Methyl-5-(2'-hydroxyethyl)-thiazole	Alkaloids
Tryptamine, 1-methyl- (2TMS)	N-Methyltryptamine	Alkaloids
Oxamide (2TMS)	Oxamide	Amides
Guanidine (3TMS)	Guanidine	Amides
Lipoamide, alpha- (2TMS)	Lipoamide	Amides
Maleimide, N-ethyl-	N-Ethylmaleimide	Amides
Urea (2TMS)	Urea	Amides
Histamine, 1-methyl- (2TMS)	1-Methylhistamine	Amines
Histidinol (4TMS)	L-Histidinol	Amines
Phenethylamine (2TMS)	Phenethylamine	Amines
Putrescine (3TMS)	Putrescine	Amines
Putrescine (4TMS)	Putrescine	Amines
Putrescine, N-acetyl- (2TMS)	N-Acetylputrescine	Amines
Spermidine (4TMS)	Spermidine	Amines
Spermidine [?] (5TMS)	Spermidine	Amines
Spermine (5TMS) MP	Spermine	Amines

Tyramine (3TMS)	Tyramine	Amines
1-Pyrroline-2-carboxylate (1TMS)	1-Pyrroline-2-carboxylate	Amino acids
Alanine (2TMS)	L-Alanine	Amino acids
Alanine, beta- (3TMS)	beta-Alanine	Amino acids
Alanine, beta-, N-(2-hydroxyethyl)- (3TMS)	3-(2-hydroxyethylamino)propanoic acid	Amino acids
Asparagine (3TMS)	Asparagine	Amino acids
Asparagine (4TMS) BP2	L-Asparagine	Amino acids
Asparagine (4TMS) MP	L-Asparagine	Amino acids
Aspartic acid (3TMS)	L-Aspartate	Amino acids
Aspartic acid, N-(3-indolylacetyl)- (4TMS)	(Indol-3-yl)acetyl-L-aspartate	Amino acids
Aspartic acid, N-(aminocarbonyl)- (3TMS)	N-Carbamoyl-L-aspartate	Amino acids
Aspartic acid, N-(aminocarbonyl)- (4TMS)	N-Carbamoyl-L-aspartate	Amino acids
Aspartic acid, N-acetyl- (3TMS)	N-Acetyl-L-aspartate	Amino acids
beta-Alanyl-lysine (6TMS)	beta-Alanyl-L-lysine	Amino acids
Butanoic acid, 2-amino- (2TMS)	D-2-Aminobutyrate	Amino acids
Butanoic acid, 4-amino- (2TMS)	gamma-Aminobutyric acid	Amino acids
Butanoic acid, 4-amino- (3TMS)	gamma-Aminobutyric acid	Amino acids
Butanoic acid, 4-amino-3-hydroxy- (3TMS)	4-Amino-3-hydroxybutanoate	Amino acids
Butyric acid, 2,4-diamino-, DL- (3TMS)	L-2,4-Diaminobutanoate	Amino acids
Citrulline (3TMS)	L-Citrulline	Amino acids
Cystathionine (3TMS)	L-Cystathionine	Amino acids
Cystathionine (4TMS)	L-Cystathionine	Amino acids
Cysteinesulfinic acid (3TMS)	L-Cysteinesulfinic acid	Amino acids
Cysteinyl-glycine (3TMS)	L-Cysteinylglycine	Amino acids
Cystine (4TMS)	Cystine	Amino acids
Glutamic acid (3TMS)	L-Glutamate	Amino acids
Glutamic acid, N-acetyl- (2TMS)	N-Acetyl-L-glutamate	Amino acids
Glutamic acid, N-methyl- (3TMS)	N-Methyl-L-glutamate	Amino acids
Glutamine [-H2O] (3TMS) MP	L-Glutamine	Amino acids
Glutamine, DL- (3TMS)	L-Glutamine	Amino acids
Glycine (2TMS)	Glycine	Amino acids

Glycine (3TMS)	Glycine	Amino acids
Glycine, 2-phenyl- (2TMS)	Phenylacetic acid	Amino acids
Hippuric acid (1TMS)	Hippuric acid	Amino acids
Histidine (3TMS)	L-Histidine	Amino acids
Histidine (4TMS)	L-Histidine	Amino acids
Histidine, N-tau-methyl- (2TMS)	1-Methylhistidine	Amino acids
Histidine, N-tau-methyl- (3TMS)	1-Methylhistidine	Amino acids
Homocysteine thiolactone (1TMS)	Homocysteine thiolactone	Amino acids
Homoserine (3TMS)	L-Homoserine	Amino acids
Homoserine lactone, N-dodecanoyl- (1TMS)	N-Dodecanoyl-DL-homoserine lactone	Amino acids
Homoserine, O-succinyl- (3TMS)	O-Succinyl-L-homoserine	Amino acids
Isobutanoic acid, 3-amino- (3TMS)	(R)-3-Amino-2-methylpropanoate	Amino acids
Isoleucine (2TMS)	L-Isoleucine	Amino acids
Leucine (2TMS)	L-Leucine	Amino acids
Levodopa (4TMS)	L-Dopa	Amino acids
Lysine (4TMS)	L-Lysine	Amino acids
Lysine, N-epsilon-acetyl- (3TMS)	N6-Acetyl-L-lysine	Amino acids
Lysine, N6-biotinyl- (5TMS)	N6-D-Biotinyl-L-lysine	Amino acids
Methionine (2TMS)	L-Methionine	Amino acids
Methionine, seleno- (1TMS)	L-Selenomethionine	Amino acids
Ornithine (4TMS)	L-Ornithine	Amino acids
Phenylalanine (2TMS)	L-Phenylalanine	Amino acids
Phenylalanine, N-acetyl- (2TMS)	N-Acetyl-L-phenylalanine	Amino acids
Proline (2TMS)	L-Proline	Amino acids
Proline, 3-hydroxy-, trans- (3TMS)	trans-3-Hydroxy-L-proline	Amino acids
Propanoic acid, 2-amino-2-methyl-3-hydroxy- (2TMS)	2-Methylserine	Amino acids
Propanoic acid, 2-amino-2-methyl-3-hydroxy- (3TMS)	2-Methylserine	Amino acids
Propanoic acid, 3-amino-3-(4-hydroxyphenyl)- (3TMS)	beta-Tyrosine	Amino acids
Saccharopine (5TMS)	L-Saccharopine	Amino acids
Serine (3TMS)	L-Serine	Amino acids
Serine, N-acetyl- (2TMS)	N-Acetylserine	Amino acids

Serine, O-phospho- (4TMS)	O-Phospho-L-serine	Amino acids
Threonine, allo- (3TMS)	L-Allothreonine	Amino acids
Threonine, O-methyl-, L- (2TMS)	O-Methyl-L-threonine	Amino acids
Tryptophan (2TMS)_1	Tryptophan	Amino acids
Tryptophan (2TMS)_2	Tryptophan	Amino acids
Tyrosine (3TMS)	L-Tyrosine	Amino acids
Tyrosine, 3-nitro- (3TMS)	3-Nitrotyrosine	Amino acids
Valine (2TMS)	L-Valine	Amino acids
Orotic acid, 4,5-dihydro- (3TMS)	L-Dihydroorotate	Amino acids
Orotic acid, 4,5-dihydro- (4TMS)	L-Dihydroorotate	Amino acids
Abscisic acid (1MEOX) (1TMS)_1	(+)-Abscisic acid	Lipids
Abscisic acid (1MEOX) (1TMS)_2	(+)-Abscisic acid	Lipids
Acetoacetic acid (1MEOX) (1TMS) BP	Acetoacetate	Lipids
Butanoic acid, 2-hydroxy- (2TMS)	2-Hydroxybutyric acid	Lipids
Butanoic acid, 2-oxo- (1MEOX) (1TMS) MP	2-Oxobutanoate	Lipids
Butanoic acid, 3-hydroxy- (2TMS)	D-beta-Hydroxybutyric acid	Lipids
Ergosterol (1TMS)	Ergosterol	Lipids
Fumaric acid, 2-methyl- (2TMS)	Mesaconate	Lipids
Gibberellin A4 (2TMS)	Gibberellin A4	Lipids
Glutaric acid (2TMS)	Glutaric acid	Lipids
Glycerol-2-phosphate (4TMS)	Glycerol 2-phosphate	Lipids
Glycerol-3-phosphate (4TMS)	Glycerol-3-phosphate	Lipids
Hexadecan-1-ol, n- (1TMS)	1-Hexadecanol	Lipids
Homoserine lactone, N-heptanoyl-	N-Heptanoylhomoserine lactone	Lipids
Homoserine lactone, N-hexanoyl-	N-Hexanoyl-L-homoserine lactone	Lipids
Homoserine lactone, N-octanoyl-	N-Octanoyl-L-homoserine lactone	Lipids
Levulinic acid, 5-amino- (1MEOX) (3TMS) MP	5-Aminolevulinate	Lipids
Lipoamide, alpha- (1TMS)	Lipoamide	Lipids
Maleamic acid (2TMS) BP	Maleamic acid	Lipids
Malic acid, 2-methyl- (3TMS)	D-Citramalate	Lipids
Menthol, L-(-)- (1TMS)	L-Menthol	Lipids

Nonacosanoic acid (1TMS)	Nonacosanoic acid	Lipids
Octanoic acid, 3-hydroxy- (2TMS)	3-Hydroxyoctanoate	Lipids
Octanoic acid, 8-amino- (4TMS)	8-Aminooctanoic acid	Lipids
Octanoic acid, n- (1TMS)	Octanoic acid	Lipids
Oxaloacetate (1MEOX) (2TMS) BP	Oxalacetic acid	Lipids
Pentanedioic acid, 3-methyl- (2TMS)	3-Methylglutaric acid	Lipids
Perillic acid (1TMS)	Perillic acid	Lipids
Phylloquinone	Phylloquinone	Lipids
Phytol (1TMS) MP	Phytol	Lipids
Secologanin (TMS) MP	Secologanin	Lipids
Sitosterol, beta- (1TMS)	Beta-Sitosterol	Lipids
Squalene, all-trans-	Squalene	Lipids
Succinic acid, 2,3-dimethyl- (2TMS)	2,3-Dimethylsuccinic acid	Lipids
4-(Methylamino)benzoic acid (2TMS)	N-Methyl-4-aminobenzoate	Other organic acids
Acetic acid, 3-hydroxyphenyl- (3TMS)	3-Hydroxyphenylacetate	Other organic acids
Aconitic acid, cis- (3TMS)	cis-Aconitate	Other organic acids
Anthranilic acid, 3-hydroxy- (3TMS)	3-Hydroxyanthranilate	Other organic acids
Benzoic acid, 3-hydroxy- (2TMS)	3-Hydroxybenzoic acid	Other organic acids
Butanoic acid, 4-hydroxy- (2TMS)	(R)-3-Hydroxybutanoate	Other organic acids
Citric acid (4TMS)	Citrate	Other organic acids
Cyclohexane-1,2-dicarboxylic acid, trans- (2TMS)	1,2-Cyclohexanedicarboxylic acid	Other organic acids
Fumaric acid (2TMS)	Fumarate	Other organic acids
Glutaric acid, 2-hydroxy- (3TMS)	(R)-2-Hydroxyglutarate	Other organic acids
Glutaric acid, 2-oxo- (1MEOX) (2TMS) MP	2-Oxoglutarate	Other organic acids
Glycolic acid (2TMS)	Glycolic acid	Other organic acids
Hippuric acid, 2-hydroxy- (3TMS)	Salicyluric acid	Other organic acids
Isocaproic acid, 2-oxo- (1MEOX) (1TMS) MP	4-Methyl-2-oxopentanoate	Other organic acids
Isovaleric acid, 2-oxo- (1MEOX) (1TMS) MP	2-Keto-3-methylbutyric acid	Other organic acids
Kynurenic acid (2TMS)	Kynurenic acid	Other organic acids
Lactic acid, DL- (2TMS)	L-Lactic acid	Other organic acids

Malic acid (3TMS)	L-Malate	Other organic acids
Malic acid, 3-oxalo- (1MEOX) (4TMS) MP	3-Oxalomalate	Other organic acids
Malic acid, 3-oxalo- (TMS)	3-Oxalomalate	Other organic acids
Nicotinic acid (1TMS)	Nicotinate	Other organic acids
Nicotinic acid, 6-hydroxy- (2TMS)	6-Hydroxynicotinate	Other organic acids
Octopamine (4TMS)	Octopamine	Other organic acids
Oxaloacetate (1MEOX) (3TMS) MP	Oxaloacetate	Other organic acids
Oxamic acid (2TMS)	Oxamate	Other organic acids
Pyrrole-2-carboxylic acid (2TMS)	Pyrrole-2-carboxylate	Other organic acids
Pyruvic acid (1MEOX) (1TMS)	Pyruvate	Other organic acids
Pyruvic acid (2TMS)	Pyruvate	Other organic acids
Pyruvic acid, 4-hydroxyphenyl- (1MEOX) (2TMS) MP	4-Hydroxyphenylpyruvate	Other organic acids
Shikimic acid (4TMS)	Shikimate	Other organic acids
Shikimic acid-3-phosphate (5TMS)	Shikimate 3-phosphate	Other organic acids
Succinic acid (2TMS)	Succinate	Other organic acids
Tartronic acid (3TMS)	Tartronic acid	Other organic acids
Tartronic acid (4TMS)	Tartronic acid	Other organic acids
Tropic acid (2TMS)	Tropic acid	Other organic acids
1,4-Naphthoquinone, 2-methyl-	Menadione	Others
3-Indoleacetic acid, 5-methoxy- (2TMS)	5-Methoxyindoleacetate	Others
4-Pyridoxic acid (3TMS)	4-Pyridoxate	Others
Acetic acid, 4-chlorophenyl- (1TMS)	4-Chlorophenylacetic acid	Others
Adenine (2TMS)	Adenine	Others
Adenosine, 2*-deoxy- (3TMS)	Deoxyadenosine	Others
Adenosine, alpha- (4TMS) BP	Adenosine	Others
Ascorbic acid (4TMS)	L-Ascorbic acid	Others
Benzimidazole, 5,6-dimethyl- (1TMS)	Dimethylbenzimidazole	Others
Benzylalcohol (1TMS)	Benzyl alcohol	Others
Butyro-1,4-lactam, 2-amino- (2TMS)	3-Aminopyrrolidin-2-one	Others
Cystineamine (4TMS)	Cystamine	Others
Dehydroascorbic acid dimer (2MEOX) MP	Dehydroascorbic acid	Others
Dihydrozeatin riboside (4TMS)	Dihydrozeatin riboside	Others

Farnesol (1TMS) BP2	2-trans,6-trans-Farnesol	Others
Galactonic acid-1,4-lactone (4TMS)	D-Galactono-1,4-lactone	Others
Glucaric acid-1,4-lactone (4TMS)	D-Glucaro-1,4-lactone	Others
Guanine (3TMS)	Guanine	Others
Guanosine-3*,5*-cyclic-monophosphate (4TMS)	2',3'-Cyclic GMP	Others
Gulonic acid-1,4-lactone (4TMS)	L-Galactono-1,4-lactone	Others
Hydantoin, 1-methyl- (2TMS)	N-Methylhydantoin	Others
Imidazole-4-acetic acid, 1-methyl- (1TMS)	Methylimidazoleacetic acid	Others
Indole-3-acetamide (3TMS)	Indole-3-acetamide	Others
Indole-3-acetonitrile (1TMS)	3-Indoleacetonitrile	Others
Inosine (4TMS)	Inosine	Others
Inosine-5*-monophosphate (5TMS)	Inosinic acid	Others
Inositol, allo- (6TMS)	Inositol	Others
Inositol, myo- (6TMS)	Inositol	Others
Kynurenine (3TMS)	L-Kynurenine	Others
myo-Inositol-1-phosphate (7TMS)	myo-Inositol 1-phosphate	Others
Octacosane, n-	Octacosane	Others
Orotic acid (3TMS)	Orotic acid	Others
Pantothenic acid, D- (4TMS)	Pantothenate	Others
Phylloquinone (2TMS)	Phylloquinone	Others
Propane-1,2-diol, 3-amino (4TMS)	3-Aminopropane-1,2-diol	Others
Propane-1,3-diol, 2-amino-2-methyl- (3TMS)	2-Amino-2-methyl-1,3-propandiol	Others
Pyridine, 2-hydroxy- (1TMS)	2-Hydroxypyridine	Others
Pyridine, 3-hydroxy- (1TMS)	3-Hydroxypyridine	Others
Pyridoxal (1MEOX) (2TMS) MP	Pyridoxal	Others
Pyridoxine (3TMS)	Pyridoxine	Others
Quinic acid (5TMS)	Quinic acid	Others
Quinoline-2-carboxylic acid, 4,8-dihydroxy- (3TMS)	Xanthurenic acid	Others
Serine, cyclo- (2TMS)	Cycloserine	Others
Tartronic acid, 2-(methylaminomethyl)- (1MEOX) (2TMS)	2-(Methylaminomethyl)tartronic acid	Others

Tartronic acid, 2-(methylaminomethyl)-(1MEOX) (3TMS)	2-(Methylaminomethyl)tartronic acid	Others
Tetracosane, n-	Lignocerane	Others
Tetratriacontane	Tetratriacontane	Others
Uracil, dihydro- (2TMS)	Dihydrouracil	Others
Uridine (3TMS)	Uridine	Others
Valeric acid, 3-hydroxy-3-methyl- (1TMS)	3-Hydroxy-3-methylpentanoic acid	Others
Valero-1,5-lactam (1TMS)	N-Methyl-2-pyrrolidinone	Others
Lactic acid, 3-(4-hydroxyphenyl)- (3TMS)	(±)-2-Hydroxy-3-(2-hydroxyphenyl)propanoic acid	Phenylpropanoids or phenolic compounds
3-(2,4-Dihydroxyphenyl)-propanoic acid (3TMS)	3-(2,4-Dihydroxyphenyl)propanoic acid	Phenylpropanoids or phenolic compounds
Acetic acid, 3,4-dihydroxyphenyl- (3TMS)	3,4-Dihydroxyphenylacetate	Phenylpropanoids or phenolic compounds
Benzoic acid, 4-hydroxy-(2TMS)	4-Hydroxybenzoate	Phenylpropanoids or phenolic compounds
Caffeic acid, cis- (3TMS)	Caffeate	Phenylpropanoids or phenolic compounds
Caffeic acid, trans- (3TMS)	trans-Caffeate	Phenylpropanoids or phenolic compounds
Catechin (5TMS)	(+)-Catechin	Phenylpropanoids or phenolic compounds
Cinnamaldehyde, 3,5-dimethoxy-4-hydroxy-, trans-(1MEOX) (1TMS) BP	Sinapaldehyde	Phenylpropanoids or phenolic compounds
Cinnamic acid, 2,5-dimethoxy-, trans- (1TMS)	2,5-Dimethoxycinnamic acid	Phenylpropanoids or phenolic compounds
Cinnamic acid, 3-methoxy-, trans- (1TMS)	3-Methoxycinnamic acid	Phenylpropanoids or phenolic compounds
Cinnamic acid, 3,4,5-trimethoxy-, trans- (1TMS)	3,4,5-Trimethoxycinnamic acid	Phenylpropanoids or phenolic compounds
Cinnamic acid, 3,5-dimethoxy-, trans- (1TMS)	3,5-Dimethoxycinnamic acid	Phenylpropanoids or phenolic compounds
Cinnamic acid, 4-hydroxy-, trans- (2TMS)	p-Coumaric acid	Phenylpropanoids or phenolic compounds
Cinnamic acid, 4-methoxy-, trans- (1TMS)	4-Methoxycinnamic acid	Phenylpropanoids or phenolic compounds
Coniferylalcohol, trans- (2TMS)	Coniferyl alcohol	Phenylpropanoids or phenolic compounds
Coniferylaldehyde, trans-(1MEOX) (1TMS) BP	Coniferyl aldehyde	Phenylpropanoids or phenolic compounds
Coumarine	Coumarin	Phenylpropanoids or phenolic compounds
Epigallocatechin (6TMS)	(-)-Epigallocatechin	Phenylpropanoids or phenolic compounds

Eriodictyol (5TMS)	Eriodictyol	Phenylpropanoids or phenolic compounds
Ethanol, 2-(4-hydroxyphenyl)- (2TMS)	4-Hydroxyphenylethanol	Phenylpropanoids or phenolic compounds
Ferulic acid, trans- (2TMS)	Ferulate	Phenylpropanoids or phenolic compounds
Flavone, 3-hydroxy- (1TMS)	3-Hydroxyflavone	Phenylpropanoids or phenolic compounds
Hydrocaffeic acid (3TMS)	3,4-Dihydroxyphenylpropanoate	Phenylpropanoids or phenolic compounds
Hydrocinnamic acid (1TMS)	3-Phenylpropanoate	Phenylpropanoids or phenolic compounds
Hydroquinone (2TMS)	Hydroquinone	Phenylpropanoids or phenolic compounds
Isoflavone, 5,7-dihydroxy-4*-methoxy- (2TMS)	Biochanin A	Phenylpropanoids or phenolic compounds
Mandelic acid, 3,4-dihydroxy- (4TMS)	3,4-Dihydroxymandelic acid	Phenylpropanoids or phenolic compounds
Myricetin (6TMS)	Myricetin	Phenylpropanoids or phenolic compounds
Piceatannol (4TMS) BP	Piceatannol	Phenylpropanoids or phenolic compounds
Propanoic acid, 3-(2-hydroxyphenyl)- (2TMS)	3-(2-Hydroxyphenyl)propanoic acid	Phenylpropanoids or phenolic compounds
Pyruvic acid, 4-hydroxyphenyl- (1MEOX) (3TMS) MP	4-Hydroxyphenylpyruvate	Phenylpropanoids or phenolic compounds
Quercetin (5TMS)	Quercetin	Phenylpropanoids or phenolic compounds
Quinic acid, 3-caffeoyl-, trans- (6TMS)	Chlorogenic acid	Phenylpropanoids or phenolic compounds
Quinic acid, 4-caffeoyl-, trans- (6TMS)	Cryptochlorogenic acid	Phenylpropanoids or phenolic compounds
Quinic acid, 5-caffeoyl-, trans- (6TMS)	Chlorogenate	Phenylpropanoids or phenolic compounds
Salicylic acid (2TMS)	Salicylic acid	Phenylpropanoids or phenolic compounds
Shikimic acid, 3-dehydro- (1MEOX) (3TMS) MP	3-Dehydroshikimate	Phenylpropanoids or phenolic compounds
Sinapic acid, trans- (2TMS)	Sinapate	Phenylpropanoids or phenolic compounds
Vanillic acid (2TMS)	Vanillate	Phenylpropanoids or phenolic compounds
1,3-Dihydroxyaceton (1MEOX) (2TMS) MP	Dihydroxyacetone	Sugars
4-Hydroxyphenyl-beta-glucopyranoside (5TMS)	Arbutin	Sugars
Allose (1MEOX) (5TMS) MP	D-Allose	Sugars
Arabitol (5TMS)	D-Arabitol	Sugars

Digitoxose (1MEOX) (3TMS) BP	D-Digitoxose	Sugars
Digitoxose (1MEOX) (3TMS) MP	D-Digitoxose	Sugars
Erythronic acid (4TMS)	D-Erythronate	Sugars
Erythrose (1MEOX) (3TMS) MP	D-Erythrose	Sugars
Fructose (1MEOX) (5TMS) BP	D-Fructose	Sugars
Fructose-1-phosphate (1MEOX) (6TMS) MP	D-Fructose 1-phosphate	Sugars
Fructose-1-phosphate (6TMS) MP	D-Fructose 1-phosphate	Sugars
Fructose-1,6-diphosphate (1MEOX) (7TMS) BP	D-Fructose 1,6-bisphosphate	Sugars
Fucose (1MEOX) (4TMS) MP	2-Deoxyglucose	Sugars
Galactaric acid (6TMS)	D-Galactarate	Sugars
Galactinol (9TMS)	Galactinol	Sugars
Galactosamine, N-acetyl- (1MEOX) (5TMS) BP	N-Acetyl-D-galactosamine	Sugars
Galactosamine, N-acetyl- (1MEOX) (5TMS) MP	N-Acetylgalactosamine	Sugars
Galactose, 3,6-anhydro- (1MEOX) (3TMS)	3,6-Anhydrogalactose	Sugars
Gentiobiose (1MEOX) (8TMS) MP	Gentiobiose	Sugars
Glucoheptose (1MEOX) (6TMS) MP	d-Glycero-d-galacto-heptose	Sugars
Gluconic acid-1,5-lactone (4TMS)	D-Glucono-1,5-lactone	Sugars
Gluconic acid, 2-amino-2-deoxy- (7TMS)	2-Amino-2-deoxy-D-gluconate	Sugars
Gluconic acid, 2-oxo- (1MEOX) (5TMS) BP	2-Dehydro-D-gluconate	Sugars
Glucopyranoside, 1-O-methyl-, alpha- (4TMS)	Methyl alpha-D-galactopyranoside	Sugars
Glucosamine, N-acetyl- (1MEOX) (4TMS)	2-Acetamido-2-deoxy-D-glucose	Sugars
Glucose (1MEOX) (5TMS) BP	D-Glucose	Sugars
Glucose (1MEOX) (5TMS) MP	D-Glucose	Sugars
Glucose-6-phosphate (1MEOX) (6TMS) MP	Glucose 6-phosphate	Sugars
Glucose-6-phosphate, 2-amino-2-deoxy- (6TMS)	D-Glucosamine 6-phosphate	Sugars
Glucose, 2-amino-2-deoxy- (5TMS) BP	2-Amino-2-deoxy-D-glucose	Sugars
Glucose, 2-amino-2-deoxy-, D- (1MEOX) (5TMS)	2-Amino-2-deoxy-D-glucose	Sugars

Glyceraldehyde (1MEOX) (2TMS) BP	D-Glyceraldehyde	Sugars
Glyceraldehyde (1MEOX) (2TMS) MP	D-Glyceraldehyde	Sugars
Glyceric acid (3TMS)	L-Glycerate	Sugars
Glycerol (3TMS)	Glycerol	Sugars
Isomaltose (1MEOX) (8TMS) MP	Isomaltose	Sugars
Kestose, 1- (11TMS)	1-Kestose	Sugars
Lactose, alpha- (1MEOX) (8TMS) MP	Lactose	Sugars
Leucrose (1MEOX) (8TMS) MP	Leucrose	Sugars
Maltotriose (1MEOX) (11TMS) MP	Maltotriose	Sugars
Melibiose (1MEOX) (8TMS) BP	Melibiose	Sugars
Melibiose (1MEOX) (8TMS) MP	Melibiose	Sugars
N,N*-Diacetylchitobiose (1MEOX) (6TMS) MP	Chitobiose	Sugars
Psicose (1MEOX) (5TMS) MP	D-Psicose	Sugars
Rhamnose (1MEOX) (4TMS) MP	L-Rhamnose	Sugars
Ribonic acid (5TMS)	D-Ribonate	Sugars
Ribonic acid-1,4-lactone (3TMS)	D-Arabinono-1,4-lactone	Sugars
Ribose-5-phosphate (1 MEOX) (5TMS) BP	D-Ribose 5-phosphate	Sugars
Ribose, 2-deoxy- (1MEOX) (3TMS) BP	Deoxyribose	Sugars
Ribose, 2-deoxy- (1MEOX) (3TMS) MP	2-Deoxy-D-ribose	Sugars
Ribulose (1MEOX) (4TMS) BP2	D-Ribulose	Sugars
Sophorose (1MEOX) (8TMS) BP	Sophorose	Sugars
Sophorose (1MEOX) (8TMS) MP	Sophorose	Sugars
Sucrose (8TMS)	Sucrose	Sugars
Tagatose (1MEOX) (5TMS) BP	D-Tagatose	Sugars
Threitol (4TMS)	D-Threitol	Sugars
Threonic acid (4TMS)	Threonate	Sugars
Xylose, D- (1MEOX) (4TMS)	alpha-D-Xylose	Sugars
Xylulose (1MEOX) (4TMS) MP	D-Xylulose	Sugars

Xylulose-5-phosphate (1MEOX) (5TMS) BP	D-Xylulose 5-phosphate	Sugars

ARTIGO 3: COMPARATIVE MORPHOANATOMICAL, METABOLOMIC, AND PROTEOMIC OF *Coffea arabica* L. CULTIVARS UNDER IRRIGATED AND RAINFED CONDITIONS IN THE CERRADO MINEIRO REGION, BRAZIL

Artigo redigido de acordo com as normas da *The Crop Journal*

Comparative Morphoanatomical, Metabolomic, and Proteomic of *Coffea arabica* L. Cultivars under Irrigated and Rainfed Conditions in the Cerrado Mineiro Region, Brazil

Jéfyne Campos Carréra[1,*], Leonor Guerra-Guimarães[2,3], Mário Lúcio Vilela Resende[1], Jenny Renault[4], Céline Leclercq[4], Paulo Henrique Sales Guimarães[1], John Charles D'Auria[5], Carla Pinheiro[6,7], Luana de Jesus Sartori[1], Vânia Aparecida Silva[8], Margarete Lordelo Volpato[8], Christiano de Sousa Machado de Matos[8], Gladyston Rodrigues Carvalho[8], Fabio Akira Mori[1]

[1]UFLA - Universidade Federal de Lavras, 37203-202 Lavras, Minas Gerais, Brasil.

[2]CIFC - Centro de Investigação das Ferrugens do Cafeeiro, Jardim Botânico da Ajuda, Instituto Superior de Agronomia, Universidade de Lisboa, Calçada da Ajuda, 1300-011 Lisboa, Portugal.

[3]LEAF - Linking Landscape, Environment, Agriculture and Food Research Centre and TERRA Associated Laboratory, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017, Lisboa, Portugal.

[4]Luxembourg Institute of Science and Technology, rue Bommel, 4940 Hautcharage, Luxemburg.

[5]Research Group Metabolic Diversity, Department of Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK Gatersleben), OT Gatersleben, Corrensstraße 3, 06466 Seeland, Germany.

[6]UCIBIO Applied Molecular Biosciences Unit, Department of Life Sciences, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.

[7]Associate Laboratory i4HB Institute for Health and Bioeconomy, NOVA School of Science and Technology, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal.

[8]EPAMIG – Empresa de Pesquisa Agropecuária de Minas Gerais, 37200-000 Lavras, Minas Gerais, Brasil.

*Correspondence author: Jéfyne Campos Carréra, Universidade Federal de Lavras, 37203-202 Lavras, Minas Gerais, Brasil. E-mail: jefynecarrera@gmail.com. https://orcid.org/0000-0001-6597-7152.

ORCID List

Leonor Guerra-Guimarães: https://orcid.org/0000-0002-9676-1036

Mário Lúcio Vilela Resende: https://orcid.org/0000-0003-3357-5057

Jenny Renault: https://orcid.org/0000-0002-0450-3866

Céline Leclercq: https://orcid.org/0000-0003-0565-4591

Paulo Henrique Sales Guimarães: https://orcid.org/0000-0001-9158-1688

John Charles D'Auria: https://orcid.org/0000-0002-4865-3938

Luana de Jesus Sartori: https://orcid.org/0000-0003-0764-977X

Carla Pinheiro: https://orcid.org/0000-0002-7799-0183

Vânia Aparecida Silva: https://orcid.org/0000-0001-6640-3644
Margarete Lordelo Volpato: https://orcid.org/0000-0002-6905-6876
Christiano de Sousa Machado de Matos: https://orcid.org/0000-0001-7844-3352
Gladyston Rodrigues Carvalho: https://orcid.org/0000-0002-4466-674X
Fabio Akira Mori: https://orcid.org/0000-0002-7468-018X

Abstract

Understanding how coffee cultivars respond to varying water availability is essential for improving crop resilience and productivity under climate variability. This study investigated the leaf water status, morphoanatomical characteristics, as well as the metabolome, proetome and enzymatic activity of *Coffea arabica* cultivars grown under irrigated and rainfed conditions in two experimental sites in the Cerrado Mineiro region, Minas Gerais, Brazil. Significant differences were observed in leaf water potential and key morphoanatomical traits, notably specific leaf area (higher in irrigated plants) and palisade parenchyma thickness (greater in rainfed plants). Invertase activity was generally higher in rainfed plants, except in irrigated Catiguá MG2 and Sarchimor. Furthermore, the cultivars exhibited distinct metabolomic and proteomic profiles under the different irrigation regimes. Among the differentially expressed metabolites, lipids, sugars, and amino acids were the predominant classes. In the proteomic analysis, the most represented biological processes were associated with protein metabolism—including biosynthesis, modification, and homeostasis—and photosynthesis-related pathways. Subcellular localization analysis revealed that most differentially expressed proteins were targeted to chloroplasts or the extracellular space. The most pronounced differences in metabolomic and proteomic profiles were observed in Paraíso 2 and Catiguá MG2, which exhibited 236 and 234 differentially expressed metabolic features and proteins, respectively. Stress-related pathways were more active in Catiguá MG2 plants with low leaf water potential, while in Paraíso 2 acclimation responses were observed under both conditions. These findings highlight cultivar-specific plasticity and offer insights for enhancing coffee resilience through breeding and water management in an important specialty coffee producing region.

Keywords: Coffee plant, crop resilience, stress-response.

1. Introduction

Periodic leaf structure and metabolic evaluations of important crops are critical for optimizing agricultural performance. Changes in leaf anatomical traits and specific leaf area can markedly influence photosynthetic efficiency, biomass distribution, and stress tolerance [1,2]. Additionally, adjustments in plant energy metabolism and protein synthesis may shift

resource allocation from growth to protective mechanisms under stress conditions [3]. This phenotypic plasticity enables crop plants to adapt their growth and developmental processes to dynamic environmental conditions [4].

The morphoanatomical characteristics and metabolism of *Coffea arabica*, a globally significant crop of which Brazil is the largest producer [5], have been extensively investigated in studies examining plant responses to stress [6, 7, 8, 9, 10, 11], nutrient variability [12, 13], implications on beverage quality [14, 15], as well as, to evaluate changes in protein profiles in response to different post-harvest processing methods [16] and elucidate mechanisms of disease resistance, including the identification of proteins associated with resistance to coffee leaf rust (*Hemileia vastatrix*) [17, 18]. Furthermore, na integrative approach of proteomic and metabolomic analyses offers a more comprehensive understanding of the key metabolic pathways affected in coffee plants subjected to stress conditions [19].

In Minas Gerais, Brazil's leading domestic coffee producer, coffee cultivation is concentrated in the Cerrado Mineiro, Sul de Minas, and Zona da Mata regions [20]. The Cerrado Mineiro region, renowned for its specialty coffee production, is characterized by a pronounced dry season, with lower rainfall rates from May to September, higher reference evapotranspiration rates from July to October, these conditions lead to less water accumulation in the soil during this period [21]. To mitigate the effects of water stress and optimize crop performance, controlled irrigation practices are commonly employed in this region to regulate the growth and reproductive development of various coffee cultivars [22].

Coffee plants in the Cerrado Mineiro region experience moderate water stress [23], characterized by leaf water potential values around -1.1 MPa, during the onset of the dry season [24]. Regarding leaf plasticity in response to water availability, certain coffee genotypes and cultivars exhibit greater variability [7, 23]. Moreover, anatomical variations contribute to maintaining leaf water status under prolonged water deficit conditions [7].

According to water availability, coffee plants activate their antioxidant defense systems [23], exhibit an increased abundance of photosynthesis-related proteins [25], and modulate carbohydrate metabolism [26, 27]. These metabolic adjustments support the production of specialized metabolites and energy, functioning as mechanisms to enhance drought tolerance [10]. Therefore, the objective of this study was to investigate potential variations in anatomical and metabolic traits of *Coffea arabica* cultivars grown under irrigated and rainfed conditions in two experimental sites in the Cerrado region of Minas Gerais during the dry season, Brazil.

2. Material and methods

2.1.Field Experiments

The experimental sites were strategically chosen upon the establishment of Coffee Demonstration units, a collaborative endeavor between EPAMIG and local farming communities. Within this project, a standardized set of cultivars was simultaneously planted in multiple municipalities within the Cerrado Mineiro region. The primary objective of this initiative is to evaluate the field performance of these cultivars *Coffea arabica* plants, that have been grown since 2016 in the experimental germplasm fields of EPAMIG, located in the municipality of Patrocínio (18° 57' S, 46°54' W) and private farm in the city of Monte Carmelo (18°56'10" S, 47°22'36" W) [21]. In the experimental site in Patrocínio, the plant cultivation followed rainfed practices, while in Monte Carmelo, irrigation was employed. The irrigation system consists of surface-level Ram drippers spaced at intervals of 0.7 meters, each delivering a flow rate of 2.3 L h^{-1}. The plants were irrigated every 48 hours over a three-hour period, whereby received approximately 9 L of water. Overall, the available water capacity (AWC) is greater in Monte Carmelo than in Patrocínio [21]. The experimental sites in Patrocínio and Monte Carmelo are approximately 1,000 meters above sea level, with an elevation difference of up to 100 meters between them. Monte Carmelo occupies the higher elevation.

2.2.Plant Material

The five *C. arabica* cultivars selected for this study were chosen based on specific characteristics, including resistance to *Hemileia vastatrix* (coffee rust) and *Meloidogyne exigua* (nematode), as well as their productivity and cup quality (Table 1). The plant spacing is 3.8 m x 0.6 m. At the two experimental sites (Monte Carmelo and Patrocinio), expanded leaves from the middle third of five plant replicates per cultivar were collected in August 2022. Leaf collection started at 5 a.m. to assess pre-dawn water potential, and a second collection occurred between 7 and 9 a.m. for metabolomic, proteomic assays and morphoanatomical analysis. The collection order of cultivars remained consistent across both experimental sites to the sun-exposed side of the plant. Subsequently, the midribs were excised from these leaves, and the remaining material was preserved using liquid nitrogen. For anatomical analysis, the collected plant material were fixed and preserved in 70% ethanol.

2.3.Water potential measurement

Pre-drawn water potential analysis and leaf spectral vegetative indices were obtained from the same leaf sample. Measurements were conducted on fully expanded leaves from the 3rd or 4th pair from the tip of an actively growing plagiotropic branch. The leaf water potential

(measured in MPa) was determined using a Scholander pressure chamber (1000 PMS Instruments Plant Moisture) [28].

2.4.Specific leaf area

The leaves utilized for assessing leaf water potential were stored in Kraft paper bags to facilitate drying and subsequent determination of specific leaf area (SLA; $cm^2\ g^{-1}$). The leaf area was quantified by scanning the leaves using a flatbed scanner, followed by image analysis with FIJI (ImageJ) software [29]. The collected material was oven-dried at 60 °C for 48 hours and weighed on a precision scale to determine the dry biomass. The specific leaf area was calculated as the ratio of the leaf area to dry biomass.

2.5.Leaf anatomical analysis

Fully expanded leaves from the third or fourth node from the tip of an actively growing plagiotropic branch were selected for anatomical analysis. Cross-sections, encompassing the midrib region, were obtained using a freehand technique. Sections were clarified with sodium hypochlorite, rinsed, stained with a mixture of astra blue and safranin [30], mounted on slides with glycerin water, and sealed. Paradermal sections were excised from the leaf margin, incubated in sodium hypochlorite to dissociate the epidermis, rinsed, stained with 1% safranin [31], and mounted on slides with 50% glycerin. Slides were examined under a Nikon Eclipse E100 microscope equipped with a Lumenera Infinity 13.1 mp camera. FIJI (ImageJ) software [29] was employed to measure the following anatomical parameters for cross-sections: Thickness of leaf, epidermis, cuticle, mesophyll, palisade parenchyma, spongy parenchyma, area of xylem, phloem and xylem vessel diameter. For paradermal sections, the following parameters were measured: stomatal pore area, stomatal polar and equatorial diameters, and stomatal index.

2.6.Leaf extract

The leaf samples (without the middle vein) were ground with liquid nitrogen, followed by lyophilization and storage in vacuum-sealed conical centrifuge tubes. The extraction procedure was performed in accordance with the methodology described by Salem and coworkers [32], with modifications for coffee leaf matrices [33]. This involved transferring 30 mg of lyophilized material into microtubes containing 1.5 mL of the extraction solvent mixture, Methyl-tert-butyl-ether: methanol (MTBE: MeOH, 3:1, v:v), along with internal standard L-Alanine-d4 (50 µL/100 mL). The microtubes were vortexed and incubated on an orbital shaker

(40 rpm) for 45 min at 4 °C followed by sonication for 15 min in a water bath at 4 °C. The microtubes were then centrifuged (10,000 × g) for 10 min at 4 °C. From the soluble fractions, aliquots (0.5 mL) were transferred to new microtubes and dried in a speed vacuum for 45 min for metabolomic analyses. The remaining solid pellets, composed of proteins and starch, were washed with 500 µL methanol using a vortex mixer for 1 min. The pellets were then centrifuged at 10,000 × g for 5 min at 4°C. This process was repeated three times following the protocol outlined by Salem et al. (2016). The protein and starch contents were determined in the pellets. Metabolic analysis, as well as protein and starch content measurements, was performed on five biological replicates for each coffee cultivar at both sites, Monte Carmelo and Patrocínio.

2.7.Metabolite profiling

Dried metabolic extracts in sealed, crimped vials were derivatized using 10 µL of a 20 mg mL -1 solution of methoxyamine hydrochloride in pyridine and 10 µL of *N*-trimethylsilyl-*N*-methyl trifluoroacetamide (MSTFA) [34] and injected using a Gerstel MPS2-XL (Gerstel, Muhlheim/Ruhr, Germany) autosampler. A LECO Pegasus BT time-of-flight mass spectrometer (LECO, St. Joseph, MI, USA) hyphenated with an Agilent 8890 gas chromatograph and helium as the carrier gas at 1.0 mL min-1 flow and linear velocity as flow control mode. The front inlet temperature was set to 230 °C, and the samples were run in splitless mode. The capillary column used was an Agilent DB-35MS (30 m × 0.25 mm × 0.25 µm). The temperature program started at 85°C in isothermal mode for 2 min. The isothermal step was followed by a 15°C/min ramp to 360°C. Fatty acid methyl esters (FAME MIX: C8-C30) were used to determine the retention time index (Kind et al. 2009). The ion-source temperature was 250 °C. The recorded mass range was m/z = 50-600 at 30 scans/s. A solvent delay of 3 min was used. The filament bias current was -70V, and the detector voltage ranged from approximately to 1900-2000 V depending on the detector age. The instrument was tuned and validated according to the EPA tune compliance.

Metabolic features were identified using LECO ChromaTOF software in conjunction with the GOLM Metabolome Database. Peak intensities were determined using the TargetSearch package [35] in R software (4.3.1 version) [36]. For normalization, the content of the internal standard L-Alanine-d4 was utilized.

2.8.Leaf protein extraction for proteomic analysis

Protein extraction from the washed pellets was conducted using a modified version of Salem et al. (2020). The pellets were resuspended in 100 µL of protein extraction buffer (30

mM Tris-HCl pH 8.8, 7 M Urea, 2M thiourea, 15 mM DTT, 4% CHAPS), for 10 to 15 minutes. The samples were then sonicated in an ice bath (10 minutes) and incubated for 30 minutes on an orbital shaker (40 rpm) at room temperature. The suspended proteins were centrifuged at 14,000 g for 15 min, and the protein supernatant was transferred to a new tube. This supernatant was used to determine protein concentration using a modified Bradford assay [37].

2.9.Proteomic analysis

Proteomic sample preparation was conducted using the PreOmics iST kit, following the manufacturer's protocol. A total of 50 µg of protein was utilized in the analysis. Liquid chromatography-tandem mass spectrometry (LC-MS/MS) was performed using an Eksigent NanoLC-425 system coupled with a TripleTOF 6600+ mass spectrometer [38]. The acquired MS and MS/MS data were processed using Progenesis QI for Proteomics software (version 4, Nonlinear Dynamics, Waters, Newcastle upon Tyne, UK). Protein and peptide identification was performed by searching the *Coffea arabica* database (53,938 sequences) from UniProtKB using Mascot Daemon (version 2.6.0, Matrix Science, London, UK). The search parameters included a peptide mass tolerance of 20 ppm, a fragment mass tolerance of 0.03 Da, and a maximum allowance of two missed cleavages. Protein identifications were considered valid only if they met the following criteria: a Mascot-calculated significance threshold corresponding to $p < 0.05$, a minimum of two peptide sequences per protein, and at least one unique peptide sequence per protein.

2.10. Protein extraction and invertase activity assay

The invertase enzymes were extracted and analyzed using a modified protocol [39]. Briefly, 150 mg of leaf material per sample were extracted with 1.5 mL of 100 mM KH_2PO_4 buffer pH 7.4, containing 7.5 mM $MgCl_2$, 20 µM $MnCl_2$, 10% glycerol, 5 mM DTT, 5 mM ascorbic acid, 1.5% sodium sulphite and 50 mg of PVPP. The samples were homogenized for 40 minutes at 4°C using a head-to-head shaker. Subsequently, the homogenates were centrifuged at 12,000 rpm for 4°C until a clear supernatant was obtained. An aliquot of 0.3 mL of the supernatant (crude extract) was transferred to new tubes and stored at -80°C. The remaining supernatant was subjected to dialysis, and the resulting extracts were used to analyze vacuolar and cytosolic invertases. The pellet was washed three times with saline buffer (40 mM Tris-HCl, pH 7.6, 3 mM $MgCl_2$, 1 mM EDTA, and 1 M NaCl), followed by centrifugation at 12,000 rpm for 10 minutes at 4°C, and the supernatant was discarded after each wash. Following the final wash, 1 mL of saline buffer was added to the pellet.

The samples were homogenized overnight at 4°C using a head-to-head shaker. On the following day, the samples were centrifuged at 12,000 rpm for 10 minutes at 4°C until the supernatant was completely clear. The supernatant was then transferred to new tubes and subjected to dialysis to obtain the cell wall extract. For the invertase activity assays, aliquots of protein extracts (up to 20 µL) were used. Vacuolar and cell wall invertase activities were assessed using 15 µL of buffer at pH 4.5 (454 mM Na_2HPO_4 and 273 mM citric acid), while cytoplasmic invertase activity was evaluated with 15 µL of buffer at pH 6.8 (772 mM Na_2HPO_4 and 114 mM citric acid). In all assays, 15 µL of sucrose solution (0.1 M) was added, and the reaction volume was adjusted to 50 µL with distilled water. Glucose content was quantified using the ®R-BIOPHARM Starch Kit (Cat. No. 10207748035, Boehringer Mannheim/R-Biopharm), adapted for microtiter plates with absorbance measured at 490 nm [40]. Calibration curves were generated using glucose standards ranging from 0 to 1.2 g/L. The results were expressed as micromoles of glucose equivalents per gram of fresh weight (FW).

2.11. Statistical analysis

Comparative visualization of key variables across cultivars and irrigation was conducted using R (version 4.3.1) [36] with the following packages: *circlize* [41], *ComplexHeatmap* [42, 43], *dplyr* [44], *ggplot2* [45], and *pacman* [46]. Statistical comparisons of leaf water potential, morphoanatomical, and invertases activity of each cultivar across the irrigation conditions were performed using the Student's *t*-test, with Welch's correction applied in cases where homogeneity of variances was not observed. For variables that did not meet the assumptions required for the *t*-test, the nonparametric Wilcoxon test was employed. Pearson's correlation analysis was conducted to examine the relationships between significant morphoanatomical characteristics. A significance level of $\alpha = 0.05$ was adopted for all statistical analyses.

All subsequent analyses were conducted using the R statistical environment and the MetaboAnalyst 6.0 platform [47]. Metabolomic and proteomic datasets were initially examined to assess potential differences among the ten experimental groups, defined by cultivar and irrigation condition. Metabolic features and proteins with constant or single values across samples were excluded from the statistical analysis. Variables with more than 50% missing values were also removed. For the remaining data, missing values were imputed using 1/5 of the minimum positive value observed for each variable. Data were normalized by log-transforming (base 10) and scaling with mean centering. Group-wise differences among the ten conditions were evaluated using one-way analysis of variance (ANOVA), followed by Tukey's Honestly Significant Difference (HSD) post hoc test, with statistical significance set at $p < 0.05$.

Significant metabolomic and proteomic variables were subsequently analyzed separately for each cultivar.

To identify the most discriminative metabolic features and proteins between the two irrigation conditions within each cultivar, volcano plots were generated using a fold-change threshold of 1.75 and a significance cut-off of $p < 0.05$. Based on the results of these analyses, the associated metabolic classes, biological processes, and subcellular localizations were determined. All metabolic class identifications were manually validated using the GOLM metabolome database [48], the Kyoto Encyclopedia of Genes and Genomes (KEGG) [49], and the Human Metabolome Database (HMDB) [50]. The functional annotation of identified proteins was based on the MapMan "Bin" ontology, utilizing the Mercator Automated Sequence Annotation Pipeline (Mercator4 v1.0) [51]. Default parameters were used for the annotation process. The localization of the significant proteins in the cell was investigated using TargetP 2.0 [52]. A principal component analysis (PCA), in conjunction with permutational multivariate analysis of variance (PERMANOVA), was conducted using all significant morphoanatomical and biochemical characteristics to evaluate the overall behavior of the cultivars concerning the irrigation.

3. Results

The irrigated and rainfed plants of the Catiguá (MG2 and MG3) and Sarchimor cultivars exhibited significant differences in leaf water potential, with lower values observed in the irrigated plants (Table 2). In contrast, no significant variation in water potential was observed between irrigation conditions for the Catuaí and Paraíso 2 cultivars. Regarding morphoanatomical traits, of the sixteen structural characteristics evaluated, only seven exhibited significant differences in at least one cultivar when comparing irrigated and rainfed conditions (Table 3). The most pronounced differences were observed in specific leaf area, which was consistently higher in irrigated plants across all cultivars, and in the thickness of the palisade parenchyma, which was higher in rainfed plants, except in Catuaí, in which no significant difference was found. Notably, Catuaí displayed considerable variation in the thickness of both the upper and lower cuticles, as well as in stomatal pore area. On the other hand, Catiguá cultivars presented differences in the equatorial diameter of the stomata. Additionally, like Catuaí, rainfed Catiguá MG3 plants exhibited increased cuticle thickness. A significant increase in xylem area was observed exclusively in rainfed Sarchimor plants.

Statistical analyses revealed significant differences in metabolomic and proteomic profiles between irrigated and non-irrigated plants across the various cultivars. Among the ten groups evaluated—comprising combinations of irrigation conditions and cultivars—a total of 268 metabolic features (MFs), out of 431 valid MFs, as well as 275 proteins from identified protein prior dataset, were identified as differentially expressed.

Subsequently, these metabolic and proteomic characteristics were analyzed within each cultivar to assess differences between irrigated and rainfed conditions. The volcano plot analyses revealed that the cultivars Catiguá MG2, Catuaí, and Paraíso 2 exhibited a higher number of significant characteristics compared to Catiguá MG3 and Sarchimor (Fig. 1) (Supplementary Tables 1 to 10). Across all cultivars, proteomic profiles demonstrated a higher degree of differential expression than metabolomic profiles. Furthermore, the Catiguá cultivars displayed a higher number of upregulated features under irrigated conditions, whereas Catuaí, Paraíso 2, and Sarchimor exhibited more upregulated features in rainfed plants (Fig. 1).

In the differentially expressed metabolome, the predominant metabolite classes included lipids, sugars, and amino acids; however, the distribution of these metabolic classes exhibited variation across the cultivars (Fig. 2). Analysis of the differentially expressed proteome revealed that protein metabolism—encompassing biosynthesis, modifications, and homeostasis—and photosynthesis-related pathways constituted the primary biological processes across all cultivars examined (Fig. 3). Regarding the enzyme classification, oxidoreductases were the most prominently represented group in nearly all cultivars, except Sarchimor, in which isomerases were more prevalent (Fig. 4). Subcellular localization analysis indicated that most differentially expressed proteins were associated with chloroplasts or possessed signal peptides targeting them to the extracellular space, consistently across all cultivars (Fig. 5).

Regarding invertase activity, rainfed plants of all cultivars exhibited elevated vacuolar invertase activity (Fig. 6). Additionally, rainfed Catuaí and Paraíso 2 plants demonstrated higher cytoplasmic and cell wall invertase activities. In contrast, increased cell wall invertase activity was observed only in the irrigated plants of Catiguá MG2 and Sarchimor.

A principal component analysis (PCA) combined with permutational multivariate analysis of variance (PERMANOVA) was performed using invertase activity and all significant morphoanatomical, metabolomic, proteomic characteristics. Principal component analysis (PCA) revealed a distinct separation between Catiguá MG2, Catuaí, and Paraíso cultivars across the two irrigation conditions. The first principal component, accounting for 31.8% of the variance, was the primary driver of this separation. Additionally, Sarchimor samples from

irrigated and rainfed plants were differentiated by the second principal component, which explained 8.4% of the variance (Table 4, Fig. 7).

4. Discussion

In the present study, the leaf water potential observed predominantly in irrigated plants of the Catiguá (MG2 and MG3) and Sarchimor. A leaf water potential -1.5 MPa is indicative of moderate water deficit in coffee plants [24]. Overall, the water potential values recorded across both irrigated and rainfed conditions for the evaluated cultivars were consistent with those typically observed in coffee plants grown in the Cerrado region of Minas Gerais, where water potential generally approaches -1.1 MPa during the dry season, particularly in August [23]. Notably, no statistically significant differences in water potential were detected between irrigated and rainfed treatments for the Catuaí and Paraíso 2 cultivars.

A plausible explanation for the observed leaf water potential in irrigated plants is that consistent irrigation promotes root development in upper soil layers where water is more readily available [53]. As a result, during periods when evapotranspiration exceeds surface soil moisture availability [21], these plants may experience higher water stress due to their limited capacity to access deeper soil moisture reserves—an adaptive advantage typically observed in rainfed plants with deeper root systems [54, 55].

Given the higher annual available water capacity at the experimental site designated for irrigated plants [21], a greater specific leaf area (SLA) was anticipated during leaf growth and expansion phases. On the other hand, nearly all rainfed plants exhibited greater palisade parenchyma thickness. This inverse relationship between SLA and palisade parenchyma thickness is consistent with the findings of Gotoh and coworkers [56], who reported that a more columnar palisade parenchyma structure is typically associated with reduced specific leaf area.

Besides that, the observed thickening of the cuticle in rainfed plants of the Catuaí and Catiguá MG3 cultivars, as well as in irrigated Catuaí plants, may represent an adaptive strategy employed by this cultivar to maintain leaf water status under variable water availability [57]. Following the establishment of leaf structural features, cultivars may further acclimate through metabolic adjustments in response to periods of reduced water availability. Notably, the upregulation of Glycine-rich proteins (*GRP*) in irrigated plants across all cultivars suggests a role for these proteins in modulating stomatal opening, as GRP gene expression is known to be induced by abscisic acid (ABA) [58]. Supporting this, the Abscisic Stress-Ripening proteins (*ASR*), which are associated with responses to water deficit [59], were also found to be upregulated in irrigated Catuaí plants. Furthermore, increased ABA accumulation was detected

in irrigated plants of the Catiguá MG2 cultivar, reinforcing the involvement of ABA-mediated signaling in the regulation of water status and stress response in these cultivars.

An evaluation of key metabolomic and proteomic traits reveals a pronounced tendency for irrigated plants to exhibit higher responses to drought stress compared to rainfed plants, potentially related to their lower water potential at the time of sampling. The Catiguá MG2 and Paraíso 2 cultivars, which exhibited more pronounced metabolic differences relative to the other cultivars, also demonstrated accumulation of metabolites with recognized osmoprotective functions [60]. In rainfed plants, there was an increased accumulation of the polyamine spermine, along with amino acids such as β-alanine and proline. Conversely, irrigated Catiguá MG2 plants exhibited elevated levels of putrescine and the sugar alcohol myo-inositol, suggesting distinct metabolic pathways may be activated in response to leaf water status. In addition to osmotic adjustment, alterations in cell wall rigidity have been identified as responses to mild drought stress [61]. The observed upregulation of the trichome birefringence-like (*TBL*) protein in irrigated Catuaí and Paraíso 2 cultivars suggests structural modifications of the cell wall, which may contribute to enhanced drought tolerance [62].

Moreover, invertase activity exhibited significant variation across different cultivars and irrigation conditions (Figure 5). The increased invertase activity is frequently associated with plant stress responses [63, 64]. Specifically, elevated activity of acid invertases, including vacuolar and cell wall invertases, has been linked to drought tolerance mechanisms in *Coffea canephora* Pierre var. *kouillou* clones [26]. Increased activity of acid and neutral invertases, alongside the accumulation of reducing sugars, in *C. arabica* cv. SIRIEMA seedlings subjected to water deficit conditions was reported before [27]. Furthermore, enhanced expression of cell wall invertase is associated with the biosynthesis of specialized metabolites, such as phenylpropanoids and alkaloids [65].

Although not as prominently represented as other biological processes, secondary metabolism—particularly pathways associated with phenolic compound biosynthesis—had a significant role in irrigated Catiguá MG2 plants, plants with lower leaf water potential. In this cultivar, all differentially expressed phenolic compounds, as well as the Polyphenol Oxidase I protein, were upregulated. The induction of phenolic biosynthesis is known to be triggered by environmental stimuli [66], and these compounds contribute to the antioxidant defense system by mitigating intracellular reactive oxygen species (ROS) [67]. Furthermore, peroxidase enzymes—members of the oxidoreductase class, which represented the most abundant enzyme group across the cultivars—are essential components of the antioxidant machinery due to their efficiency in ROS scavenging [68, 69]. Supporting this observation, redox-regulatory proteins

[70, 71], including Thioredoxin-like 1-2 and Glutathione Peroxidase, were also significantly upregulated in irrigated Catiguá MG2 plants. Conversely, proteins associated with redox homeostasis in the Catuaí and Paraíso 2 cultivars were predominantly upregulated in rainfed plants.

Protein metabolism was the most representative of the biological processes. Post-translational protein modifications are recognized as rapid and early responses of plants to environmental fluctuations [72]. Additionally, proteins containing signal peptides—key constituents of the plant secretome—play essential roles in mediating responses to environmental stress [73]. It is also important to consider that, beyond water availability, the phenological stage of the coffee plants differed among cultivars, with the biennial cycle representing a significant factor influencing physiological responses [74, 75]. In particular, coffee plants in a post-harvest dormant phase, such as those analyzed in this study, may develop flowering (or pre-flowering) in response to abrupt water availability, typically following a drought period [76]. Notably, in the week preceding leaf sample collection, rainfall events were recorded (data not shown) at the Patrocínio experimental site, where rainfed plants were cultivated. This transient increase in soil moisture likely influenced plant water status, potentially activating flowering-related signaling pathways, especially in rainfed plants of Catuaí and Paraíso plants. This hypothesis is supported by the observed upregulation of FT-interacting protein 1-like, a key protein involved in the transport of the florigen protein FT, which is essential for the initiation of flowering [77].

Conclusion

This study highlights the complex interplay between water availability, morphoanatomy, physiological responses, and molecular regulation in *Coffea arabica* cultivars grown under contrasting irrigation regimes in the Cerrado region of Minas Gerais. Despite irrigation, no substantial anatomical differences were observed among the cultivars, with the most notable variations limited to palisade parenchyma. Water availability in the months preceding leaf sampling likely influenced the specific leaf area (SLA) of irrigated plants. However, irrigated plants exhibited lower leaf water potential. In parallel, distinct responses were identified in the metabolomic, proteomic, and enzymatic (invertase) profiles of the cultivars. Irrigated and rainfed plants demonstrated the accumulation of metabolites associated with osmotic adjustment and the upregulation of proteins involved in stress response pathways. Among the evaluated cultivars, Catiguá MG2 and Paraíso 2 exhibited the most pronounced differences in metabolomic and proteomic profiles. Notably, in Catiguá MG2, the

activation of stress-related metabolic pathways was more evident in irrigated plants, which also exhibited lower water potential. In contrast, Paraíso 2 plants, regardless of irrigation treatment, appeared to employ acclimatization mechanisms to mitigate the effects of reduced water availability.Collectively, these findings underscore the complex and cultivar-specific nature of coffee plant acclimation to water availability, wherein both structural and metabolic plasticity play critical roles. The results offer valuable insights into the physiological and molecular mechanisms underlying responses to irrigation conditions in Cerrado Mineiro, with important implications for breeding and agronomic strategies to improve coffee cultivars resilience in the field.

References

[1] V. Velikova, C. Arena, L.G. Izzo, T. Tsonev, D. Koleva, M. Tattini, O. Roeva, A.D. Maio, F. Loreto, Functional and Structural Leaf Plasticity Determine Photosynthetic Performances during Drought Stress and Recovery in Two *Platanus orientalis* Populations from Contrasting Habitats, International Journal of Molecular Sciences (2020) 3912. doi:10.3390/ijms21113912.

[2] G. C. Stotz, C. Salgado-Luarte, V. M. Escobedo, F. Valladares, E. Gianoli, Phenotypic plasticity and the leaf economics spectrum: Plasticity is positively associated with specific leaf area, Oikos (2022) e09342. doi:10.1111/oik.09342.

[3] G. R. Cramer, K. Urano, S. Delrot, M. Pezzotti, K. Shinozaki, Effects of abiotic stress on plants: A systems biology perspective, BMC Plant Biology (2011) 163. doi:10.1186/1471-2229-11-163.

[4] R. A. Laitinen, Importance of phenotypic plasticity in crop resilience, Journal of Experimental Botany (2024) 670-673. doi:10.1093/jxb/erad465.

[5] International Coffee Organization, Trade Statistic Tables, Production, Coffee Trade Status. [Online]. http://www.ico.org/ trade_statistics.asp, 2024 (acessed 20 June 2024).

[6] F.F. De Oliveira, J. P. Tomaz, B.S.R Da Silva, T. B. Dos Santos, S. T. Ivamoto-Suzuki, M. B. Dos Santos Scholz, L. F. P. Pereira, *Coffea arabica* L. Genes from isoprenoid metabolic pathways are more expressed in full sun cultivation systems than in agroforestry systems, Plant Gene (2021) 100287. doi:10.1016/j.plgene.2021.100287.

[7] L. S. Coelho, G. A.T. Tassone, G. R. Carvalho, V. A. Silva, M. T. R. Viana, F. A. C Pereira, D. H. S. Nadaleti, H. R. O. Silveira, C. E. Botelho, Morphological, physiological, and agronomic traits of crossings of 'Icatu' x 'Catimor' coffee tree subjected to water déficit, Pesquisa Agropecuária Brasileira (2022) e02788. doi:10.1590/S1678-3921.pab2022.v57.02788.

[8] I. Marques, A. P. Rodrigues, D. Gouveia, F. C. Lidon, S. Martins, M. C. Semedo, J. Gaillard, I.P. Pais, J. N. Semedo, P. Scotti-Campos, F.H. Reboredo, F.L. Partelli, F.M. DaMatta, J. Armengaud, A. I. Ribeiro-Barros, J. C. Ramalho, 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. Journal of Plant Physiology (2022) 153788. doi:10.1016/j.jplph.2022.153788.

[9] V. A. Silva, J. C. Abrahão, A. M. Reis, M.D. Santos, A.A. Pereira, C. E. Botelho, G.R. Carvalho, E. M. Castro, J. P. Barbosa, G. P. Botega, A. C. Oliveira, Strategy for Selection of Drought-Tolerant Arabica Coffee Genotypes in Brazil.Agronomy 12 (2022) 2167. doi:10.3390/agronomy12092167.

[10] H. Chekol, B. Warkineh, T. Shimber, G. B. Dąbrowska, A. Degu, Drought Stress Responses in Arabica Coffee Genotypes: Physiological and Metabolic Insights, Plants 13 (2024) 828. doi:10.3390/plants13060828.

[11] I. Marques, I. Fernandes, O.S. Paulo, D. Batista, F. C. Lidon, A.P. Rodrigues, F. L. Partelli, F. M. DaMatta, A. I. Ribeiro-Barros, J. C. Ramalho JC. Transcriptomic Analyses Reveal That *Coffea arabica* and *Coffea canephora* Have More Complex Responses under Combined Heat and Drought than under Individual Stressors. International Journal ofMolecular Sciences 25 (2024) 7995. doi:10.3390/ijms25147995.

[12] A. R. Reis, J. L. Favarin, P. L. Gratão, F. R. Capaldi, R. A. Azevedo, Antioxidant metabolism in coffee (*Coffea arabica* L.) plants in response to nitrogen supply, Theor. Exp. Plant Physiol. 27 (2015) 203–213. doi:10.1007/s40626-015-0045-3.

[13] D.D.S. Soares, E.A. da Silva, M.A.D.F. Carvalho, F.A.C. Pereira, R.J. Guimarães, Leaf anatomy, physiology and vegetative growth of fertigated *Coffea arabica* L. trees after exposure to pruning, Coffee Science (2021) e161962.doi:10.25186/.v16i.1962.

[14] J.P. Pérez-Molina, E.A.D.T. Picoli, L.A. Oliveira, B.T. Silva, G.A. De Souza, J.L.D.S. Rufino, A.A. Pereira, M.D.F. Ribeiro, G.L. Malvicini, L. Turello, S.C. D´ Alessandro, N.S. Sakiyama, W.P.M. Ferreira, Treasured exceptions: Association of morphoanatomical leaf traits with cup quality of *Coffea arabica* L. Cv. “Catuaí”, Food Research International 141 (2021) 110118. doi:10.1016/j.foodres.2021.110118.

[15] R. Bollen, O. Rojo-Poveda, L. Verleysen, R. Ndezu, E.A. Tshimi, H. Mavar, T. Ruttink, O. Honnay, P. Stoffelen, C. Stévigny, F. Souard, C. Delporte, Metabolite profiles of green leaves and coffee beans as predictors of coffee sensory quality in Robusta (*Coffea canephora*) germplasm from the Democratic Republic of Congo, Applied Food Research 4 (2024) 100560. doi:10.1016/j.afres.2024.100560.

[16] K.G.D. Livramento, F. M. Borém, A. C. José, A. V. Santos, D. E. D. Livramento, J. D. Alves, L. V. Paiva, Proteomic analysis of coffee grains exposed to different drying process, Food Chem. 221 (2017) 1874-1882. doi: 10.1016/j.foodchem.2016.10.069.

[17] L. Guerra-Guimarães, R. Tenente, C. Pinheiro, I. Chaves, M. C. Silva, F. M. Cardoso, S. Planchon, D. R. Barros, J. Renaut, C. P. Ricardo, Proteomic analysis of apoplastic fluid of *Coffea arabica* leaves highlights novel biomarkers for resistance against *Hemileia vastatrix*. Front Plant Sci. 6 (2015) 478. doi: 10.3389/fpls.2015.00478.

[18] K. F. Possa, J. A. Silva, M. L. Resende, R. Tenente, C. Pinheiro, I. Chaves, S. Planchon, A. C. Monteiro, J. Renaut, M. A. Carvalho, C. P. Ricardo, Primary Metabolism Is Distinctly Modulated by Plant Resistance Inducers in *Coffea arabica* Leaves Infected by Hemileia vastatrix. Frontiers in Plant Science 11 (2020) 481646. https://doi.org/10.3389/fpls.2020.00309.

[19] Z. Li, B. Zhou, T. Zheng, C. Zhao, X. Shen, X. Wang, M. Qiu, J. Fan, Integrating Metabolomics and Proteomics Technologies Provides Insights into the Flavor Precursor Changes at Different Maturity Stages of Arabica Coffee Cherries, Foods 12 (2023) 1432. doi: 10.3390/foods12071432.

[20] M. F. V. Pereira, Globalização, especialização territorial e divisão do trabalho: Patrocínio e o café do Cerrado Mineiro, Cuadernos de Geografía: Revista Colombiana de Geografía 23 (2014). doi:10.15446/rcdg.v23n2.37333.

[21] G. B. Voltolini, G. R. Carvalho, V. T. Andrade, A. D. Ferreira, F. V. Raposo, J. P. Carvalho, D. J. Vilela, C. A. Da Silva, J. D. Costa, G. B. Abreu, J. C. Abrahão, C. E. Botelho, D. H. Nadaleti, M. R. Malta, V. A. Silva, S. M. Salgado, R. S. De Faria, A. C. De Oliveira, A. A. Pereira, Agronomic Performance of Irrigated and Rainfed Arabica Coffee Cultivars in the Cerrado Mineiro Region. Agronomy 15 (2025) 222. https://doi.org/10.3390/agronomy15010222.

[22] A.F. Guerra, J.D.F. Santos, L.T. Ferreira, O.C. Rocha, (2021) Cafés do Brasil: Pesquisa, sustentabilidade e inovação. in: S. F. P. Telhado, G. Capdeville (Eds.), Tecnologias poupa-terra 2021. Embrapa, Brasília (DF), 2021, pp. 63-75.

[23] C.S. Dos Santos, A.F. De Freitas, G.H.B. Da Silva, M.A.D.F. Carvalho, M.D.O. Santos, G.R. Carvalho, V.A. Silva, Adaptations to the drought season and impacts on the yield of 'Híbrido de Timor' coffee tree in the Minas Gerais State Cerrado (Brazilian Savanna), Pesq. Agropec. Trop. 52 (2022) e72448. doi:10.1590/1983-40632022v5272448.

[24] C.S. Dos Santos, A. F. De Freitas, G.H.B. da Silva, J.P. Pennacchi, M.A. Figueiredo de Carvalho, M.D.O. Santos, T.S.J. Moraes, J.C.R. Abrahão, A.A. Pereira, G.R.C. Carvalho, C.E.

Botelho, V.A. Silva, Phenotypic Plasticity Index as aStrategy for Selecting Water-Stress-Adapted Coffee Genotypes, Plants 12 (2023) 4029. doi:10.3390/plants12234029.

[25] J.N. Semedo, A.P. Rodrigues, F.C. Lidon, I.P. Pais, I. Marques, D. Gouveia, J. Armengaud, M.J. Silva, S. Martins, M.C. Semedo, D. Dubberstein, F. L. Partelli, F. H. Reboredo, P. Scotti-Campos, A. I. Ribeiro-Barros, F. M. DaMatta, J. C. Ramalho, Intrinsic non-stomatal resilience to drought of the photosynthetic apparatus in *Coffea* spp. Is strengthened by elevated air [CO2], Tree Physiology 41(2021) 708-727. doi:10.1093/treephys/tpaa158.

[26] S. C. Praxedes, F. M. DaMatta, M. E. Loureiro, M. A. G. Ferrão, A. T. Cordeiro, Effects of long-term soil drought on photosynthesis and carbohydrate metabolism in mature robusta coffee (*Coffea canephora* Pierre var. *Kouillou*) leaves, Environmental and Experimental Botany, 56 (2006) 263-273. doi:10.1016/j.envexpbot.2005.02.008.

[27] C.N.F. Brum, E.F. Melo, L.O.B. Barquero, J.D. Alves, A. Chalfun-Júnior, MODIFICATIONS IN THE METABOLISM OF CARBOHYDRATES IN (*Coffea arabica* L. cv. SIRIEMA) SEEDLINGS UNDERDROUGHT CONDITIONS, Coffee Science 8 (2013)140-147.

[28] P.F. Scholander, H.T. Hammel, E.A. Hemmingsenm, E.D. Bradstreet, HYDROSTATIC PRESSURE AND OSMOTIC POTENTIAL IN LEAVES OF MANGROVES AND SOME OTHER PLANTS, Proc Natl Acad Sci. 52 (2006) 119-125. doi:10.1073/pnas.52.1.119.

[29] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji – an Open platformfor biological image analysis, Nature Methods 9 (2012) 676–82. doi:10.1038/nmeth.2019.Fiji.

[30] J.E. Kraus, M. Arduin, Manual básico de métodos em morfologia vegetal. EDUR, Seropédica, 1997.

[31] H.C. Melo, E.M.D Castro, Â.M. Soares, L.A.D. Melo, J.D. Alves, Alterações anatômicas e fisiológicas em *Setaria anceps* Stapf ex Massey e *Paspalum paniculatum* L. sob condições de déficit hídrico, Hoehnea 34 (2007) 145-153.

[32] M. A. Salem, L. Perez de Souza, A. Serag, A. R. Fernie, M .A. Farag, S. M. Ezzat, S. Alseekh. Metabolomics in the Context of Plant Natural Products Research: From Sample Preparation to Metabolite Analysis, Metabolites 10 (2020) 37.doi:10.3390/metabo10010037.

[33] J. C. Carréra, L. Guerra-Guimarães, J. C. D'Auria, L. J. Sartori, C. Pinheiro, V. A. Silva, M. L. Volpato, G. R. Carvalho, F. A. Mori, NON TARGETED METABOLOMIC ANALYSIS OF FIELD-GROWN *Coffea arabica* CULTIVARS REVEALS DISTINCT LEAF

METABOLIC SIGNATURES. Theoretical and Experimental Plant Physiology 2025; (In Press).

[34] A. Erban, N. Schauer, A.R. Fernie, J. Kopka, Nonsupervised Construction and Application of Mass Spectral and Retention Time Index Libraries From Time-of-Flight Gas Chromatography-Mass Spectrometry Metabolite Profiles. In: W. Weckwerth (ed) Metabolomics - Methods in Molecular Biology™, 358. Humana Press, 2007, pp. 19-38. doi:10.1007/978-1-59745-244-1_2.

[35] Á. Cuadros-Inostroza, C. Caldana, H. Redestig, M. Kusano, J. Lisec, H. Peña-Cortés, L. Willmitzer, M. A. Hannah, TargetSearch - a Bioconductor package for the efficient preprocessing of GC-MS metabolite profiling data, BMC Bioinformatics, 10 (2009) 1-12. doi:10.1186/1471-2105-10-428.

[36] RCoreTeam, R: a language and environment for statistical computing. version 4.3.1. R Foundation for statistical computing, Vienna, Austria, 2023 https://www.R-project.org/. (acessed 27 October 2023).

[37] L. S. Ramagli LS, Quantifying protein in 2-D PAGE solubilization buffers. Methods Mol Biol. 112 (1999) 99-103. doi:10.1385/1-59259-584-7:99.

[38] X. Xu, S. Legay, K. Sergeant, S. Zorzan, C.C. Leclercq, S. Charton, V. Giarola, X. Liu, D. Challabathula, J. Renaut, J.F. Hausman, D. Bartels, G. Guerriero, Molecular insights into plant desiccation tolerance: Transcriptomics, proteomics and targeted metabolite profiling in *Craterostigma plantagineum*. The Plant Journal 107 (2021) 377-398. doi:10.1111/tpj.15294.

[39] A. Jammer, A. Gasperl, N. Luschin-Ebengreuth, E. Heyneke, H. Chu, E. Cantero-Navarro, D.K. Großkinsky, A.A. Albacete, E. Stabentheiner, J. Franzaring, A. Fangmeier, E. van der Graaff, T. Roitsch, Simple and robust determination of the activity signature of key carbohydrate metabolism enzymes for physiological phenotyping in model and crop plants, J Exp Bot. 66 (2015) 5531-42. doi: 10.1093/jxb/erv228.

[40] C. Pinheiro, G. Emiliani, G. Marino, A.S. Fortunato, M. Haworth, A. De Carlo, M.M. Chaves, F. Loreto, M. Centritto. Metabolic Background, Not Photosynthetic Physiology, Determines Drought and Drought Recovery Responses in C3 and C2 Moricandias. International Journal of Molecular Sciences 24 (2023) 4094.doi:10.3390/ijms24044094.

[41] Z. Gu, L. Gu, R. Eils, M. Schlesner, B. Brors, circlize Implements and enhances circular visualization in R, Bioinformatics 30 (2014) 2811-2. doi:10.1093/bioinformatics/btu393.

[42] Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data, Bioinformatics 32 (2016) 2847-2849. doi:10.1093/bioinformatics/btw313.

[43] Z. Gu, Complex heatmap visualization, IMeta 1 (2022) e43. doi:10.1002/imt2.43.

[44] H. Wickham, R. François, L. Henry, K. Müller, D. Vaughan (2023). dplyr: A Grammar of Data Manipulation. R package version 1.1.4. https://CRAN.R-project.org/package=dplyr. 2023 (acessed 15 november 2024).

[45] H. Wickham, ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York, 2016.

[46] T. W. Rinker, D. Kurkiewicz. pacman: Package Management for R. version 0.5.0. Buffalo, New York. http://github.com/trinker/pacman, 2017 (acessed 15 november 2024).

[47] Z. Pang, Y. Lu, G. Zhou, F. Hui, L. Xu, C. Viau, A.F. Spigelman, P.E. MacDonald, D.S. Wishart, S. Li, J. Xia, MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation. Nucleic Acids Research 52 (W1) (2024), W398–W406.doi:10.1093/nar/gkae253.

[48] I. Hummel, F. Pantin, R. Sulpice, M. Piques, G. Rolland, M. Dauzat, A. Christophe, M. Pervent, M. Bouteillé, M. Stitt, Y. Gibon, B. Muller B, Arabidopsis Plants Acclimate to Water Deficit at Low Cost through Changes of Carbon Usage: An Integrated Perspective Using Growth, Metabolite, Enzyme, and Gene Expression Analysis. Plant Physiology 154 (2010) 357-372. doi:10.1104/pp.110.157008.

[49] M. Kanehisa, M. Furumichi, Y. Sato, M. Kawashima, M. Ishiguro-Watanabe, KEGG for taxonomy-based analysis of pathways and genomes, Nucleic Acids Research 51 (D1) (2023) D587-D592.

[50] D. S. Wishart, A. Guo, E. Oler, F. Wang, A. Anjum, H. Peters, R. Dizon, Z. Sayeeda, S. Tian, B. L. Lee, M. Berjanskii, R. Mah, M. Yamamoto, J. Jovel, C. Torres-Calzada, M. Hiebert-Giesbrecht, V. W. Lui, D. Varshavi, D. Varshavi, D. Allen, D. Arndt, N. Khetarpal, A. Sivakumaran, K. Harford, S. Sanford, K. Yee, X. Cao, Z. Budinski, J. Liigand, L. Zhang, J. Zheng, R. Mandal, N. Karu, M. Dambrova, H.B. Schiöth, R. Greiner, V. Gautam, HMDB 5.0: the Human Metabolome Database for 2022, Nucleic Acids Res. 50 (D1) (2022) D622-D631. doi: 10.1093/nar/gkab1062.

[51] R. Schwacke, G.Y. Ponce-Soto, K. Krause, A.M. Bolger, B. Arsova, A. Hallab, K. Gruden, M. Stitt, M.E. Bolger, B. Usadel, MapMan4: A Refined Protein Classification and Annotation Framework Applicable to Multi-Omics Data Analysis, Mol Plant. 12 (2019) 879-892. doi: 10.1016/j.molp.2019.01.003.

[52] J. J. Almagro Armenteros, M. Salvatore, O. Emanuelsson, O. Winther, G. von Heijne, A. Elofsson, H. Nielsen, Detecting sequence signals in targeting peptides using deep learning, Life Sci Alliance 30 (2019) e201900429. doi: 10.26508/lsa.201900429.

[53] Z. Li, R. Zong, T. Wang, Z. Wang, J. Zhang, Adapting Root Distribution and Improving Water Use Efficiency via Drip Irrigation in a Jujube (Zizyphus jujube Mill.) Orchard after Long-Term Flood Irrigation, Agriculture 11 (2021) 1184. https://doi.org/10.3390/agriculture11121184.

[54] E. Adriano, J.P. Laclau, J.D. Rodrigues, Deep rooting of rainfed and irrigated orange trees in Brazil, Trees 31 (2017) 285–297. https://doi.org/10.1007/s00468-016-1483-5.

[55] A. Odone, O. Popovic, K. Thorup-Kristensen, Deep roots: implications for nitrogen uptake and drought tolerance among winter wheat cultivars. Plant Soil 500 (2024) 13–32. https://doi.org/10.1007/s11104-023-06255-5.

[56] E. Gotoh, N. Suetsugu, T. Higa, T. Matsushita, H. Tsukaya, M. Wada, Palisade cell shape affects the light-induced chloroplast movements and leaf photosynthesis. Scientific Reports 8 (2018) 1-9. doi:10.1038/s41598-018-19896-9.

[57] D. K. Kosma, B. Bourdenx, A. Bernard, E. P. Parsons, M. A. Jenks, The Impact of Water Deficiency on Leaf Cuticle Lipids of Arabidopsis. Plant Physiology 151 (2009) 1918-1929. https://doi.org/10.1104/pp.109.141911.

[58] M. Czolpinska, M. Rurek. Plant Glycine-Rich Proteins in Stress Response: An Emerging, Still Prospective Story. Frontiers in Plant Science, 9 (2018) 327614. https://doi.org/10.3389/fpls.2018.00302.

[59] I. Yacoubi, A. Gadaleta, N. Mathlouthi, K. Hamdi, A. Giancaspro, Abscisic Acid-Stress-Ripening Genes Involved in Plant Response to High Salinity and Water Deficit in Durum and Common Wheat. Frontiers in Plant Science 13 (2022) 789701. https://doi.org/10.3389/fpls.2022.789701.

[60] G. Kaur, S.K. Sanwal, A. Kumar, R.K. Pundir, M. Yadav, N. Sehrawat, Role of osmolytes dynamics in plant metabolism to cope with salinity induced osmotic stress, Discov Agric 2, 59 (2024). https://doi.org/10.1007/s44279-024-00070-x.

[61] H. Al-Yasi, H. Attia, K. Alamer, F. Hassan, E. Ali, S. Elshazly, K.H. Siddique, K. Hessini, Impact of drought on growth, photosynthesis, osmotic adjustment, and cell wall elasticity in Damask rose, Plant Physiology and Biochemistry 150 (2020) 133-139. https://doi.org/10.1016/j.plaphy.2020.02.038.

[62] S. Wang, H. Su, J. Jin, J. Tao, Z. Li, P. Cao, J. Zhang, P. Lu, Genome-wide identification and analysis of the TBL genes reveals NtTBL31 increases drought resistance of tobacco (*Nicotiana tabacum*), Chemical and Biological Technologies in Agriculture 12 (2025) 1-17. https://doi.org/10.1186/s40538-025-00752-8.

[63] A.S. Tauzin, T. Giardina. Sucrose and invertases, a part of the plant defense response to the biotic stresses.Frontiers in Plant Science 5 (2014) 86818. doi:10.3389/fpls.2014.00293.

[64] N. Jiang, P. Yu, W. Fu, G. Li, B. Feng, T. Chen, H. Li, L. Tao, G. Fu, Acid invertase confers heat tolerance in rice plants by maintaining energy homoeostasis of spikelets. Plant, Cell & Environment 43 (2020) 1273-1287.doi:10.1111/pce.13733.

[65] M.J. Nishanth, S.A. Sheshadri, S.S. Rathore, S. Srinidhi, B. Simon, Expression analysis of Cell wall invertase under abiotic stress conditions influencing specialized metabolism in *Catharanthus roseus*, Scientific Reports 8 (2018) 1-15. doi:10.1038/s41598-018-33415-w.

[66] A. Sharma, B. Shahzad, A. Rehman, R. Bhardwaj, M. Landi, B. Zheng, Response of Phenylpropanoid Pathway and the Role of Polyphenols in Plants under Abiotic Stress. Molecules, 24 (2019), 2452. https://doi.org/10.3390/molecules24132452.

[67] P. Cosme, A. B. Rodríguez, J. Espino, M. Garrido, Plant Phenolics: Bioavailability as a Key Determinant of Their Potential Health-Promoting Applications. Antioxidants, 9 (2020), 1263. https://doi.org/10.3390/antiox9121263.

[68] G. Czégény, L. Kőrösi, Å. Strid, É. Hideg, Multiple roles for Vitamin B6 in plant acclimation to UV-B, Scientific Reports 9 (2019) 1-9. https://doi.org/10.1038/s41598-018-38053-w.

[69] V. D. Rajput, R. K. Singh, K. K. Verma, L. Sharma, F. Roberto, M. Meena, V. S. Gour, T. Minkina, S. Sushkova, S. Mandzhieva, Recent Developments in Enzymatic Antioxidant Defence Mechanism in Plants with Special Reference to Abiotic Stress, Biology 10 (2021) 267. https://doi.org/10.3390/biology10040267.

[70] M. Luíza, A. Silva Santos, C. P. Pirovani, F. Micheli, The family of glutathione peroxidase proteins and their role against biotic stress in plants: A systematic review, Frontiers in Plant Science 16 (2025) 1425880. https://doi.org/10.3389/fpls.2025.1425880.

[71] K. Yoshida, T. Hisabori, Current Insights into the Redox Regulation Network in Plant Chloroplasts. Plant and Cell Physiology 64 (2023) 704. https://doi.org/10.1093/pcp/pcad049.

[72] V. Muleya, L. M. Lois, H. Chahtane, L. Thomas, M. Chiapello, C. Marondedze, (De)Activation (Ir)Reversibly or Degradation: Dynamics of Post-Translational Protein Modifications in Plants, Life (Basel) 12 (2022) 324. doi: 10.3390/life12020324.

[73] E. Alexandersson, A. Ashfaq, S. Resjö, E. Andreasson, Plant secretome proteomics. Frontiers in Plant Science, 4 (2013) 39538. https://doi.org/10.3389/fpls.2013.00009.

[74] F. M. DaMatta, C. P. Ronchi, M. Maestri, R. S. Barros, Ecophysiology of coffee growth and production, Braz. J. Plant Physiol. 19 (2007) https://doi.org/10.1590/S1677-04202007000400014.

[75] L. S. Soares, T. T. Rezende, L. A. Beijo, K. S. Franco Júnior, Interaction between climate, flowering and production of dry coffee (*Coffea arabica* L.) in Minas Gerais. Coffee Science 16 (2021) e161786, 2021. doi: 10.25186/.v16i.1786.

[76] G.A.L.Torres, C.N. Campos, M.V. Salomon, A.P. Pantano, J. A. S. Almeida, Coffea arabica L: History, phenology and climatic aptitude of the state of São Paulo, Brazil, Arq. Inst. Biol. 88 (2021). https://doi.org/10.1590/1808-1657000602020.

[77] L. Liu, C. Liu, X. Hou, W. Xi, L. Shen, Z. Tao, Y. Wang, Y. Hao, FTIP1 Is an Essential Regulator Required for Florigen Transport, PLOS Biology 10 (2012) e1001313.https://doi.org/10.1371/journal.pbio.1001313.

Figures

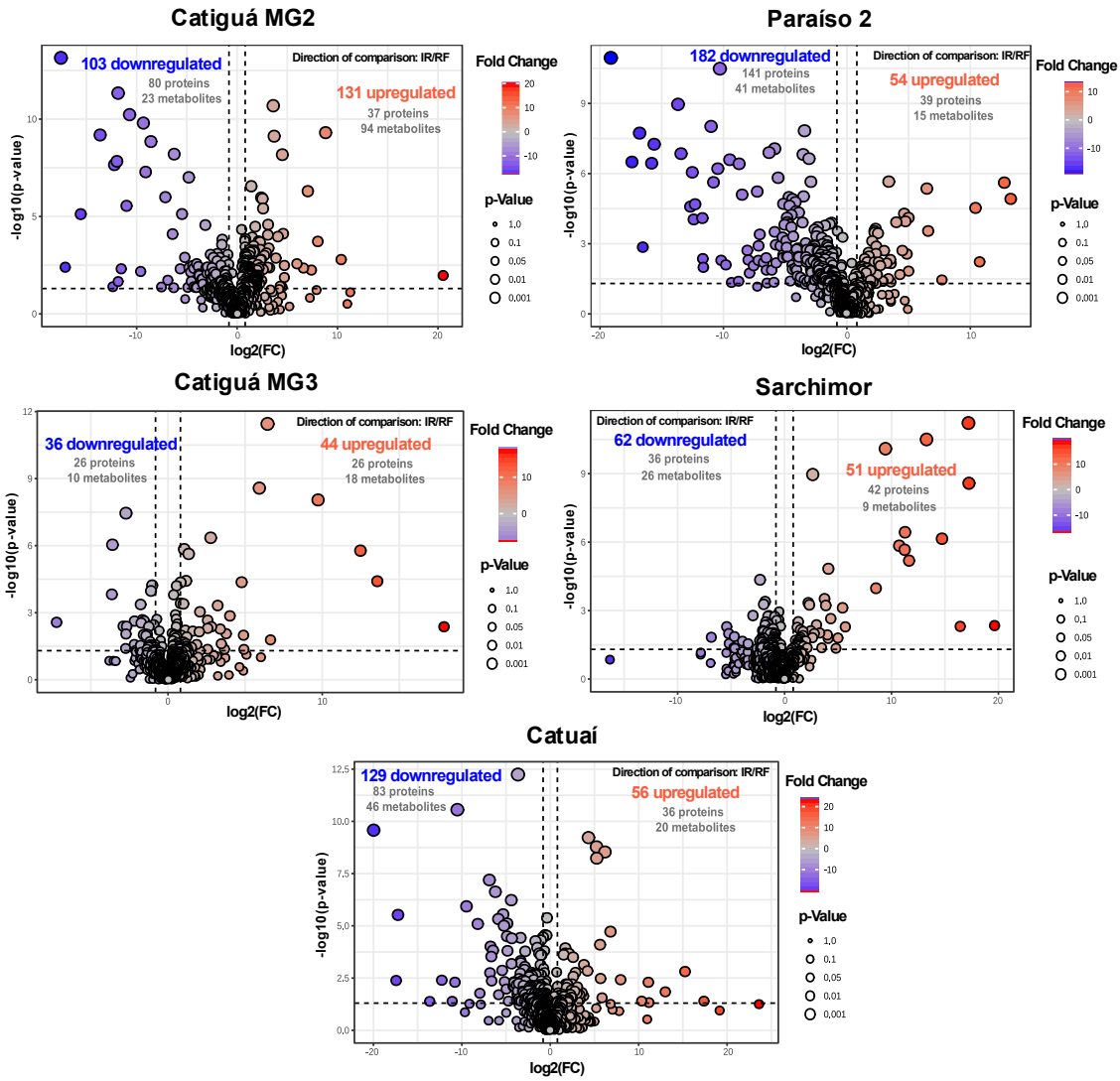


Fig. 1. Volcano plot illustrating the significant características (metabolome and proteome) that differentiate the irrigated (IR) and rainfed (RF) plants of *C. arabica* L. cultivars. The volcano plot was created using Fold Change limits (FC > 1.75) and p-value (p > 0.05).

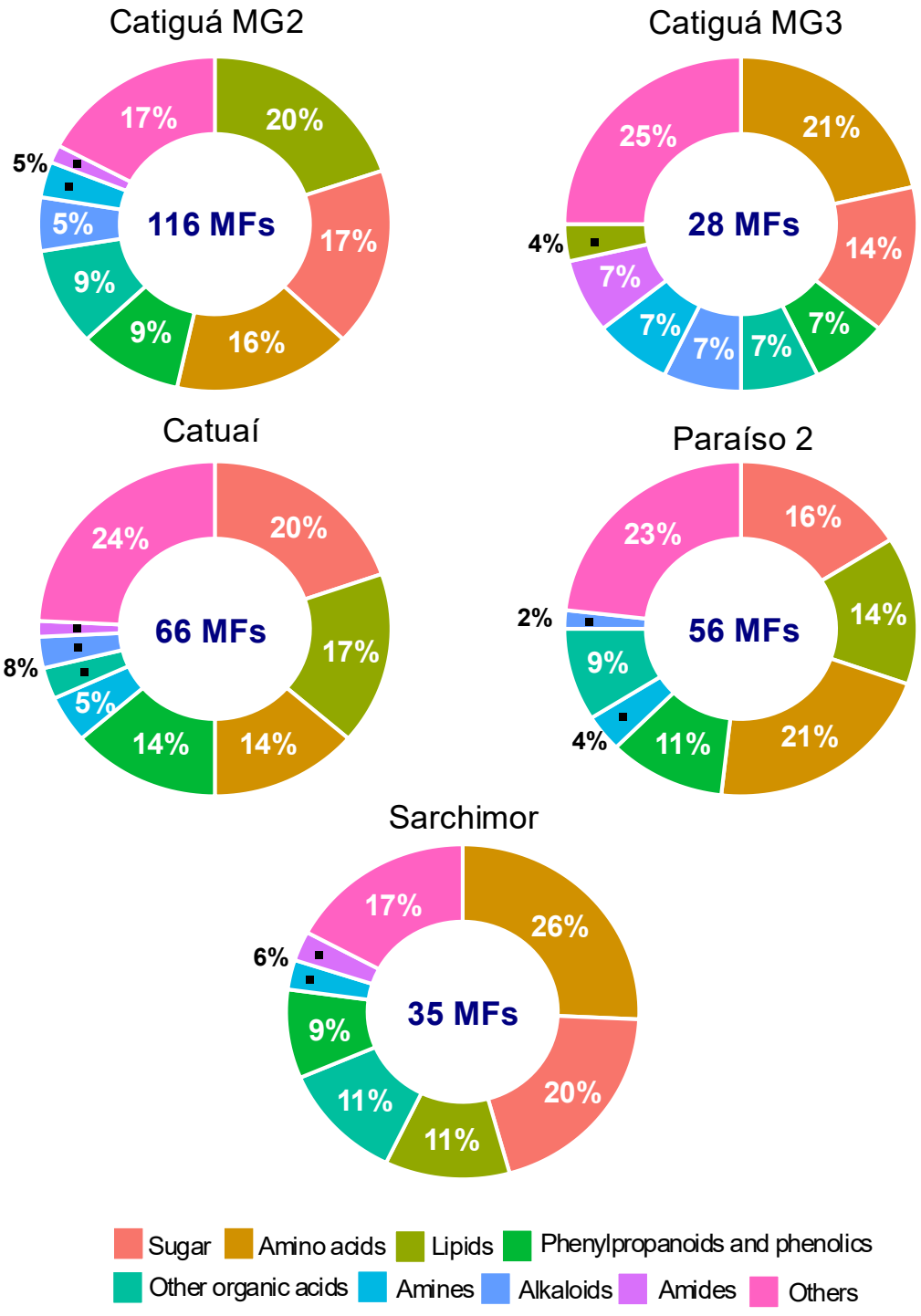


Fig. 2. Metabolic classes of the significant metabolic features (MFs) identified in *C. arabica* cultivars.

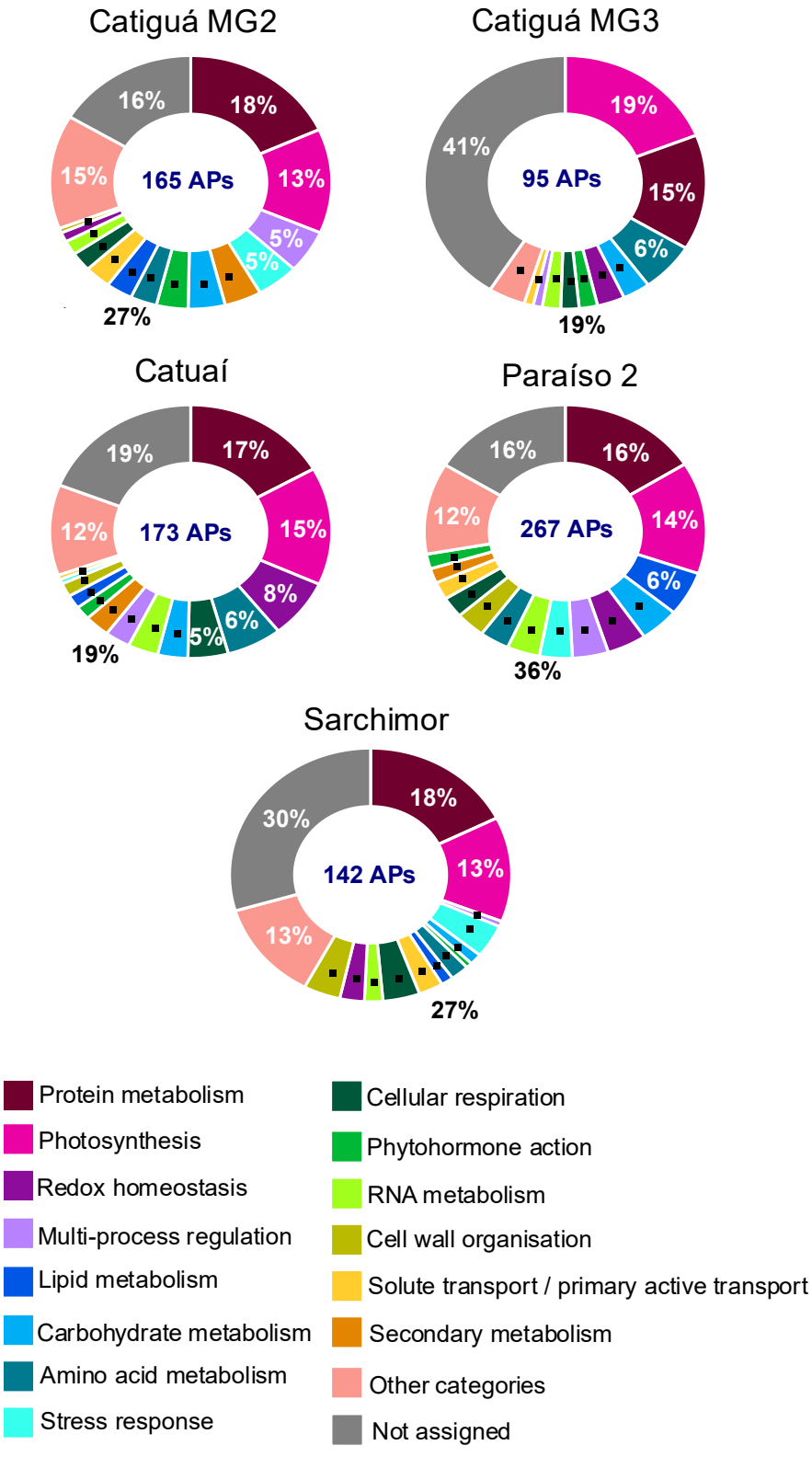


Fig. 3. Functional analysis of the biological processes in which the most significant proteins in *C. arabica* cultivars, based on the MapMan "Bin" and GO ontology.

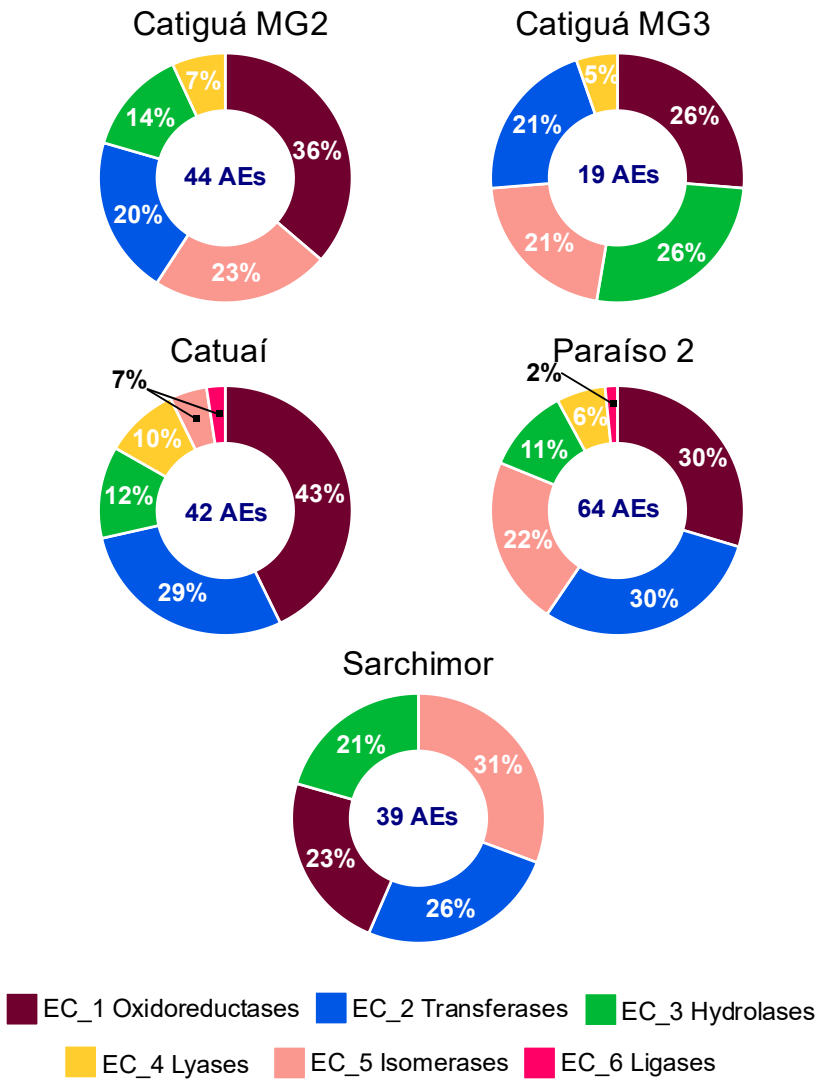


Fig. 4. Functional analysis of the enzyme groups to which the most significant proteins in *C. arabica* cultivars, based on the MapMan "Bin" and GO ontology.

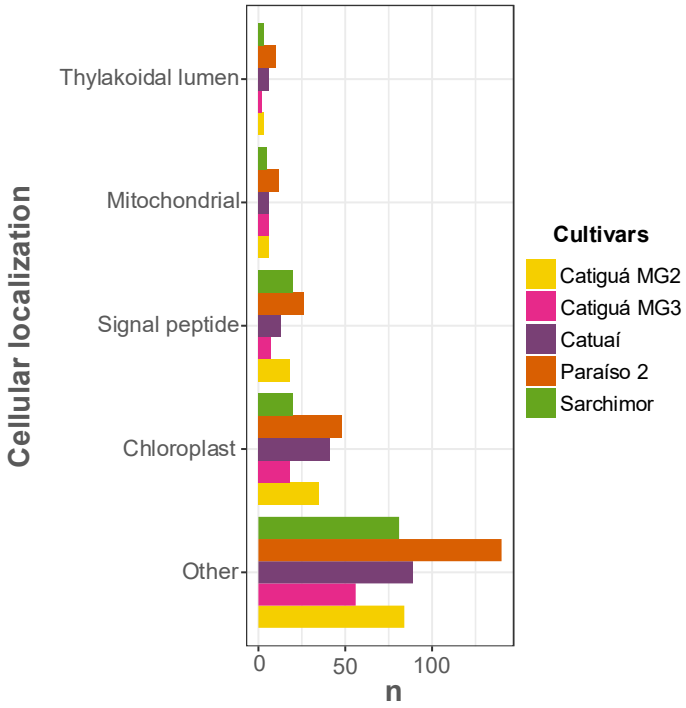


Fig. 5. Subcellular localization of the most significant proteins in distinguishing the cultivars in *C. arabica* cultivars, based on TargetP 2.0.

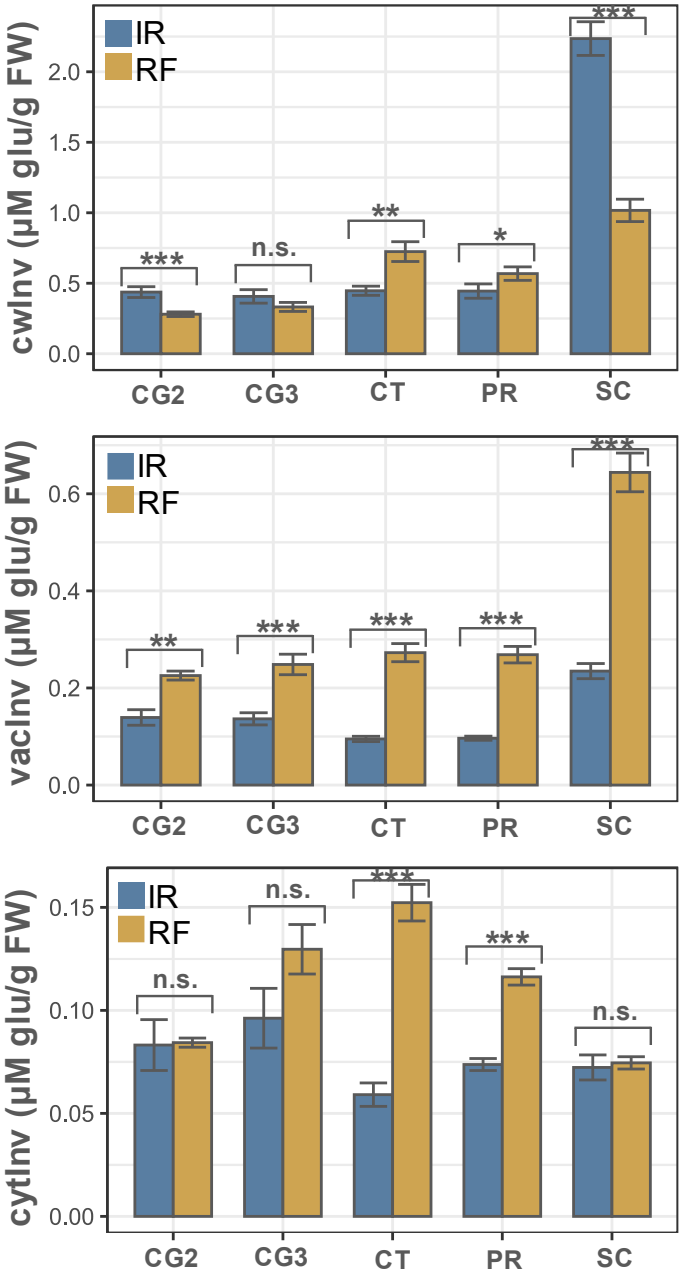


Fig. 6. Invertases activities of *Coffea arabica* cultivars field-grown in the Cerrado Mineiro Region, expressed as micromoles of glucose equivalents per gram of fresh weight (FW); Bars represent standard error. CG2: Catiguá MG2, CG3: Catiguá MG3, cytInv: Cytosolic invertase, CT: Catuaí, cwInv: Cell wall invertase, IR: irrigated plants, PR: Paraíso, RF: Rainfed plants, SC: Sarchimor, vacInv: Vacuolar invertase. (*) p-value < 0.05 and (*) p-value < 0.01, Wilcoxon test.

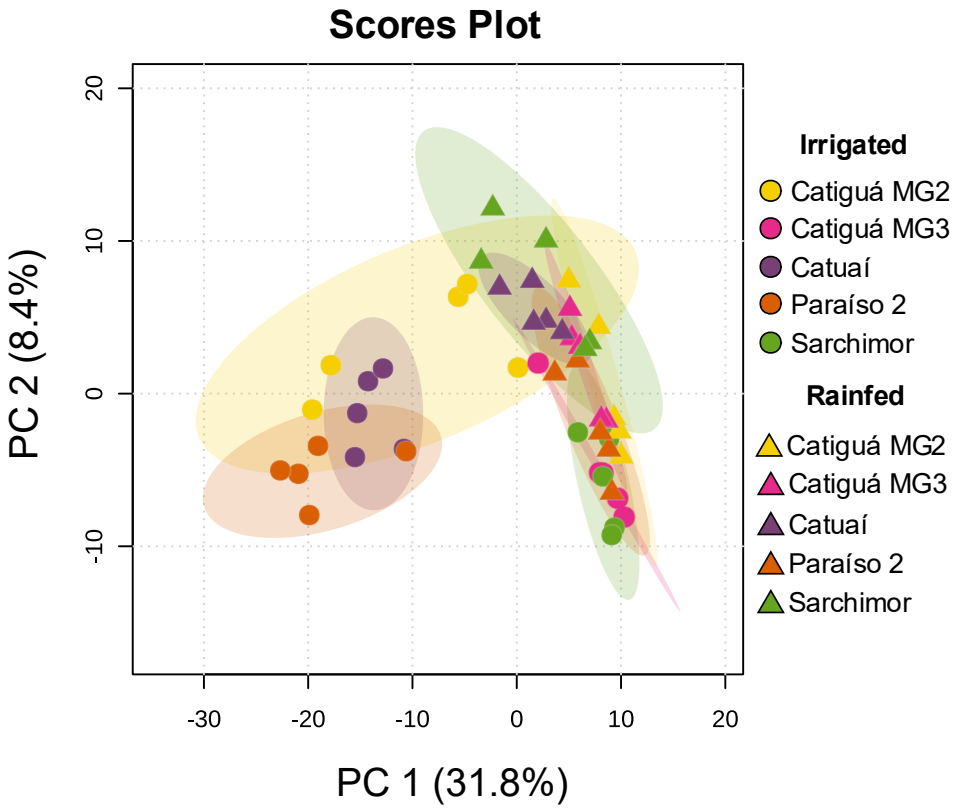


Fig. 7. A principal component analysis (PCA) combined with permutational multivariate analysis of variance (PERMANOVA) of all 554 significant variables (morphoanatomical, metabolomic, proteomic, and invertase activity). The analysis using this dataset obtained a F-value: 22.996, R2: 0.84, and p-value (based on 999 permutations): 0.001.

Tables

Table 1. Classification of the five *Coffea arabica* cultivars studied: genetic origin, resistance to coffee leaf rust (CLR) and nematode, cup quality and coffee production.

Cultivar	Genetic origin (1)	CLR resistance	Nematode resistance	Cup quality	Coffee production
Catuaí Vermelho IAC 144	Caturra amarelo IAC 476-11 x Mundo Novo IAC 374-19	Susceptible (2,3)	Susceptible (2)	Regular (2)	High (2)
Catiguá MG2	Catuaí amarelo IAC 86 x Híbrido de Timor UFV 440-10	Resistant (2,3)	Susceptible (2,4)	Distinct (2)	High (4)
Catiguá MG3	Catuaí amarelo IAC 86 x Híbrido de Timor UFV 440-10	Resistant (2,3)	Resistant (2,4)	Good (4)	High (2,4)
Paraíso 2	Catuaí amarelo IAC 30 x Híbrido de Timor UFV 445-46	Resistant (3)	Susceptible (2)	Distinct (2)	High (2)
Sarchimor MG 8840	Villa Sarchi CIFC 971/10 x Híbrido de Timor CIFC 832/2	Resistant (2)	Susceptible (2)	Regular (2)	High (2)

(1) IAC - Instituto Agronômico de Campinas (Brazil); MG - Minas Gerais (Brazil); CIFC - Centro de Investigação das Ferrugens do Cafeeiro (Portugal); UFV - Universidade Federal de Viçosa (Brazil). (2) De Carvalho et al. (2022); (3) Oliveira et al. (2021); (4) Fazuoli et al. (2008).

Table 2. Leaf Water Potential of *Coffea arabica* cultivar at the onset of the dry season in the Cerrado Mineiro Region. (*) Values represent means ± standard error. Means followed by the same letter within each cultivar, when comparing irrigation treatments, are not significantly different according to Student's *t*-test ($p < 0.05$).

Cultivar	**Leaf water potential (MPa)**	
	Irrigated	**Rainfed**
Catiguá MG2	-1.60 (±0.13) a*	- 0.72 (±0.20) b
Catiguá MG3	-1.34 (±0.21) a	- 0.66 (±0.12) b
Catuaí	-1.38 (±0.10) a	-1.12 (±0.14) a
Paraíso 2	-1.38 (±0.20) a	- 0.82 (±0.15) a
Sarchimor	-1.44 (±0.10) a	- 0.92 (±0.14) b

Table 3. Morphoanatomical characteristics of irrigated (IR) and rainfed (RF) *Coffea arabica* L. cultivars grown under field conditions in the Cerrado Mineiro region, Brazil. (*) Values represent means ± standard error. Means followed by the same letter within each cultivar, when comparing irrigation treatments, are not significantly different according to Student's *t*-test ($p < 0.05$).

Cultivar	Irrigation condition	Leaf	Epidermis Upper	Lower Epidermis (log)	Upper Cuticle	Lower Cuticle (log)	Mesophyll	Palisade parenchyma	Spongy parenchyma (log)
		Thickness (μm)							
Catiguá MG2	IR	327.94 (±12.4)a*	32.80 (±1.23) a	2.99 (±0.06) a	3.90 (±0.08) a	1.21 (±0.05) a	272.47 (±11.1) a	**56.72 (±5.73) b**	5.37 (±0.03) a
	RF	338.00 (±16.5) a	30.39 (±1.03) a	3.05 (±0.04) a	3.94 (±0.20) a	1.17 (±0.03) a	286.62 (±14.2) a	**81.20 (±4.24) a**	5.33 (±0.03) a
Catiguá MG3	IR	351.78 (±8.04) a	30.85 (±0.93) a	3.05 (±0.05) a	3.57 (±0.10) a	**1.02 (±0.08) b**	292.25 (±7.48) a	**58.84 (±2.68) b**	5.46 (±0.04) a
	RF	354.26 (±13.3) a	29.26 (±0.74) a	3.00 (±0.04) a	3.69 (±0.12) a	**1.28 (±0.04) a**	296.19 (±11.5) a	**70.96 (±2.97) a**	5.40 (±0.03) a
Catuaí	IR	301.80 (±4.21) a	25.59 (±0.91) a	2.84 (±0.04) a	**4.17 (±0.08) a**	**0.94 (±0.09) b**	251.97 (±4.62) a	50.99 (±1.96) a	5.30 (±0.02) a
	RF	309.53 (±12.4) a	27.10 (±1.36) a	2.90 (±0.05) a	**3.60 (±0.04) b**	**1.21 (±0.04) a**	261.17 (±12.7) a	58.93 (±7.96) a	5.30 (±0.03) a
Paraíso 2	IR	327.67 (±6.44) a	29.37 (±0.84) a	3.07 (±0.07) a	3.78 (±0.13) a	1.07 (±0.05) a	276.91 (±5.50) a	**53.72 (±3.69) b**	5.42 (±0.01) a
	RF	358.32 (±18.1) a	29.68 (±0.96) a	3.05 (±0.03) a	3.75 (±0.07) a	0.97 (±0.08) a	306.54 (±18.9) a	**85.37 (±9.46) a**	5.41 (±0.04) a
Sarchimor	IR	356.99 (±7.24) a	29.10 (±1.90) a	2.91 (±0.03) a	3.64 (±0.08) a	1.07 (±0.08) a	300.95 (±4.17) a	**57.83 (±3.77) b**	5.43 (±0.05) a
	RF	373.60 (±6.52) a	29.00 (±2.00) a	2.95 (±0.05) a	3.34 (±0.11) a	0.98 (±0.03) a	318. 49 (±7.03) a	**78.49 (±5.25) a**	5.50 (±0.03) a
Cultivar	**Irrigation condition**	**Stomatal polar**	**Stomatal equatorial (log)**	**Xylem Vessel element**	**Xylem (Vascular bundle)**	**Phloem (Vascular bundle)**	**Stomatal pore (log)**	**Stomatal index (log)**	**Specific leaf area**

		--- Diameter (μm) ---			--- Area (μm^2) ---			(%)	Area (cm^2)
Catiguá MG2	IR	23.8 (±0.54) a	**2.69 (±0.02) b**	15.31 (±0.59) a	0.30 (±0.03) a	0.20 (±0.02) a	4.10 (±0.08) a	2.73 (±0.05) a	**101.53 (±3.29) a**
	RF	24.4 (±0.37) a	**2.77 (±0.01) a**	15.04 (±0.86) a	0.39 (±0.05) a	0.28 (±0.04) a	4.08 (±0.02) a	2.74 (±0.05) a	**87.77 (±1.68) b**
Catiguá MG3	IR	23.7 (±0.45) a	**2.69 (±0.01) b**	15.54 (±0.31) a	0.10 (±0.01) a	0.12 (±0.01) a	4.01 (±0.03) a	2.65 (±0.03) a	**106.69 (±1.65) a**
	RF	24.8 (±0.21) a	**2.80 (±0.01) a**	15.49 (±0.58) a	0.16 (±0.03) a	0.14 (±0.01) a	4.06 (±0.05) a	2.72 (±0.08) a	**87.45 (±2.00) b**
Catuaí	IR	23.5 (±0.36) a	2.71 (±0.02) a	13.20 (±0.58) a	0.25 (±0.02) a	0.17 (±0.02) a	**3.97 (±0.04) b**	2.73 (±0.02) a	**113.44 (±0.69) a**
	RF	24.3 (±0.29) a	2.75 (±0.03) a	14.26 (±0.45) a	0.25 (±0.02) a	0.17 (±0.02) a	**4.11 (±0.04) a**	2.81 (±0.04) a	**96.89 (±0.78) b**
Paraíso 2	IR	23.9 (±0.55) a	2.75 (±0.02) a	16.21 (±0.72) a	0.31 (±0.03) a	0.20 (±0.02) a	4.05 (±0.06) a	2.75 (±0.06) a	**112.76 (±5.35) a**
	RF	23.4 (±0.65) a	2.74 (±0.03) a	14.80 (±0.31) a	0.33 (±0.03) a	0.21 (±0.02) a	4.03 (±0.06) a	2.79 (±0.03) a	**93.67 (±3.65) b**
Sarchimor	IR	24.6 (±0.32) a	2.72 (±0.01) a	16.13 (±0.76) a	**0.11 (±0.02) b**	0.11 (±0.01) a	4.03 (±0.04) a	2.71 (±0.06) a	**129.77 (±3.14) a**
	RF	24.2 (±0.44) a	2.71 (±0.02) a	15.47 (±0.36) a	**0.22 (±0.03) a**	0.13 (±0.02) a	3.99 (±0.03) a	2.72 (±0.05) a	**96.88 (±3.91) b**

Table 4. Results of the Permutational multivariate analysis of variance (PERMANOVA) considering the comparison of each cultivar at the two experimental sites. F. Model: F-value. P.adj: adjusted p-value. IR: Irrigated. RF: Rainfed.

Comparison IR and RF	**F.Model**	**R^2**	**p-value**	**p.adj**
Catiguá MG2	14.20	0.64	0.01	0.02
Catiguá MG3	5.32	0.40	0.06	0.08
Catuaí	79.15	0.91	0.01	0.02
Paraíso	77.40	0.91	0.01	0.02
Sarchimor	20.78	0.72	0.01	0.02

Supplementary Table 1 Classification of statistically significant proteins of Catiguá MG2 (*C. arabica* cultivar) in irrigated vs. rainfed plants, according to the results of the volcano plot analysis.

Uniprot code	Protein name	Condition
A0A6P6SDN8_COFAR	Acidic endochitinase-like_2	upregulated
A0A6P6SBP5_COFAR	Anthocyanidin reductase	upregulated
A0A6P6WNF9_COFAR	Auxin-induced protein PCNT115-like isoform X2	upregulated
A0A6P6XC29_COFAR	Calcium-binding allergen Ole e 8-like	upregulated
A0A6P6XGU5_COFAR	Calmodulin	upregulated
A0A6P6TKD8_COFAR	Calmodulin-lysine N-methyltransferase	upregulated
A7IZK7_COFCA	Cell-wall invertase	upregulated
A0A6P6T7D3_COFAR	Citrate-binding protein-like	upregulated
A0A6P6UK35_COFAR	CO(2)-response secreted protease-like	upregulated
A0A6P6TRS3_COFAR	Enhancer of polycomb-like protein	upregulated
A0A6P6SS35_COFAR	Flavin-containing monooxygenase	upregulated
A0A6P6TYE3_COFAR	Glutathione peroxidase_1	upregulated
A0A6P6VH11_COFAR	Glyceraldehyde-3-phosphate dehydrogenase_1	upregulated
A0A6P6WNB0_COFAR	Glycine-rich cell wall structural protein-like_1	upregulated
A0A6P6T5Q0_COFAR	Granule-bound starch synthase 2 chloroplastic/amyloplastic	upregulated
A0A6P6VSY4_COFAR	Heat shock 70 kDa protein mitochondrial	upregulated
A0A6P6SC41_COFAR	LOW QUALITY PROTEIN: protein LHCP TRANSLOCATION DEFECT-like	upregulated
A0A6P6TLX4_COFAR	Non-specific lipid-transfer protein 2-like	upregulated
A0A6P6UBZ5_COFAR	non-specific serine/threonine protein kinase_1	upregulated
A0A6P6V2B1_COFAR	peptidylprolyl isomerase_1	upregulated
A0A6P6UK18_COFAR	Phosphatidylinositol/phosphatidylcholine transfer protein SFH6-like isoform X3	upregulated
A0A1W6CCF1_COFAR	Photosystem I iron-sulfur center	upregulated
A0A1W6CCA3_COFAR	Photosystem II reaction center protein H	upregulated
A0A6P6WI95_COFAR	Pleiotropic drug resistance protein 3-like isoform X6	upregulated
A0A6P6V179_COFAR; A0A6P6VJH1_COFAR	Polyphenol oxidase I chloroplastic-like	upregulated
A0A6P6SI46_COFAR	Prefoldin subunit 1-like	upregulated
A0A6P6U2Y7_COFAR	Probable aquaporin TIP1-1	upregulated
A0A6P6UCL3_COFAR; A0A096VHZ8_9GENT	Probable caffeine synthase 3	upregulated
A0A6P6SND7_COFAR	Proteasome subunit alpha type	upregulated
A0A6P6X5W1_COFAR	Protein NBR1 homolog isoform X1	upregulated

A0A6P6VWN1_COFAR	pyridoxal 5'-phosphate synthase (glutamine hydrolyzing)	upregulated
A0A6P6VPJ7_COFAR; A0A6P6W4L6_COFAR	Remorin-like	upregulated
A0A6P6VV14_COFAR	Ribosomal protein_1	upregulated
A0A6P6U6Z5_COFAR	RING-type E3 ubiquitin transferase_2	upregulated
A0A6P6W8F2_COFAR	Thiamine thiazole synthase chloroplastic	upregulated
A0A6P6X7J7_COFAR	Thioredoxin-like 1-2, chloroplastic	upregulated
A0A6P6SLW8_COFAR	Transmembrane 9 superfamily member	upregulated
A0A6P6SM02_COFAR	1,4-alpha-glucan branching enzyme	downregulated
A0A6P6TNW5_COFAR	14 kDa zinc-binding protein-like isoform X1	downregulated
A0A6P6WHU3_COFAR	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	downregulated
A0A6P6TUF9_COFAR	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase	downregulated
A0A6P6UIQ0_COFAR	30S ribosomal protein S6 alpha, chloroplastic isoform X1	downregulated
A0A6P6SC70_COFAR	ABC transporter G family member 25-like	downregulated
A0A6P6V1H5_COFAR	Acidic endochitinase-like_1	downregulated
A0A6P6WKV7_COFAR	Acidic endochitinase-like_3	downregulated
A0A6P6VY99_COFAR	Acyl carrier protein_2	downregulated
A0A6P6VY99_COFAR	ADP/ATP translocase	downregulated
A0A6P6TM10_COFAR	alanine--glyoxylate transaminase	downregulated
A0A6P6ULZ1_COFAR	alcohol dehydrogenase	downregulated
A0A1W6CCB7_COFAR	ATP synthase CF1 epsilon subunit	downregulated
A0A6P6VNJ3_COFAR	ATP synthase subunit delta' mitochondrial-like	downregulated
A0A6P6VSP7_COFAR	BAHD acyltransferase At5g47980-like isoform X2	downregulated
A0A6P6WG71_COFAR	beta-galactosidase_2	downregulated
A0A6P6SZK1_COFAR	Blue copper protein-like	downregulated
A0A6P6SDY0_COFAR; A0A6P6SD75_COFAR; A0A6P6URI6_COFAR	chitinase_2	downregulated
A0A6P6X7K8_COFAR	Chlorophyll a-b binding protein chloroplastic_6	downregulated
A0A6P6V4E0_COFAR	Chlorophyll a-b binding protein chloroplastic_7	downregulated
A0A6P6V0C6_COFAR	Chlorophyllase-1-like	downregulated
A0A6P6TM83_COFAR	COP1-interactive protein 1-like	downregulated
A0A6P6SI29_COFAR	Cysteine proteinase 3-like	downregulated
A0A1W6CCH4_COFAR	Cytochrome f	downregulated
A0A6P6UDD2_COFAR	Cytochrome P450 97B2 chloroplastic-like	downregulated
A0A6P6U1X1_COFAR	Delta-aminolevulinic acid dehydratase	downregulated

A0A6P6VZU1_COFAR	DNA damage-repair/toleration protein DRT102-like	downregulated
A0A6P6TUZ3_COFAR	Elongation factor 1-beta 2-like	downregulated
A0A6P6X7C7_COFAR	Elongation factor 1-delta-like	downregulated
A0A6P6T3Q2_COFAR; A0A6P6VAG5_COFAR	Eukaryotic translation initiation factor 5A	downregulated
A0A6P6T7E4_COFAR	Ferredoxin--NADP reductase chloroplastic	downregulated
A0A6P6XNL1_COFAR	Fructose-bisphosphate aldolase_2	downregulated
A0A6P6SHR3_COFAR	Glucan endo-1, 3-beta-glucosidase acidic-like_1	downregulated
A0A6P6UPQ6_COFAR; A0A6P6V2M1_COFAR; A0A6P6VC84_COFAR	Glyceraldehyde-3-phosphate dehydrogenase_3	downregulated
A0A6P6U1Q5_COFAR	Glycine cleavage system P protein	downregulated
A0A6P6VRR9_COFAR	Glycine-rich RNA-binding abscisic acid-inducible protein-like	downregulated
A0A6P6SGL2_COFAR	Heme-binding protein 2-like	downregulated
A0A6P6SYE0_COFAR	Histone H2B	downregulated
A0A6P6TSN2_COFAR	IQ domain-containing protein IQM1-like	downregulated
A0A6P6SGT0_COFAR	Ketol-acid reductoisomerase	downregulated
A0A6P6ULJ0_COFAR	L-ascorbate oxidase homolog	downregulated
A0A6P6VK88_COFAR	Late embryogenesis abundant protein At5g17165-like	downregulated
A0A6P6U5G9_COFAR	Light-induced protein chloroplastic-like	downregulated
A0A6P6V1Z2_COFAR	Lipoxygenase	downregulated
A0A6P6S9I3_COFAR	LOW QUALITY PROTEIN: chloroplastic lipocalin-like	downregulated
A0A6P6UJW0_COFAR	LRR receptor-like serine/threonine-protein kinase RCH1	downregulated
A0A6P6SQ38_COFAR	Macrophage migration inhibitory factor homolog isoform X1	downregulated
A0A6P6V3F1_COFAR	Mitochondrial outer membrane protein porin of 36 kDa-like	downregulated
A0A6P6UZK2_COFAR	MLP-like protein 34	downregulated
A0A6P6VXI3_COFAR	NAD(P)H dehydrogenase (quinone)	downregulated
A0A6P6WV53_COFAR	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 6-like	downregulated
A0A6P6X1U0_COFAR	non-specific serine/threonine protein kinase_2	downregulated
A0A6P6VEG9_COFAR	Non-symbiotic hemoglobin 2-like	downregulated
A0A6P6UGW4_COFAR; A0A6P6SGA0_COFAR	Patatin	downregulated
A0A6P6VI58_COFAR	Pentatricopeptide repeat-containing protein At3g46790, chloroplastic-like	downregulated
A0A6P6VFW0_COFAR	peptidylprolyl isomerase_6	downregulated
A0A6P6T4J5_COFAR	Peroxidase_1	downregulated
A0A6P6ULL6_COFAR	Plastocyanin	downregulated

A0A6P6SS20_COFAR	Polyadenylate-binding protein 5-like isoform X1	downregulated
A0A6P6VSB3_COFAR	Presequence protease 1 chloroplastic/mitochondrial-like	downregulated
A0A6P6W8S4_COFAR	Probable calcium-binding protein CML13	downregulated
A0A6P6VHK7_COFAR	Probable carboxylesterase 9	downregulated
A0A6P6TYG4_COFAR	Probable steroid-binding protein 3	downregulated
A0A6P6U277_COFAR	Protein HOTHEAD-like	downregulated
A0A6P6XC49_COFAR	Protein MET1 chloroplastic-like isoform X1	downregulated
A0A6P6S6N3_COFAR	Protein THYLAKOID FORMATION1 chloroplastic	downregulated
A0A6P6TR78_COFAR	Protein unc-13 homolog	downregulated
A0A6P6SQZ6_COFAR	protein-serine/threonine phosphatase	downregulated
A0A6P6S7E2_COFAR	Reactive Intermediate Deaminase A chloroplastic-like	downregulated
A0A6P6U4J4_COFAR	ribose-5-phosphate isomerase	downregulated
A0A6P6SHB8_COFAR	Ribosomal protein_2	downregulated
A0A6P6VV71_COFAR	Ribulose bisphosphate carboxylase/oxygenase activase chloroplastic_2	downregulated
A0A6P6SQ02_COFAR	Ricin B-like lectin R40G3	downregulated
A0A6P6XBB8_COFAR	RuBisCO large subunit-binding protein subunit beta chloroplastic-like_2	downregulated
A0A6P6U6S6_COFAR	RuvB-like helicase	downregulated
A0A6P6UKU0_COFAR	Serine hydroxymethyltransferase	downregulated
A0A6P6WMJ5_COFAR	Stress-response A/B barrel domain-containing protein HS1-like	downregulated
A0A6P6VMW9_COFAR	Thylakoid lumenal 15 kDa protein 1 chloroplastic-like	downregulated
A0A6P6XIM1_COFAR	TPR repeat-containing thioredoxin TTL1 isoform X2	downregulated
A0A6P6TF28_COFAR	UDP-arabinopyranose mutase_2	downregulated

Supplementary Table 2 Classification of the statistically significant metabolic features of Catiguá MG2 (*C. arabica* cultivar) in irrigated vs. rainfed plants, according to the results of the volcano plot analysis.

Metabolic feature	Metabolite	Metabolic class	Condition
1-Pyrroline-2-carboxylate (1TMS)	1-Pyrroline-2-carboxylate	Amino acids	upregulated
1,3-Dihydroxyaceton (1MEOX) (2TMS) MP	Dihydroxyacetone	Sugars	upregulated
1,4-Naphthoquinone, 2-methyl-	Menadione	Others	upregulated
4-(Methylamino)benzoic acid (2TMS)	N-Methyl-4-aminobenzoate	Other organic acids	upregulated
Abscisic acid (1MEOX) (1TMS)_1	(+)-Abscisic acid	Lipids	upregulated
Abscisic acid (1MEOX) (1TMS)_2	(+)-Abscisic acid	Lipids	upregulated
Acetic acid, 3,4-dihydroxyphenyl- (3TMS)	3,4-Dihydroxyphenylacetate	Phenylpropanoids or phenolic compounds	upregulated
Acetoacetic acid (1MEOX) (1TMS) BP	Acetoacetate	Lipids	upregulated
Allose (1MEOX) (5TMS) MP	D-Allose	Sugars	upregulated
Arabitol (5TMS)	D-Arabitol	Sugars	upregulated
Benzoic acid, 3-hydroxy- (2TMS)	3-Hydroxybenzoic acid	Other organic acids	upregulated
Benzylalcohol (1TMS)	Benzyl alcohol	Others	upregulated
Butanoic acid, 2-hydroxy- (2TMS)	2-Hydroxybutyric acid	Lipids	upregulated
Butanoic acid, 2-oxo- (1MEOX) (1TMS) MP	2-Oxobutanoate	Lipids	upregulated
Butanoic acid, 3-hydroxy- (2TMS)	D-beta-Hydroxybutyric acid	Lipids	upregulated
Butyro-1,4-lactam, 2-amino- (2TMS)	3-Aminopyrrolidin-2-one	Others	upregulated
Cinnamaldehyde, 3,5-dimethoxy-4-hydroxy-, trans- (1MEOX) (1TMS) BP	Sinapaldehyde	Phenylpropanoids or phenolic compounds	upregulated
Citrulline (3TMS)	L-Citrulline	Amino acids	upregulated
Cyclohexane-1,2-dicarboxylic acid, trans- (2TMS)	1,2-Cyclohexanedicarboxylic acid	Other organic acids	upregulated
Cystineamine (4TMS)	Cystamine	Others	upregulated
Dehydroascorbic acid dimer (2MEOX) MP	Dehydroascorbic acid	Others	upregulated
Digitoxose (1MEOX) (3TMS) MP	D-Digitoxose	Sugars	upregulated
Epigallocatechin (6TMS)	(-)-Epigallocatechin	Phenylpropanoids or phenolic compounds	upregulated
Eriodictyol (5TMS)	Eriodictyol	Phenylpropanoids or phenolic compounds	upregulated

Erythrose-4-phosphate (1MEOX) (4TMS) MP	D-Erythrose 4-phosphate	Sugars	upregulated
Fructose-1-phosphate (6TMS) MP	D-Fructose 1-phosphate	Sugars	upregulated
Fucose (1MEOX) (4TMS) MP	2-Deoxyglucose	Sugars	upregulated
Fumaric acid, 2-methyl- (2TMS)	Mesaconate	Lipids	upregulated
Glucopyranoside, 1-O-methyl-, alpha- (4TMS)	Methyl alpha-D-galactopyranoside	Sugars	upregulated
Glucose (1MEOX) (5TMS) MP	D-Glucose	Sugars	upregulated
Glucose, 2-amino-2-deoxy-, D- (1MEOX) (5TMS)	2-Amino-2-deoxy-D-glucose	Sugars	upregulated
Glutamic acid, N-acetyl- (2TMS)	N-Acetyl-L-glutamate	Amino acids	upregulated
Glutamic acid, N-acetyl- (3TMS)	N-Acetyl-L-glutamate	Amino acids	upregulated
Glutaric acid (2TMS)	Glutaric acid	Lipids	upregulated
Glycerol (3TMS)	Glycerol	Sugars	upregulated
Glycerol-3-phosphate (4TMS)	Glycerol-3-phosphate	Lipids	upregulated
Glycolic acid (2TMS)	Glycolic acid	Other organic acids	upregulated
Glycylglycine, DL- (4TMS)	Glycyl-glycine	Amino acids	upregulated
Guanine (3TMS)	Guanine	Others	upregulated
Guanosine-2*,3*-cyclic-monophosphate (4TMS)	2',3'-Cyclic GMP	Others	upregulated
Hexadecan-1-ol, n- (1TMS)	1-Hexadecanol	Lipids	upregulated
Hippuric acid (1TMS)	Hippuric acid	Amino acids	upregulated
Histidine (3TMS)	L-Histidine	Amino acids	upregulated
Histidine, N-tau-methyl- (3TMS)	1-Methylhistidine	Amino acids	upregulated
Homoserine lactone, N-heptanoyl-	N-Heptanoylhomoserine lactone	Lipids	upregulated
Homoserine lactone, N-hexanoyl-	N-Hexanoyl-L-homoserine lactone	Lipids	upregulated
Hydantoin, 5-propionate- (3TMS)	Hydantoin-5-propionate	Others	upregulated
Hydrocaffeic acid (3TMS)	3,4-Dihydroxyphenylpropanoate	Phenylpropanoids or phenolic compounds	upregulated
Hydroquinone (2TMS)	Hydroquinone	Phenylpropanoids or phenolic compounds	upregulated
Indole-3-acetamide (3TMS)	Indole-3-acetamide	Others	upregulated
Inositol, myo- (6TMS)	Inositol	Others	upregulated
Isoflavone, 5,7-dihydroxy-4*-methoxy- (2TMS)	Biochanin A	Phenylpropanoids or phenolic compounds	upregulated
Isomaltose (1MEOX) (8TMS) MP	Isomaltose	Sugars	upregulated
Lactic acid, 3-(4-hydroxyphenyl)- (3TMS)	2-Hydroxy-3-(4-hydroxyphenyl)propenoate	Other organic acids	upregulated

Lipoamide, alpha- (1TMS)	Lipoamide	Lipids	upregulated
Lysine, N-epsilon-acetyl- (3TMS)	N6-Acetyl-L-lysine	Amino acids	upregulated
Malic acid, 3-oxalo- (1MEOX) (4TMS) MP	3-Oxalomalate	Other organic acids	upregulated
Menthol, L-(-)- (1TMS)	L-Menthol	Lipids	upregulated
Octacosane, n-	Octacosane	Others	upregulated
Octanoic acid, n- (1TMS)	Octanoic acid	Lipids	upregulated
Orotic acid (3TMS)	Orotic acid	Others	upregulated
Orotic acid, 4,5-dihydro- (3TMS)	L-Dihydroorotate	Others	upregulated
Oxalic acid (2TMS)	Oxalate	Other organic acids	upregulated
Oxaloacetate (1MEOX) (2TMS) BP	Oxalacetic acid	Lipids	upregulated
Pentanedioic acid, 3-methyl- (2TMS)	3-Methylglutaric acid	Lipids	upregulated
Perillic acid (1TMS)	Perillic acid	Lipids	upregulated
Phenethylamine (2TMS)	Phenethylamine	Amines	upregulated
Phytol (1TMS) MP	Phytol	Lipids	upregulated
Piceatannol (4TMS) MP	Piceatannol	Phenylpropanoids or phenolic compounds	upregulated
Proline, 3-hydroxy-, trans- (3TMS)	trans-3-Hydroxy-L-proline	Amino acids	upregulated
Propanoic acid, 2-amino-2-methyl-3-hydroxy- (2TMS)	2-Methylserine	Amino acids	upregulated
Psicose (1MEOX) (5TMS) BP	D-Psicose	Sugars	upregulated
Putrescine (3TMS)	Putrescine	Amines	upregulated
Pyridine, 3-hydroxy- (1TMS)	3-Hydroxypyridine	Others	upregulated
Pyruvic acid, 4-hydroxyphenyl- (1MEOX) (3TMS) MP	4-Hydroxyphenylpyruvate	Phenylpropanoids or phenolic compounds	upregulated
Quinic acid (5TMS)	Quinic acid	Others	upregulated
Quinic acid, 3-caffeoyl-, trans- (6TMS)	Chlorogenic acid	Phenylpropanoids or phenolic compounds	upregulated
Ribose, 2-deoxy- (1MEOX) (3TMS) BP	Deoxyribose	Sugars	upregulated
Ribulose (1MEOX) (4TMS) BP2	D-Ribulose	Sugars	upregulated
Saccharopine (5TMS)	L-Saccharopine	Amino acids	upregulated
Senecionine	Senecionine	Alkaloids	upregulated
Senecionine (1TMS) BP2	Senecionine	Alkaloids	upregulated
Shikimic acid (4TMS)	Shikimate	Other organic acids	upregulated
Squalene, all-trans-	Squalene	Lipids	upregulated
Succinic acid, 2,3-dimethyl- (2TMS)	2,3-Dimethylsuccinic acid	Lipids	upregulated

Sucrose (8TMS)	Sucrose	Sugars	upregulated
Tartronic acid, 2-(methylaminomethyl)- (1MEOX) (3TMS)	2-(Methylaminomethyl)tartronic acid	Others	upregulated
Taxifolin (1MEOX) (5TMS) MP	Taxifolin	Phenylpropanoids or phenolic compounds	upregulated
Tetracosane, n-	Lignocerane	Others	upregulated
Thiazole, 4-methyl-5-hydroxyethyl- (1TMS)	4-Methyl-5-(2'-hydroxyethyl)-thiazole	Alkaloids	upregulated
Tryptamine, 1-methyl- (2TMS)	N-Methyltryptamine	Alkaloids	upregulated
Tyramine (3TMS)	Tyramine	Amines	upregulated
Urea (2TMS)	Urea	Amides	upregulated
Valero-1,5-lactam (1TMS)	N-Methyl-2-pyrrolidinone	Others	upregulated
Alanine, beta- (3TMS)	beta-Alanine	Amino acids	downregulated
Aspartic acid (3TMS)	L-Aspartate	Amino acids	downregulated
Caffeine	Caffeine	Alkaloids	downregulated
Cystathionine (4TMS)	L-Cystathionine	Amino acids	downregulated
Cysteine, S-methyl-, DL- (2TMS)	S-N-Methylcysteine	Amino acids	downregulated
Gulonic acid-1,4-lactone (4TMS)	L-Galactono-1,4-lactone	Others	downregulated
Gulonic acid, 2-oxo-, DL- (1MEOX) (5TMS) MP	2-Keto-L-gluconate	Sugars	downregulated
Hypoxanthine (2TMS)	Hypoxanthine	Alkaloids	downregulated
Lipoamide, alpha- (2TMS)	Lipoamide	Amides	downregulated
Maleamic acid (2TMS) BP	Maleamic acid	Lipids	downregulated
Malic acid, 2-isopropyl- (3TMS)	2-Isopropylmalic acid	Lipids	downregulated
Maltose (1MEOX) (8TMS) BP	D-Maltose	Sugars	downregulated
Mannopyranoside, 1-O-methyl-, alpha- (4TMS)	Methyl alpha-D-galactopyranoside	Sugars	downregulated
Melibiose (1MEOX) (8TMS) BP	Melibiose	Sugars	downregulated
Oxaloacetate (1MEOX) (3TMS) MP	Oxaloacetate	Other organic acids	downregulated
Pantothenic acid, D- (4TMS)	Pantothenate	Others	downregulated
Phenylalanine (2TMS)	L-Phenylalanine	Amino acids	downregulated
Phenylpyruvic acid (2TMS)	Phenylpyruvate	Other organic acids	downregulated
Spermine (6TMS)	Spermine	Amines	downregulated
Tropic acid (2TMS)	Tropic acid	Other organic acids	downregulated
Tryptophan (2TMS)_1	Tryptophan	Amino acids	downregulated
Valine (2TMS)	L-Valine	Amino acids	downregulated

Supplementary Table 3 Classification of statistically significant proteins of Catiguá MG3 (*C. arabica* cultivar) in irrigated vs. rainfed plants, according to the results of the volcano plot analysis.

Uniprot code	Protein name	Condition
A0A6P6UH18_COFAR	ATP synthase delta chain chloroplastic-like	upregulated
A0A6P6TAF6_COFAR	Caltractin-like	upregulated
A0A6P6TEJ5_COFAR	Carbonic anhydrase	upregulated
A0A6P6SI29_COFAR	Cysteine proteinase 3-like	upregulated
A0A6P6WQV5_COFAR	Cysteine synthase_2	upregulated
A0A1W6CCG4_COFAR	Cytochrome b6	upregulated
A0A6P6T7E4_COFAR	Ferredoxin--NADP reductase chloroplastic	upregulated
A0A6P6SMW1_COFAR	Glucan endo-1,3-beta-glucosidase acidic-like_2	upregulated
A0A6P6UJK4_COFAR	glutathione transferase_4	upregulated
A0A6P6UDN8_COFAR	Glycine cleavage system H protein_2	upregulated
A0A6P6WWZ4_COFAR	Glycine-rich cell wall structural protein-like_2	upregulated
A0A6P6VRR9_COFAR	Glycine-rich RNA-binding abscisic acid-inducible protein-like	upregulated
A0A6P6UNQ4_COFAR	Low-temperature-induced cysteine proteinase-like	upregulated
A0A6P6V3F6_COFAR	Major allergen Pru ar 1-like_2	upregulated
A0A6P6WWQ6_COFAR	Mitochondrial proton/calcium exchanger protein	upregulated
A0A6P6UZK2_COFAR	MLP-like protein 34	upregulated
A0A6P6VI58_COFAR	Pentatricopeptide repeat-containing protein At3g46790, chloroplastic-like	upregulated
A0A6P6SBQ7_COFAR	Photosystem II repair protein PSB27-H1 chloroplastic-like	upregulated
A0A6P6UKU7_COFAR	Probable cysteine protease RD19C	upregulated
A0A6P6VV71_COFAR	Ribulose bisphosphate carboxylase/oxygenase activase chloroplastic_2	upregulated
A0A6P6XCE1_COFAR	ribulose-phosphate 3-epimerase	upregulated
A0A6P6W748_COFAR	Sedoheptulose-1,7-bisphosphatase chloroplastic	upregulated
A0A6P6T2L6_COFAR	serine O-acetyltransferase	upregulated
A0A6P6VV67_COFAR	transketolase	upregulated
A0A6P6W7I5_COFAR	Triosephosphate isomerase cytosolic	upregulated
A0A6P6VEX3_COFAR	Triosephosphate isomerase, cytosolic-like	upregulated
A0A6P6VUZ0_COFAR	Alpha-N-acetylglucosaminidase-like isoform X2	downregulated
A0A6P6XIP3_COFAR	ATP-dependent Clp protease ATP-binding subunit ClpA homolog CD4B chloroplastic	downregulated

A0A6P6TM83_COFAR	COP1-interactive protein 1-like	downregulated
A0A6P6VD82_COFAR	Cytochrome b561 domain-containing protein At4g18260-like isoform X1	downregulated
A0A6P6WQ90_COFAR	Disease resistance protein RGA3 isoform X1	downregulated
A0A6P6TRS3_COFAR	Enhancer of polycomb-like protein	downregulated
A0A6P6SGT0_COFAR	Ketol-acid reductoisomerase	downregulated
A0A6P6SR15_COFAR	LOW QUALITY PROTEIN: ATP-dependent zinc metalloprotease FTSH 8 mitochondrial-like	downregulated
A0A6P6UHZ3_COFAR	Monooxygenase 2-like	downregulated
A0A6P6UBZ5_COFAR	non-specific serine/threonine protein kinase_1	downregulated
A0A6P6TAR2_COFAR	peptidylprolyl isomerase_3	downregulated
A0A6P6S4Q8_COFAR	peptidylprolyl isomerase_4	downregulated
A0A6P6SSQ7_COFAR	Peroxiredoxin Q chloroplastic	downregulated
A0A6P6ULL6_COFAR	Plastocyanin	downregulated
A0A6P6SS20_COFAR	Polyadenylate-binding protein 5-like isoform X1	downregulated
A0A6P6U2Y7_COFAR	Probable aquaporin TIP1-1	downregulated
A0A6P6V4A5_COFAR	Protein C2-DOMAIN ABA-RELATED 7-like isoform X2	downregulated
A0A6P6VXX5_COFAR	Protein SUPPRESSOR OF QUENCHING 1 chloroplastic-like	downregulated
A0A6P6WTK0_COFAR	Protein transport protein SEC23	downregulated
A0A6P6WE36_COFAR	Putative rRNA methyltransferase	downregulated
A0A6P6SHB8_COFAR	Ribosomal protein_2	downregulated
A0A6P6XBB8_COFAR	RuBisCO large subunit-binding protein subunit beta chloroplastic-like_2	downregulated
A0A6P6UBY8_COFAR	Thioredoxin O2, mitochondrial-like	downregulated
A0A6P6XIM1_COFAR	TPR repeat-containing thioredoxin TTL1 isoform X2	downregulated
A0A6P6SLW8_COFAR	Transmembrane 9 superfamily member	downregulated
A0A6P6U6L6_COFAR	Tyrosine decarboxylase 1-like	downregulated

Supplementary Table 4 Classification of the statistically significant metabolic features of Catiguá MG3 (*C. arabica* cultivar) in irrigated vs. rainfed plants, according to the results of the volcano plot analysis.

Metabolic feature	Metabolite	Metabolic class	Condition
Ascorbic acid (4TMS)	L-Ascorbic acid	Others	upregulated
Butanoic acid, 4-amino- (2TMS)	gamma-Aminobutyric acid	Amino acids	upregulated
Dopamine (4TMS)	Dopamine	Amines	upregulated
Galactonic acid-1,4-lactone (4TMS)	D-Galactono-1,4-lactone	Others	upregulated
Glyceric acid (3TMS)	L-Glycerate	Sugars	upregulated
Guanine (3TMS)	Guanine	Others	upregulated
Inosine-5*-monophosphate (5TMS)	Inosinic acid	Others	upregulated
Inositol, myo- (6TMS)	Inositol	Others	upregulated
Isocaproic acid, 2-oxo- (1MEOX) (1TMS) MP	4-Methyl-2-oxopentanoate	Other organic acids	upregulated
Malic acid, 2-methyl- (3TMS)	D-Citramalate	Lipids	upregulated
Octopamine (4TMS)	Octopamine	Other organic acids	upregulated
Ornithine, N2-acetyl- (3TMS)	N2-Acetylornithine	Amino acids	upregulated
Phenethylamine (2TMS)	Phenethylamine	Amines	upregulated
Quinic acid, 1-caffeoyl-, trans- (6TMS)	1-Caffeoylquinic acid	Phenylpropanoids or phenolic compounds	upregulated
Ribonic acid (5TMS)	D-Ribonate	Sugars	upregulated
Shikimic acid, 3-dehydro- (1MEOX) (3TMS) MP	3-Dehydroshikimate	Phenylpropanoids or phenolic compounds	upregulated
Tryptophan (2TMS)_1	Tryptophan	Amino acids	upregulated
Urea (2TMS)	Urea	Amides	upregulated
Caffeine	Caffeine	Alkaloids	downregulated
Glucoheptose (1MEOX) (6TMS) MP	d-Glycero-d-galacto-heptose	Sugars	downregulated
Glutamic acid, N-acetyl- (3TMS)	N-Acetyl-L-glutamate	Amino acids	downregulated
Indole-3-acetic acid, 5-hydroxy-, 1H- (2TMS)	5-Methoxyindoleacetate	Others	downregulated
Inosine (4TMS)	Inosine	Others	downregulated
Levodopa (4TMS)	L-Dopa	Amino acids	downregulated

Lipoamide, alpha- (2TMS)	Lipoamide	Amides	downregulated
Lyxose (1MEOX) (4TMS) MP	D-lyxose	Sugars	downregulated
Propanoic acid, 3-amino-3-(4-hydroxyphenyl)- (3TMS)	beta-Tyrosine	Amino acids	downregulated
Theophylline (1TMS)	Theophylline	Alkaloids	downregulated

Supplementary Table 5 Classification of statistically significant proteins of Catuaí (*C. arabica* cultivar) in irrigated vs. rainfed plants, according to the results of the volcano plot analysis.

Uniprot code	Protein name	Condition
A0A6P6WHU3_COFAR	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase	upregulated
A0A6P6ULI3_COFAR	Abscisic stress-ripening protein 2-like	upregulated
A0A6P6ULZ1_COFAR	alcohol dehydrogenase	upregulated
A0A6P6SGA1_COFAR	AUGMIN subunit 1-like	upregulated
A0A6P6VSP7_COFAR	BAHD acyltransferase At5g47980-like isoform X2	upregulated
A0A6P6TKD8_COFAR	Calmodulin-lysine N-methyltransferase	upregulated
A0A6P6TAF6_COFAR	Caltractin-like	upregulated
A0A6P6V280_COFAR	Chloroplast stem-loop binding protein of 41 kDa a chloroplastic-like	upregulated
A0A6P6T7D3_COFAR	Citrate-binding protein-like	upregulated
A0A6P6VXK6_COFAR A0A6P6VXI4_COFAR A0A6P6VYB8_COFAR	glutathione transferase_1	upregulated
A0A6P6VH11_COFAR	Glyceraldehyde-3-phosphate dehydrogenase_1	upregulated
A0A6P6S6L4_COFAR	Glyceraldehyde-3-phosphate dehydrogenase_2	upregulated
A0A6P6WNB0_COFAR	Glycine-rich cell wall structural protein-like_1	upregulated
A0A6P6T5Q0_COFAR	Granule-bound starch synthase 2 chloroplastic/amyloplastic	upregulated
A0A6P6S9I3_COFAR	LOW QUALITY PROTEIN: chloroplastic lipocalin-like	upregulated
A0A6P6SC41_COFAR	LOW QUALITY PROTEIN: protein LHCP TRANSLOCATION DEFECT-like	upregulated
A0A6P6WQQ5_COFAR;A0A6P6UAU8_COFAR	Major allergen Pru ar 1-like_1	upregulated
A0A6P6VEG9_COFAR	Non-symbiotic hemoglobin 2-like	upregulated
A0A6P6VFW0_COFAR	peptidylprolyl isomerase_6	upregulated
A0A6P6UK18_COFAR	Phosphatidylinositol/phosphatidylcholine transfer protein SFH6-like isoform X3	upregulated
A0A1W6CCF1_COFAR	Photosystem I iron-sulfur center	upregulated
A0A6P6XEF9_COFAR	Photosystem I reaction center subunit VI	upregulated
A0A6P6UQK6_COFAR	Photosystem II 22 kDa protein chloroplastic-like	upregulated
A0A6P6VSB9_COFAR;A0A6P6VTW1_COFAR	Prostaglandin reductase-3-like isoform X1	upregulated
A0A6P6V9A1_COFAR	Protein trichome birefringence-like 6	upregulated
A0A6P6TR78_COFAR	Protein unc-13 homolog	upregulated

A0A6P6VWN1_COFAR	pyridoxal 5'-phosphate synthase (glutamine hydrolyzing)	upregulated
A0A6P6VPJ7_COFAR;A0A6P6W4L6_COFAR	Remorin-like	upregulated
A0A6P6W2L2_COFAR	RING-type E3 ubiquitin transferase_1	upregulated
A0A6P6U6Z5_COFAR	RING-type E3 ubiquitin transferase_2	upregulated
A0A6P6T2L6_COFAR	serine O-acetyltransferase	upregulated
A0A6P6UMW7_COFAR	Serpin-ZX-like	upregulated
A0A6P6VD95_COFAR	Temperature-induced lipocalin-1-like isoform X2	upregulated
A0A6P6W8F2_COFAR	Thiamine thiazole synthase chloroplastic	upregulated
A0A6P6UY08_COFAR	Thioredoxin X, chloroplastic-like	upregulated
A0A6P6SLW8_COFAR	Transmembrane 9 superfamily member	upregulated
A0A6P6TNW5_COFAR	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	downregulated
A0A6P6SGR8_COFAR	20 kDa chaperonin chloroplastic	downregulated
A0A6P6SZ87_COFAR	29 kDa ribonucleoprotein A chloroplastic	downregulated
A0A6P6TUF9_COFAR	30S ribosomal protein S6 alpha, chloroplastic isoform X1	downregulated
A0A6P6UIQ0_COFAR	ABC transporter G family member 25-like	downregulated
A0A6P6SDN8_COFAR	Acidic endochitinase-like_3	downregulated
A0A6P6VUZ1_COFAR	Actin-7-like	downregulated
A0A6P6TM10_COFAR	alanine--glyoxylate transaminase	downregulated
A0A6P6V602_COFAR	Aldehyde dehydrogenase family 2 member B7 mitochondrial-like	downregulated
A0A6P6X541_COFAR	Alpha-mannosidase	downregulated
A0A6P6VUZ0_COFAR	Alpha-N-acetylglucosaminidase-like isoform X2	downregulated
A0A6P6V1Z7_COFAR	Alpha-xylosidase 1-like	downregulated
A0A6P6TY02_COFAR	Amine oxidase_2	downregulated
A0A1W6CCE4_COFAR	ATP synthase CF0 B subunit	downregulated
A0A6P6UH18_COFAR	ATP synthase delta chain chloroplastic-like	downregulated
A0A6P6VNJ3_COFAR	ATP synthase subunit delta' mitochondrial-like	downregulated
A0A6P6XIP3_COFAR	ATP-dependent Clp protease ATP-binding subunit ClpA homolog CD4B chloroplastic	downregulated
A0A6P6UI56_COFAR	ATP-dependent DNA helicase_1	downregulated
A0A6P6S4T8_COFAR	ATP-dependent zinc metalloprotease FTSH 2 chloroplastic	downregulated
A0A6P6W7A5_COFAR;A0A6P6UPG2_COFAR	CDGSH iron-sulfur domain-containing protein NEET-like isoform X1	downregulated
A0A6P6SQN5_COFAR	Cell wall hydroxyproline-rich glycoprotein	downregulated

A0A6P6WT68_COFAR	Chaperonin CPN60-2 mitochondrial	downregulated
A0A6P6U9B0_COFAR	Chlorophyll a-b binding protein chloroplastic_1	downregulated
A0A6P6SD20_COFAR	Chlorophyll a-b binding protein chloroplastic_5	downregulated
A0A6P6SXI1_COFAR	Chromodomain-helicase-DNA-binding protein 5-like	downregulated
A0A6P6TM83_COFAR	COP1-interactive protein 1-like	downregulated
A0A6P6VD82_COFAR	Cytochrome b561 domain-containing protein At4g18260-like isoform X1	downregulated
A0A1W6CCG4_COFAR	Cytochrome b6	downregulated
A0A6P6UDD2_COFAR	Cytochrome P450 97B2 chloroplastic-like	downregulated
A0A6P6S5I2_COFAR	Disease resistance protein RGA3	downregulated
A0A6P6VZU1_COFAR	DNA damage-repair/toleration protein DRT102-like	downregulated
A0A6P6ULR3_COFAR	Endo-1, 31, 4-beta-D-glucanase-like isoform X2	downregulated
A0A6P6T3Q2_COFAR;A0A6P6VAG5_COFAR	Eukaryotic translation initiation factor 5A	downregulated
A0A6P6XNL1_COFAR	Fructose-bisphosphate aldolase_2	downregulated
A0A6P6SSC1_COFAR	FT-interacting protein 1-like	downregulated
A0A6P6VSR5_COFAR	GDSL esterase/lipase At1g29670-like isoform X1	downregulated
A0A6P6TMJ4_COFAR	glutathione transferase_2	downregulated
A0A6P6V6U4_COFAR	glutathione transferase_3	downregulated
A0A6P6UPQ6_COFAR;A0A6P6V2M1_COFAR;A0A6P6VC84_COFAR	Glyceraldehyde-3-phosphate dehydrogenase_3	downregulated
A0A6P6U1Q5_COFAR	Glycine cleavage system P protein	downregulated
A0A6P6V142_COFAR	Heat shock protein 90-5 chloroplastic	downregulated
A0A6P6SYE0_COFAR	Histone H2B	downregulated
A0A6P6WKM3_COFAR	isoleucine--tRNA ligase	downregulated
A0A6P6SGT0_COFAR	Ketol-acid reductoisomerase	downregulated
A0A6P6V351_COFAR;A0A6P6W4C1_COFAR;A0A6P6V5E4_COFAR	Kirola-like	downregulated
A0A6P6VST3_COFAR;H2EQR2_9GENT	L-ascorbate peroxidase 3-like	downregulated
A0A6P6VK88_COFAR	Late embryogenesis abundant protein At5g17165-like	downregulated
A0A6P6U5G9_COFAR	Light-induced protein chloroplastic-like	downregulated
A0A6P6SR15_COFAR	LOW QUALITY PROTEIN: ATP-dependent zinc metalloprotease FTSH 8 mitochondrial-like	downregulated

A0A6P6UJW0_COFAR	LRR receptor-like serine/threonine-protein kinase RCH1	downregulated
A0A6P6SQ38_COFAR	Macrophage migration inhibitory factor homolog isoform X1	downregulated
A0A6P6VSV4_COFAR	Malate dehydrogenase_4	downregulated
A0A6P6SQ25_COFAR	Mitochondrial-processing peptidase subunit alpha-like	downregulated
A0A6P6UHZ3_COFAR	Monooxygenase 2-like	downregulated
A0A6P6VXI3_COFAR	NAD(P)H dehydrogenase (quinone)	downregulated
A0A6P6X1U0_COFAR	non-specific serine/threonine protein kinase_2	downregulated
A0A6P6X3P4_COFAR	Pentatricopeptide repeat-containing protein At3g48250 chloroplastic-like	downregulated
A0A6P6T4J5_COFAR	Peroxidase_1	downregulated
A0A6P6V011_COFAR	Phosphoglycerate kinase_2	downregulated
A0A6P6T3Q6_COFAR	phosphopyruvate hydratase	downregulated
A0A1W6CCG6_COFAR	Photosystem II D2 protein	downregulated
A0A6P6SBQ7_COFAR	Photosystem II repair protein PSB27-H1 chloroplastic-like	downregulated
A0A6P6SS20_COFAR	Polyadenylate-binding protein 5-like isoform X1	downregulated
A0A6P6VHK7_COFAR	Probable carboxylesterase 9	downregulated
A0A6P6TBH4_COFAR	Probable transcription-associated protein 1	downregulated
A0A6P6TY94_COFAR	Protein ACCUMULATION AND REPLICATION OF CHLOROPLASTS 3-like	downregulated
A0A6P6XC49_COFAR	Protein MET1 chloroplastic-like isoform X1	downregulated
A0A6P6VXX5_COFAR	Protein SUPPRESSOR OF QUENCHING 1 chloroplastic-like	downregulated
A0A6P6WTK0_COFAR	Protein transport protein SEC23	downregulated
A0A6P6SQZ6_COFAR	protein-serine/threonine phosphatase	downregulated
A0A6P6WE36_COFAR	Putative rRNA methyltransferase	downregulated
A0A6P6WP45_COFAR	Rhodanese-like domain-containing protein 4 chloroplastic	downregulated
A0A6P6VV71_COFAR	Ribulose bisphosphate carboxylase/oxygenase activase chloroplastic_2	downregulated
A0A6P6SIE6_COFAR	rRNA biogenesis protein RRP5-like isoform X1	downregulated
A0A6P6U6S6_COFAR	RuvB-like helicase	downregulated
A0A6P6XAA1_COFAR	Subtilisin-like protease SBT1.4	downregulated
A0A6P6VWL9_COFAR	Superoxide dismutase [Cu-Zn]	downregulated
A0A6P6UU30_COFAR	Thioredoxin Y2, chloroplastic-like	downregulated
A0A6P6XIM1_COFAR	TPR repeat-containing thioredoxin TTL1 isoform X2	downregulated

A0A6P6W7I5_COFAR	Triosephosphate isomerase cytosolic	downregulated
A0A6P6U6L6_COFAR	Tyrosine decarboxylase 1-like	downregulated
A0A6P6TP50_COFAR	UDP-sulfoquinovose synthase chloroplastic-like	downregulated
A0A6P6SJK7_COFAR	uracil phosphoribosyltransferase	downregulated

Supplementary Table 6 Classification of the statistically significant metabolic features of Catuaí (*C. arabica* cultivar) in irrigated vs. rainfed plants, according to the results of the volcano plot analysis.

Metabolic feature	Metabolite	Metabolic class	Condition
Adenine (2TMS)	Adenine	Others	upregulated
Ergosterol (1TMS)	Ergosterol	Lipids	upregulated
Fumaric acid, 2-methyl- (2TMS)	Mesaconate	Lipids	upregulated
Guanine (3TMS)	Guanine	Others	upregulated
Homoserine (3TMS)	L-Homoserine	Amino acids	upregulated
Inositol, allo- (6TMS)	Inositol	Others	upregulated
Lysine, N6-biotinyl- (5TMS)	N6-D-Biotinyl-L-lysine	Amino acids	upregulated
Malic acid, 3-isopropyl-, threo- (3TMS)	3-Isopropylmalate	Lipids	upregulated
Nonacosanoic acid (1TMS)	Nonacosanoic acid	Lipids	upregulated
Prolyl-glycine (2TMS)	Prolylglycine	Amino acids	upregulated
Propane-1,3-diol, 2-amino-2-methyl- (3TMS)	2-Amino-2-methyl-1,3-propandiol	Others	upregulated
Propanoic acid, 3-amino-3-(4-hydroxyphenyl)- (3TMS)	beta-Tyrosine	Amino acids	upregulated
Pyridoxine (3TMS)	Pyridoxine	Others	upregulated
Pyruvic acid, 4-hydroxyphenyl- (1MEOX) (2TMS) BP	4-Hydroxyphenylpyruvic acid	Other organic acids	upregulated
Quinic acid, 1-caffeoyl-, trans- (6TMS)	1-Caffeoylquinic acid	Phenylpropanoids or phenolic compounds	upregulated
Quinic acid, 3-caffeoyl-, cis- (6TMS)	Cis-5-Caffeoylquinic acid	Phenylpropanoids or phenolic compounds	upregulated
Ribonic acid-1,4-lactone (3TMS)	D-Arabinono-1,4-lactone	Sugars	upregulated
Theophylline (1TMS)	Theophylline	Alkaloids	upregulated
Tyrosine, meta- (3TMS)	Meta-Tyrosine	Amino acids	upregulated
Urea (2TMS)	Urea	Amides	upregulated
1,4-Naphthoquinone, 2-methyl-	Menadione	Others	downregulated
2-Piperidinecarboxylic acid (2TMS)	Pipecolic acid	Alkaloids	downregulated
3-Indoleacetic acid, 5-methoxy- (2TMS)	5-Methoxyindoleacetate	Others	downregulated
4-Pyridoxic acid (3TMS)	4-Pyridoxate	Others	downregulated

Acetic acid, 3,4-dihydroxyphenyl- (3TMS)	3,4-Dihydroxyphenylacetate	Phenylpropanoids or phenolic compounds	downregulated
Acetoacetic acid (1MEOX) (1TMS) BP	Acetoacetate	Lipids	downregulated
Asparagine (3TMS)	Asparagine	Amino acids	downregulated
Butanoic acid, 4-amino-3-hydroxy- (3TMS)	4-Amino-3-hydroxybutanoate	Amino acids	downregulated
Cellobiose, D- (1MEOX) (8TMS) MP	Cellobiose	Sugars	downregulated
Digitoxose (1MEOX) (3TMS) BP	D-Digitoxose	Sugars	downregulated
Digitoxose (1MEOX) (3TMS) MP	D-Digitoxose	Sugars	downregulated
Dihydrozeatin riboside (4TMS)	Dihydrozeatin riboside	Others	downregulated
Dopamine (4TMS)	Dopamine	Amines	downregulated
Erythrose-4-phosphate (1MEOX) (4TMS) MP	D-Erythrose 4-phosphate	Sugars	downregulated
Fucose (1MEOX) (4TMS) MP	2-Deoxyglucose	Sugars	downregulated
Galactonic acid-1,4-lactone (4TMS)	D-Galactono-1,4-lactone	Others	downregulated
Galactosamine, N-acetyl- (1MEOX) (5TMS) BP	N-Acetyl-D-galactosamine	Sugars	downregulated
Glucaric acid-1,4-lactone (4TMS)	D-Glucaro-1,4-lactone	Others	downregulated
Glucose, 2-amino-2-deoxy-, D- (1MEOX) (5TMS)	2-Amino-2-deoxy-D-glucose	Sugars	downregulated
Glycerol-2-phosphate (4TMS)	Glycerol 2-phosphate	Lipids	downregulated
Guanosine-2*,3*-cyclic-monophosphate (4TMS)	2',3'-Cyclic GMP	Others	downregulated
Hexadecan-1-ol, n- (1TMS)	1-Hexadecanol	Lipids	downregulated
Histidinol (4TMS)	L-Histidinol	Amines	downregulated
Hydrocaffeic acid (3TMS)	3,4-Dihydroxyphenylpropanoate	Phenylpropanoids or phenolic compounds	downregulated
Kynurenine (3TMS)	L-Kynurenine	Others	downregulated
Malic acid, 2-isopropyl- (3TMS)	alpha-Isopropylmalate	Lipids	downregulated
Maltose (1MEOX) (8TMS) BP	D-Maltose	Sugars	downregulated
Maltotriose (1MEOX) (11TMS) MP	Maltotriose	Sugars	downregulated
myo-Inositol-1-phosphate (7TMS)	myo-Inositol 1-phosphate	Others	downregulated
Octopamine (4TMS)	Octopamine	Other organic acids	downregulated
Pantothenic acid, D- (4TMS)	Pantothenate	Others	downregulated

Phenethylamine (2TMS)	Phenethylamine	Amines	downregulated
Phylloquinone (2TMS)	Phylloquinone	Others	downregulated
Piceatannol (4TMS) BP	Piceatannol	Phenylpropanoids or phenolic compounds	downregulated
Proline, 4-hydroxy-, trans- (3TMS)	trans-4-Hydroxy-L-proline	Amino acids	downregulated
Salicylic acid (2TMS)	Salicylic acid	Phenylpropanoids or phenolic compounds	downregulated
Shikimic acid, 3-dehydro- (1MEOX) (3TMS) MP	3-Dehydroshikimate	Phenylpropanoids or phenolic compounds	downregulated
Sitosterol, beta- (1TMS)	Beta-Sitosterol	Lipids	downregulated
Sophorose (1MEOX) (8TMS) BP	Sophorose	Sugars	downregulated
Sophorose (1MEOX) (8TMS) MP	Sophorose	Sugars	downregulated
Sphingosine, O-1-beta-galactosyl- (6TMS)	Glucosylsphingosine	Lipids	downregulated
Squalene, all-trans-	Squalene	Lipids	downregulated
Taxifolin (1MEOX) (5TMS) MP	Taxifolin	Phenylpropanoids or phenolic compounds	downregulated
Taxifolin (5TMS)	Taxifolin	Phenylpropanoids or phenolic compounds	downregulated
Threonic acid (4TMS)	Threonate	Sugars	downregulated
Tyrosine, N-acetyl- (3TMS)	N-Acetyltyrosine	Amino acids	downregulated

Supplementary Table 7 Classification of statistically significant proteins of Paraíso 2 (*C. arabica* cultivar) in irrigated vs. rainfed plants, according to the results of the volcano plot analysis.

Uniprot code	Protein name	Condition
A0A6P6T2L6_COFAR	serine O-acetyltransferase	upregulated
A0A6P6UK35_COFAR	CO(2)-response secreted protease-like	upregulated
A0A6P6SN62_COFAR	Pathogen-associated molecular patterns-induced protein A70-like	upregulated
A0A6P6S6L4_COFAR	Glyceraldehyde-3-phosphate dehydrogenase_2	upregulated
A0A6P6V5L2_COFAR	Acidic endochitinase-like_5	upregulated
A0A6P6VSY4_COFAR	Heat shock 70 kDa protein mitochondrial	upregulated
A0A6P6UK18_COFAR	Phosphatidylinositol/phosphatidylcholine transfer protein SFH6-like isoform X3	upregulated
A0A6P6XC29_COFAR	Calcium-binding allergen Ole e 8-like	upregulated
A0A6P6U007_COFAR	Conserved oligomeric Golgi complex subunit 6	upregulated
A0A6P6TAF6_COFAR	Caltractin-like	upregulated
A0A6P6TKD8_COFAR	Calmodulin-lysine N-methyltransferase	upregulated
A0A6P6UBZ5_COFAR	non-specific serine/threonine protein kinase_1	upregulated
A0A6P6SS35_COFAR	Flavin-containing monooxygenase	upregulated
A0A6P6S4Q8_COFAR	peptidylprolyl isomerase_4	upregulated
A0A6P6ULI3_COFAR	Abscisic stress-ripening protein 2-like	upregulated
A0A6P6TUL0_COFAR	V-type proton ATPase subunit E	upregulated
A0A6P6UCL3_COFAR A0A096VHZ8_9GENT	Probable caffeine synthase 3	upregulated
A0A6P6VPJ7_COFAR A0A6P6W4L6_COFAR	Remorin-like	upregulated
A0A6P6V280_COFAR	Chloroplast stem-loop binding protein of 41 kDa a chloroplastic-like	upregulated
A0A6P6SM86_COFAR	Dicer-like protein 4 isoform X5	upregulated
A0A6P6SBP5_COFAR	Anthocyanidin reductase	upregulated
A0A6P6T5Q0_COFAR	Granule-bound starch synthase 2 chloroplastic/amyloplastic	upregulated
A0A6P6V1H5_COFAR	Acidic endochitinase-like_2	upregulated
A0A6P6SLW8_COFAR	Transmembrane 9 superfamily member	upregulated
A0A6P6U2Y7_COFAR	Probable aquaporin TIP1-1	upregulated
A0A6P6XGU5_COFAR	Calmodulin	upregulated
A0A6P6WI95_COFAR	Pleiotropic drug resistance protein 3-like isoform X6	upregulated
A0A6P6V9A1_COFAR	Protein trichome birefringence-like 6	upregulated
A0A6P6VWN1_COFAR	pyridoxal 5'-phosphate synthase (glutamine hydrolyzing)	upregulated
A0A6P6WNF9_COFAR	Auxin-induced protein PCNT115-like isoform X2	upregulated
A0A6P6U6Z5_COFAR	RING-type E3 ubiquitin transferase_2	upregulated
A0A1W6CCF1_COFAR	Photosystem I iron-sulfur center	upregulated
A0A6P6V2B1_COFAR	peptidylprolyl isomerase_1	upregulated

A0A6P6TAR2_COFAR	peptidylprolyl isomerase_3	upregulated
A0A6P6TYE3_COFAR	Glutathione peroxidase_1	upregulated
A0A6P6WNB0_COFAR	Glycine-rich cell wall structural protein-like_1	upregulated
A0A6P6TLX4_COFAR	Non-specific lipid-transfer protein 2-like	upregulated
A0A6P6TRS3_COFAR	Enhancer of polycomb-like protein	upregulated
A0A6P6UQK6_COFAR	Photosystem II 22 kDa protein chloroplastic-like	upregulated
A0A6P6TEJ5_COFAR	Carbonic anhydrase	downregulated
A0A6P6VV67_COFAR	transketolase	downregulated
A0A6P6S5I2_COFAR	Disease resistance protein RGA3	downregulated
A0A6P6VSZ6_COFAR	Oxygen-evolving enhancer protein 3-2 chloroplastic-like	downregulated
A0A6P6S8U9_COFAR	Heat shock cognate 70 kDa protein-like isoform X1	downregulated
A0A6P6TR19_COFAR	Elongation factor 2-like	downregulated
A0A6P6WP45_COFAR	Rhodanese-like domain-containing protein 4 chloroplastic	downregulated
A0A6P6SQ02_COFAR	Ricin B-like lectin R40G3	downregulated
A0A6P6UHS3_COFAR	ATP-dependent zinc metalloprotease FTSH 9 chloroplastic isoform X1	downregulated
A0A6P6SD20_COFAR	Chlorophyll a-b binding protein chloroplastic_5	downregulated
A0A6P6UEC1_COFAR	Chlorophyll a-b binding protein chloroplastic_3	downregulated
A0A2Z5WP99_COFAR Q31949_COFAR A0A8F2J190_COFAR R4I7C6_COFAR	Ribulose bisphosphate carboxylase large chain (Fragment)	downregulated
A0A6P6TBH4_COFAR	Probable transcription-associated protein 1	downregulated
A0A6P6TDV9_COFAR	Protein CROWDED NUCLEI 4-like	downregulated
A0A6P6UMW7_COFAR	Serpin-ZX-like	downregulated
A0A6P6WR50_COFAR	chitinase_1	downregulated
K7PAI2_COFAR	ATP synthase subunit alpha (Fragment)	downregulated
A0A6P6SQ38_COFAR	Macrophage migration inhibitory factor homolog isoform X1	downregulated
A0A6P6VGL8_COFAR	Eukaryotic translation initiation factor 5B-like	downregulated
A0A1W6CCG6_COFAR	Photosystem II D2 protein	downregulated
A0A6P6V4E0_COFAR	Chlorophyll a-b binding protein chloroplastic_7	downregulated
A0A6P6XIM1_COFAR	TPR repeat-containing thioredoxin TTL1 isoform X2	downregulated
A0A6P6UNT8_COFAR	1,4-alpha-glucan branching enzyme	downregulated
A0A6P6SGL2_COFAR	Heme-binding protein 2-like	downregulated
A0A6P6VX74_COFAR A0A6P6T3A8_COFAR	Histone H2A_2	downregulated
A0A6P6S6N3_COFAR	Protein THYLAKOID FORMATION1 chloroplastic	downregulated
A0A6P6UJW0_COFAR	LRR receptor-like serine/threonine-protein kinase RCH1	downregulated
A0A6P6UGW4_COFAR A0A6P6SGA0_COFAR	Patatin	downregulated
A0A6P6WLA7_COFAR	Glutaredoxin-dependent peroxiredoxin_2	downregulated

A0A6P6WLH1_COFAR A0A6P6VWH3_COFAR	Cysteine synthase_1	downregulated
A0A6P6W4U9_COFAR	Cytochrome b5	downregulated
A0A6P6U6L6_COFAR	Tyrosine decarboxylase 1-like	downregulated
A0A6P6U1Q5_COFAR	Glycine cleavage system P protein	downregulated
A0A6P6WQV5_COFAR	Cysteine synthase_2	downregulated
A0A6P6TYA7_COFAR	Aspartyl protease family protein At5g10770-like	downregulated
A0A6P6WKM3_COFAR	isoleucine--tRNA ligase	downregulated
A0A6P6SM02_COFAR	14 kDa zinc-binding protein-like isoform X1	downregulated
A0A6P6UI56_COFAR	ATP-dependent DNA helicase_1	downregulated
A0A6P6ULL6_COFAR	Plastocyanin	downregulated
A0A6P6SDY0_COFAR A0A6P6SD75_COFAR A0A6P6URI6_COFAR	chitinase_2	downregulated
A0A6P6UD77_COFAR A0A6P6S8X4_COFAR	Photosystem I reaction center subunit III	downregulated
A0A6P6SR15_COFAR	LOW QUALITY PROTEIN: ATP-dependent zinc metalloprotease FTSH 8 mitochondrial-like	downregulated
A0A6P6V602_COFAR	Aldehyde dehydrogenase family 2 member B7 mitochondrial-like	downregulated
A0A6P6V011_COFAR	Phosphoglycerate kinase_2	downregulated
A0A6P6UBY8_COFAR	Thioredoxin O2, mitochondrial-like	downregulated
A0A6P6SHB8_COFAR	Ribosomal protein_2	downregulated
A0A6P6UQT6_COFAR	Probable GTP-binding protein OBGM mitochondrial isoform X1	downregulated
A0A6P6SZK1_COFAR	Blue copper protein-like	downregulated
A0A6P6UH18_COFAR	ATP synthase delta chain chloroplastic-like	downregulated
A0A6P6UHZ3_COFAR	Monooxygenase 2-like	downregulated
A0A6P6SGR8_COFAR	20 kDa chaperonin chloroplastic	downregulated
A0A6P6SBQ7_COFAR	Photosystem II repair protein PSB27-H1 chloroplastic-like	downregulated
A0A6P6SLT0_COFAR	thioredoxin-dependent peroxiredoxin	downregulated
A0A6P6T4J5_COFAR	Peroxidase_1	downregulated
A0A6P6TF28_COFAR	UDP-arabinopyranose mutase_2	downregulated
A0A6P6TMJ4_COFAR	glutathione transferase_2	downregulated
A0A6P6S4T8_COFAR	ATP-dependent zinc metalloprotease FTSH 2 chloroplastic	downregulated
A0A6P6TY02_COFAR	Amine oxidase_2	downregulated
A0A6P6TM83_COFAR	COP1-interactive protein 1-like	downregulated
A0A6P6TM10_COFAR	alanine--glyoxylate transaminase	downregulated
A0A6P6VYA3_COFAR	Germin-like protein_2	downregulated
A0A6P6W748_COFAR	Sedoheptulose-1,7-bisphosphatase chloroplastic	downregulated
A0A6P6SDN8_COFAR	Acidic endochitinase-like_3	downregulated
A0A6P6UZK2_COFAR	MLP-like protein 34	downregulated
A0A6P6VFW0_COFAR	peptidylprolyl isomerase_6	downregulated

A0A6P6SIE6_COFAR	rRNA biogenesis protein RRP5-like isoform X1	downregulated
A0A6P6WKV7_COFAR	Acyl carrier protein_2	downregulated
A0A6P6X1U0_COFAR	non-specific serine/threonine protein kinase_2	downregulated
A0A6P6UPQ6_COFAR A0A6P6V2M1_COFAR A0A6P6VC84_COFAR	Glyceraldehyde-3-phosphate dehydrogenase_3	downregulated
A0A6P6SI69_COFAR	L-dopachrome isomerase	downregulated
A0A6P6UDD2_COFAR	Cytochrome P450 97B2 chloroplastic-like	downregulated
A0A6P6VXX5_COFAR	Protein SUPPRESSOR OF QUENCHING 1 chloroplastic-like	downregulated
A0A6P6X4H3_COFAR	Zinc finger protein BRUTUS-like At1g74770	downregulated
A0A6P6VD82_COFAR	Cytochrome b561 domain-containing protein At4g18260-like isoform X1	downregulated
A0A6P6SSC1_COFAR	FT-interacting protein 1-like	downregulated
A0A6P6TYG4_COFAR	Probable steroid-binding protein 3	downregulated
A0A6P6SS20_COFAR	Polyadenylate-binding protein 5-like isoform X1	downregulated
A0A6P6V142_COFAR	Heat shock protein 90-5 chloroplastic	downregulated
A0A6P6VZU1_COFAR	DNA damage-repair/toleration protein DRT102-like	downregulated
A0A6P6UU30_COFAR	Thioredoxin Y2, chloroplastic-like	downregulated
A0A6P6S8J9_COFAR	30S ribosomal protein S10 chloroplastic-like isoform X2	downregulated
A0A6P6VK88_COFAR	Late embryogenesis abundant protein At5g17165-like	downregulated
A0A6P6W7I5_COFAR	Triosephosphate isomerase cytosolic	downregulated
A0A6P6X359_COFAR	Vinorine synthase-like	downregulated
A0A6P6UJK4_COFAR	glutathione transferase_4	downregulated
A0A6P6SSZ2_COFAR	EG45-like domain containing protein	downregulated
A0A6P6V351_COFAR A0A6P6W4C1_COFAR A0A6P6V5E4_COFAR	Kirola-like	downregulated
A0A6P6SC70_COFAR	Acidic endochitinase-like_1	downregulated
A0A6P6W8S4_COFAR	Probable calcium-binding protein CML13	downregulated
A0A6P6VNJ3_COFAR	ATP synthase subunit delta' mitochondrial-like	downregulated
A0A1W6CCE4_COFAR	ATP synthase CF0 B subunit	downregulated
A0A6P6XIP3_COFAR	ATP-dependent Clp protease ATP-binding subunit ClpA homolog CD4B chloroplastic	downregulated
A0A6P6SQZ6_COFAR	protein-serine/threonine phosphatase	downregulated
A0A6P6TNW5_COFAR	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	downregulated
A0A6P6TP50_COFAR	UDP-sulfoquinovose synthase chloroplastic-like	downregulated
A0A6P6SGT0_COFAR	Ketol-acid reductoisomerase	downregulated
A0A6P6V6U4_COFAR	glutathione transferase_3	downregulated
A0A6P6WE36_COFAR	Putative rRNA methyltransferase	downregulated
A0A6P6ST40_COFAR	Pentatricopeptide repeat-containing protein At4g31070 mitochondrial-like	downregulated

A0A6P6VSV4_COFAR	Malate dehydrogenase_4	downregulated
A0A6P6VXI3_COFAR	NAD(P)H dehydrogenase (quinone)	downregulated
A0A6P6VST3_COFAR H2EQR2_9GENT	L-ascorbate peroxidase 3-like	downregulated
A0A6P6VRR9_COFAR	Glycine-rich RNA-binding abscisic acid-inducible protein-like	downregulated
A0A6P6VUZ1_COFAR	Actin-7-like	downregulated
A0A6P6SL46_COFAR	Disease resistance RPP13-like protein 1	downregulated
A0A6P6VEX3_COFAR	Triosephosphate isomerase, cytosolic-like	downregulated
A0A6P6VUZ0_COFAR	Alpha-N-acetylglucosaminidase-like isoform X2	downregulated
A0A6P6SYE0_COFAR	Histone H2B	downregulated
A0A6P6XC49_COFAR	Protein MET1 chloroplastic-like isoform X1	downregulated
A0A6P6ULR3_COFAR	Endo-1, 31, 4-beta-D-glucanase-like isoform X2	downregulated
A0A6P6VSN6_COFAR	Cysteine proteinase inhibitor 4-like	downregulated
A0A6P6XNL1_COFAR	Fructose-bisphosphate aldolase_2	downregulated
A0A1W6CCG4_COFAR	Cytochrome b6	downregulated
A0A6P6U1X1_COFAR	Delta-aminolevulinic acid dehydratase	downregulated
A0A6P6WHU3_COFAR	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase	downregulated
A0A6P6UIQ0_COFAR	ABC transporter G family member 25-like	downregulated
A0A6P6U9B0_COFAR	Chlorophyll a-b binding protein chloroplastic_1	downregulated
A0A6P6WV53_COFAR	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 6-like	downregulated
A0A6P6XCE1_COFAR	ribulose-phosphate 3-epimerase	downregulated
A0A6P6TSN2_COFAR	IQ domain-containing protein IQM1-like	downregulated
A0A6P6VI58_COFAR	Pentatricopeptide repeat-containing protein At3g46790, chloroplastic-like	downregulated
A0A6P6ULZ1_COFAR	alcohol dehydrogenase	downregulated
A0A6P6T6B6_COFAR	PLAT domain-containing protein 3-like_1	downregulated
A0A6P6TY94_COFAR	Protein ACCUMULATION AND REPLICATION OF CHLOROPLASTS 3-like	downregulated
A0A6P6V0C6_COFAR	Chlorophyllase-1-like	downregulated
A0A6P6WEB0_COFAR	Malate dehydrogenase_3	downregulated
A0A6P6V3F1_COFAR	Mitochondrial outer membrane protein porin of 36 kDa-like	downregulated
A0A6P6U6S6_COFAR	RuvB-like helicase	downregulated
A0A6P6SI29_COFAR	Cysteine proteinase 3-like	downregulated
A0A6P6ULJ0_COFAR	L-ascorbate oxidase homolog	downregulated
A0A6P6VEG9_COFAR	Non-symbiotic hemoglobin 2-like	downregulated
A0A6P6VSB3_COFAR	Presequence protease 1 chloroplastic/mitochondrial-like	downregulated
A0A6P6S579_COFAR	14-3-3 protein 9 isoform X1	downregulated
A0A6P6X7C7_COFAR	Elongation factor 1-delta-like	downregulated
A0A6P6U5G9_COFAR	Light-induced protein chloroplastic-like	downregulated

A0A6P6U4J4_COFAR	ribose-5-phosphate isomerase	downregulated
A0A6P6V3F6_COFAR	Major allergen Pru ar 1-like_2	downregulated
A0A6P6SMZ0_COFAR	Carboxypeptidase	downregulated
A0A6P6VSR5_COFAR	GDSL esterase/lipase At1g29670-like isoform X1	downregulated
A0A6P6U1E3_COFAR A0A6P6SX37_COFAR	60S acidic ribosomal protein P2A-like	downregulated
A0A6P6SQN5_COFAR	Cell wall hydroxyproline-rich glycoprotein	downregulated

Supplementary Table 8 Classification of the statistically significant metabolic features of Paraíso 2 (*C. arabica* cultivar) in irrigated vs. rainfed plants, according to the results of the volcano plot analysis.

Metabolic feature	Metabolite	Metabolic class	Condition
Digitoxose (1MEOX) (3TMS) MP	D-Digitoxose	Sugars	upregulated
Dopamine (4TMS)	Dopamine	Amines	upregulated
Ferulic acid, trans- (2TMS)	Ferulate	Phenylpropanoids or phenolic compounds	upregulated
Fumaric acid, 2-methyl- (2TMS)	Mesaconate	Lipids	upregulated
Galactonic acid (6TMS)	D-Galactonate	Other organic acids	upregulated
Galactose (1MEOX) (5TMS) BP	D-Galactose	Sugars	upregulated
Glutaric acid, 2-oxo- (1MEOX) (2TMS) MP	2-Oxoglutarate	Other organic acids	upregulated
Glyceric acid (3TMS)	L-Glycerate	Sugars	upregulated
Malic acid, 2-methyl- (3TMS)	D-Citramalate	Lipids	upregulated
Malic acid, 3-isopropyl-, threo- (3TMS)	3-Isopropylmalate	Lipids	upregulated
Oxalic acid (2TMS)	Oxalate	Other organic acids	upregulated
Oxaloacetate (1MEOX) (3TMS) MP	Oxaloacetate	Other organic acids	upregulated
Prolyl-glycine (2TMS)	Prolylglycine	Amino acids	upregulated
Quinic acid, 3-caffeoyl-, cis- (6TMS)	Cis-5-Caffeoylquinic acid	Phenylpropanoids or phenolic compounds	upregulated
Tetracosane, n-	Lignocerane	Others	upregulated
1,4-Naphthoquinone, 2-methyl-	Menadione	Others	downregulated
3-Indoleacetic acid, 5-methoxy- (2TMS)	5-Methoxyindoleacetate	Others	downregulated
Acetic acid, 3-hydroxyphenyl- (3TMS)	3-Hydroxyphenylacetate	Other organic acids	downregulated
Asparagine (3TMS)	Asparagine	Amino acids	downregulated
Butanoic acid, 4-amino-3-hydroxy- (3TMS)	4-Amino-3-hydroxybutanoate	Amino acids	downregulated
Cellobiose, D- (1MEOX) (8TMS) MP	Cellobiose	Sugars	downregulated
Digitoxose (1MEOX) (3TMS) BP	D-Digitoxose	Sugars	downregulated
Dihydrozeatin riboside (4TMS)	Dihydrozeatin riboside	Others	downregulated
Galactonic acid-1,4-lactone (4TMS)	D-Galactono-1,4-lactone	Others	downregulated
Glucose (1MEOX) (5TMS) BP	D-Glucose	Sugars	downregulated
Glucose, 2-amino-2-deoxy-, D- (1MEOX) (5TMS)	2-Amino-2-deoxy-D-glucose	Sugars	downregulated

Glycerol-2-phosphate (4TMS)	Glycerol 2-phosphate	Lipids	downregulated
Glycerol-3-phosphate (4TMS)	Glycerol-3-phosphate	Lipids	downregulated
Guanosine-2*,3*-cyclic-monophosphate (4TMS)	2',3'-Cyclic GMP	Others	downregulated
Indole-3-acetic acid, 5-hydroxy-, 1H-(2TMS)	5-Hydroxyindoleacetate	Others	downregulated
Isoleucine (2TMS)	L-Isoleucine	Amino acids	downregulated
Kynurenine (3TMS)	L-Kynurenine	Others	downregulated
Leucine (2TMS)	L-Leucine	Amino acids	downregulated
Maltose (1MEOX) (8TMS) BP	D-Maltose	Sugars	downregulated
Maltotriose (1MEOX) (11TMS) MP	Maltotriose	Sugars	downregulated
myo-Inositol-1-phosphate (7TMS)	myo-Inositol 1-phosphate	Others	downregulated
Orotic acid, 4,5-dihydro- (4TMS)	L-Dihydroorotate	Others	downregulated
Phenylalanine (2TMS)	L-Phenylalanine	Amino acids	downregulated
Phylloquinone (2TMS)	Phylloquinone	Others	downregulated
Piceatannol (4TMS) BP	Piceatannol	Phenylpropanoids or phenolic compounds	downregulated
Proline (2TMS)	L-Proline	Amino acids	downregulated
Pyruvic acid, 4-hydroxyphenyl-(1MEOX) (2TMS) BP	4-Hydroxyphenylpyruvate	Phenylpropanoids or phenolic compounds	downregulated
Salicylic acid (2TMS)	Salicylic acid	Phenylpropanoids or phenolic compounds	downregulated
Secologanin (TMS) MP	Secologanin	Lipids	downregulated
Senecionine (1TMS) BP2	Senecionine	Alkaloids	downregulated
Shikimic acid, 3-dehydro- (1MEOX) (3TMS) MP	3-Dehydroshikimate	Phenylpropanoids or phenolic compounds	downregulated
Sitosterol, beta- (1TMS)	Beta-Sitosterol	Lipids	downregulated
Spermine (5TMS) MP	Spermine	Amines	downregulated
Squalene, all-trans-	Squalene	Lipids	downregulated
Tetratriacontane	Tetratriacontane	Others	downregulated
Tryptophan (2TMS)_1	Tryptophan	Amino acids	downregulated
Tryptophan (2TMS)_2	Tryptophan	Amino acids	downregulated
Tyrosine (3TMS)	L-Tyrosine	Amino acids	downregulated
Tyrosine, N-acetyl- (3TMS)	N-Acetyltyrosine	Amino acids	downregulated
Uracil, dihydro- (2TMS)	Dihydrouracil	Others	downregulated
Valine (2TMS)	L-Valine	Amino acids	downregulated

Supplementary Table 9 Classification of statistically significant proteins of Sarchimor (*C. arabica* cultivar) in irrigated vs. rainfed plants, according to the results of the volcano plot analysis.

Uniprot code	Protein name	Condition
A0A6P6SM02_COFAR	14 kDa zinc-binding protein-like isoform X1	upregulated
A0A6P6SZ87_COFAR	29 kDa ribonucleoprotein A chloroplastic	upregulated
A0A6P6S8J9_COFAR	30S ribosomal protein S10 chloroplastic-like isoform X2	upregulated
A0A6P6TUF9_COFAR	30S ribosomal protein S6 alpha, chloroplastic isoform X1	upregulated
A0A6P6SLC1_COFAR	36.4 kDa proline-rich protein-like	upregulated
A0A6P6UIQ0_COFAR	ABC transporter G family member 25-like	upregulated
A0A6P6VUZ1_COFAR	Actin-7-like	upregulated
A0A6P6VY99_COFAR	ADP/ATP translocase	upregulated
A0A6P6TM10_COFAR	alanine--glyoxylate transaminase	upregulated
A0A6P6UH18_COFAR	ATP synthase delta chain chloroplastic-like	upregulated
A0A6P6VNJ3_COFAR	ATP synthase subunit delta' mitochondrial-like	upregulated
A0A6P6WN40_COFAR	Beta-galactosidase_1	upregulated
A0A6P6TEJ5_COFAR	Carbonic anhydrase	upregulated
A0A6P6U9B0_COFAR	Chlorophyll a-b binding protein chloroplastic_1	upregulated
A0A6P6SKT1_COFAR	Chlorophyll a-b binding protein chloroplastic_2	upregulated
A0A6P6SI29_COFAR	Cysteine proteinase 3-like	upregulated
A0A6P6VSN6_COFAR	Cysteine proteinase inhibitor 4-like	upregulated
A0A6P6WQV5_COFAR	Cysteine synthase_2	upregulated
A0A1W6CCG4_COFAR	Cytochrome b6	upregulated
A0A6P6T7E4_COFAR	Ferredoxin--NADP reductase chloroplastic	upregulated
A0A6P6SMW1_COFAR	Glucan endo-1,3-beta-glucosidase acidic-like_2	upregulated
A0A6P6UJK4_COFAR	glutathione transferase_4	upregulated
A0A6P6WWZ4_COFAR	Glycine-rich cell wall structural protein-like_2	upregulated
A0A6P6V142_COFAR	Heat shock protein 90-5 chloroplastic	upregulated
A0A6P6SYE0_COFAR	Histone H2B	upregulated
A0A6P6VK88_COFAR	Late embryogenesis abundant protein At5g17165-like	upregulated
A0A6P6V3F6_COFAR	Major allergen Pru ar 1-like_2	upregulated
A0A6P6VQX4_COFAR	Malate dehydrogenase [NADP] chloroplastic	upregulated

A0A6P6V3F1_COFAR	Mitochondrial outer membrane protein porin of 36 kDa-like	upregulated
A0A6P6SN62_COFAR	Pathogen-associated molecular patterns-induced protein A70-like	upregulated
A0A6P6V011_COFAR	Phosphoglycerate kinase_2	upregulated
A0A1W6CCA3_COFAR	Photosystem II reaction center protein H	upregulated
A0A6P6T6B6_COFAR	PLAT domain-containing protein 3-like_1	upregulated
A0A6P6VSB3_COFAR	Presequence protease 1 chloroplastic/mitochondrial-like	upregulated
A0A6P6UKU7_COFAR	Probable cysteine protease RD19C	upregulated
A0A6P6SND7_COFAR	Proteasome subunit alpha type	upregulated
A0A6P6V1H8_COFAR	Protein SIEVE ELEMENT OCCLUSION B-like	upregulated
A0A6P6SHB8_COFAR	Ribosomal protein_2	upregulated
A0A6P6VV71_COFAR	Ribulose bisphosphate carboxylase/oxygenase activase chloroplastic_2	upregulated
A0A6P6XCE1_COFAR	ribulose-phosphate 3-epimerase	upregulated
A0A6P6W7I5_COFAR	Triosephosphate isomerase cytosolic	upregulated
A0A6P6VEX3_COFAR	Triosephosphate isomerase, cytosolic-like	upregulated
A0A6P6SC70_COFAR	Acidic endochitinase-like_1	downregulated
A0A6P6V1H5_COFAR	Acidic endochitinase-like_2	downregulated
A0A6P6TY02_COFAR	Amine oxidase_2	downregulated
A0A6P6S4T8_COFAR	ATP-dependent zinc metalloprotease FTSH 2 chloroplastic	downregulated
A0A6P6WNF9_COFAR	Auxin-induced protein PCNT115-like isoform X2	downregulated
A0A6P6SZK1_COFAR	Blue copper protein-like	downregulated
A0A6P6TKD8_COFAR	Calmodulin-lysine N-methyltransferase	downregulated
A0A6P6W7A5_COFAR A0A6P6UPG2_COFAR	CDGSH iron-sulfur domain-containing protein NEET-like isoform X1	downregulated
A0A6P6SQN5_COFAR	Cell wall hydroxyproline-rich glycoprotein	downregulated
A7IZK7_COFCA	Cell-wall invertase	downregulated
A0A6P6SDY0_COFAR A0A6P6SD75_COFAR A0A6P6URI6_COFAR	chitinase_2	downregulated
A0A6P6T7D3_COFAR	Citrate-binding protein-like	downregulated
A0A6P6VD82_COFAR	Cytochrome b561 domain-containing protein At4g18260-like isoform X1	downregulated
A0A6P6WQ90_COFAR	Disease resistance protein RGA3 isoform X1	downregulated
A0A6P6TRS3_COFAR	Enhancer of polycomb-like protein	downregulated

A0A6P6U015_COFAR	Glycine cleavage system H protein_1	downregulated
A0A6P6SGL2_COFAR	Heme-binding protein 2-like	downregulated
A0A6P6WEB0_COFAR	Malate dehydrogenase_3	downregulated
A0A6P6WV53_COFAR	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 6-like	downregulated
A0A6P6UBZ5_COFAR	non-specific serine/threonine protein kinase_1	downregulated
A0A6P6V2B1_COFAR	peptidylprolyl isomerase_1	downregulated
A0A6P6TAR2_COFAR	peptidylprolyl isomerase_3	downregulated
A0A6P6S4Q8_COFAR	peptidylprolyl isomerase_4	downregulated
A0A6P6TCJ6_COFAR	Peroxidase_2	downregulated
A0A6P6SSQ7_COFAR	Peroxiredoxin Q chloroplastic	downregulated
A0A1W6CCF1_COFAR	Photosystem I iron-sulfur center	downregulated
A0A6P6SS20_COFAR	Polyadenylate-binding protein 5-like isoform X1	downregulated
A0A6P6UCL3_COFAR A0A096VHZ8_9GENT	Probable caffeine synthase 3	downregulated
A0A6P6TYG4_COFAR	Probable steroid-binding protein 3	downregulated
A0A6P6VSB9_COFAR A0A6P6VTW1_COFAR	Prostaglandin reductase-3-like isoform X1	downregulated
A0A6P6SQZ6_COFAR	protein-serine/threonine phosphatase	downregulated
A0A6P6VPJ7_COFAR A0A6P6W4L6_COFAR	Remorin-like	downregulated
A0A6P6W2L2_COFAR	RING-type E3 ubiquitin transferase_1	downregulated
A0A6P6U6Z5_COFAR	RING-type E3 ubiquitin transferase_2	downregulated
A0A6P6XAA1_COFAR	Subtilisin-like protease SBT1.4	downregulated
A0A6P6SJK7_COFAR	uracil phosphoribosyltransferase	downregulated

Supplementary Table 10 Classification of the statistically significant metabolic features of Sarchimor (*C. arabica* cultivar) in irrigated vs. rainfed plants, according to the results of the volcano plot analysis.

Metabolic feature	Metabolite	Metabolic class	Condition
Aconitic acid, cis- (3TMS)	cis-Aconitate	Other organic acids	upregulated
Butanoic acid, 4-amino-3-hydroxy- (3TMS)	4-Amino-3-hydroxybutanoate	Amino acids	upregulated
Fumaric acid, 2-methyl- (2TMS)	Mesaconate	Lipids	upregulated
Gibberellin A4 (2TMS)	Gibberellin A4	Lipids	upregulated
Inosine-5*-monophosphate (5TMS)	Inosinic acid	Others	upregulated
Phenethylamine (2TMS)	Phenethylamine	Amines	upregulated
Pyruvic acid, 4-hydroxyphenyl- (1MEOX) (3TMS) MP	4-Hydroxyphenylpyruvate	Phenylpropanoids or phenolic compounds	upregulated
Quinic acid, 3-caffeoyl-, cis- (6TMS)	Cis-5-Caffeoylquinic acid	Phenylpropanoids or phenolic compounds	upregulated
Squalene, all-trans-	Squalene	Lipids	upregulated
Cellobiose, D- (1MEOX) (8TMS) MP	Cellobiose	Sugars	downregulated
Citric acid (4TMS)	Citrate	Other organic acids	downregulated
Coumarine	Coumarin	Phenylpropanoids or phenolic compounds	downregulated
Cystine (4TMS)	Cystine	Amino acids	downregulated
Galactonic acid-1,4-lactone (4TMS)	D-Galactono-1,4-lactone	Others	downregulated
Glucoheptose (1MEOX) (6TMS) MP	d-Glycero-d-galacto-heptose	Sugars	downregulated
Glucose, 2-amino-2-deoxy- (5TMS) BP	2-Amino-2-deoxy-D-glucose	Sugars	downregulated
Glutamic acid, N-acetyl- (3TMS)	N-Acetyl-L-glutamate	Amino acids	downregulated
Hippuric acid (2TMS)	Hippurate	Other organic acids	downregulated
Inositol, allo- (6TMS)	Inositol	Others	downregulated
Kynurenine (3TMS)	L-Kynurenine	Others	downregulated
Levulinic acid, 5-amino- (1MEOX) (3TMS) MP	5-Aminolevulinate	Lipids	downregulated
Lipoamide, alpha- (2TMS)	Lipoamide	Amides	downregulated
Maltose (1MEOX) (8TMS) BP	D-Maltose	Sugars	downregulated

Melibiose (1MEOX) (8TMS) MP	Melibiose	Sugars	downregulated
Ornithine, N2-acetyl- (3TMS)	N2-Acetylornithine	Amino acids	downregulated
Orotic acid, 4,5-dihydro- (4TMS)	L-Dihydroorotate	Others	downregulated
Oxaloacetate (1MEOX) (3TMS) MP	Oxaloacetate	Other organic acids	downregulated
Phenylalanine (2TMS)	L-Phenylalanine	Amino acids	downregulated
Phenylalanine, N-acetyl- (2TMS)	N-Acetyl-L-phenylalanine	Amino acids	downregulated
Pyridoxine (3TMS)	Pyridoxine	Others	downregulated
Ribonic acid-1,4-lactone (3TMS)	D-Arabinono-1,4-lactone	Sugars	downregulated
Saccharopine (5TMS)	L-Saccharopine	Amino acids	downregulated
Sophorose (1MEOX) (8TMS) MP	Sophorose	Sugars	downregulated
Tryptophan (2TMS)_1	Tryptophan	Amino acids	downregulated
Tryptophan (2TMS)_2	Tryptophan	Amino acids	downregulated

TERCEIRA PARTE

CONSIDERAÇÕES FINAIS

Nesta tese, foram apresentados estudos sobre a plasticidade fenotípica de cultivares de *Coffea arabica* L., com potencial para resistência a estresses e para a obtenção de elevada produtividade e qualidade da bebida. A análise anatômica e outras características foliares, como os índices espectrais das cinco cultivares selecionadas de *C. arabica*, revelaram menos variações em comparação com as características bioquímicas das folhas. Ademais, a variação no metaboloma, no proteoma e na atividade de invertases pode estar relacionada ao estado hídrico foliar no momento da coleta.

Nos dois locais experimentais selecionados na região do Cerrado Mineiro para a condução deste estudo, as plantas foram cultivadas em regime de sequeiro (em Patrocínio) e com irrigação (em Monte Carmelo). Apesar da presença de irrigação, as plantas cultivadas em Monte Carmelo apresentaram menor potencial hídrico. Embora os locais apresentem baixa variabilidade climática, sendo as maiores diferenças observadas na velocidade do vento e na umidade relativa do ar, os dados obtidos por meio do modelo ecometabolômico indicaram que o clima, mais do que o solo, exerceu maior influência sobre as alterações no metaboloma das cultivares de *C. arabica*.

A caracterização do metaboloma e do proteoma das cultivares possibilitou a identificação das principais classes metabólicas e dos processos biológicos mais relevantes nos dois ambientes experimentais. Verificou-se, inclusive, que vias metabólicas relacionadas à resposta ao estresse foram diferencialmente expressas entre as cultivares, com destaque para Catiguá MG2 e Paraíso 2. No entanto, entre essas duas cultivares, apenas Paraíso 2 apresentou menor variação nas características agronômicas, mantendo elevados índices de produtividade e qualidade da bebida em ambos os locais experimentais. Metabólitos e vias metabólicas possivelmente associados à produtividade e à qualidade da bebida dessa cultivar também foram identificados e discutidos nesta tese. Embora nenhuma doença comum do cafeeiro tenha sido detectada nas folhas coletadas, é importante considerar a possível ativação de vias metabólicas relacionadas à resposta ao estresse biótico nas cultivares avaliadas.

Ressalta-se que esta é a primeira vez que o metaboloma e o proteoma de cultivares de cafeeiro são explorados em condições de campo, o que proporciona uma representação mais fiel da realidade enfrentada pelos produtores de café. Além disso, o protocolo de extração integrada de metaboloma, proteoma e amido, previamente testado em outras espécies vegetais, foi adaptado pela primeira vez para o cafeeiro. As técnicas de aprendizado de máquina também se mostraram valiosas para a interpretação abrangente dos conjuntos de dados altamente complexos gerados a partir da integração de diferentes abordagens analíticas.

Esta pesquisa contribui significativamente para uma compreensão mais aprofundada da tolerância ao estresse e da aclimatação ambiental de cultivares de café, estabelecendo conexões entre características agronômicas e perfis metabólicos. Os resultados obtidos oferecem subsídios relevantes para o desenvolvimento de estratégias voltadas à otimização da produtividade, à melhoria da qualidade da bebida e à seleção de cultivares resilientes, capazes de enfrentar os desafios ambientais futuros.