

SCREENING FOR POTENTIAL AGRONOMIC RELEVANT METABOLITES IN EIGHT *COFFEA* SPP. GENOTYPES

ISABEL MORENO-DE-PALMA¹, ROGER GUEVARA², OLIVIER BONATO³,
 JOSÉ A. GUERRERO-ANALCO¹, LUIS ARTURO IBARRA-JUÁREZ^{1*},
AND JUAN L. MONRIBOT-VILLANUEVA^{1*}

¹ Red de Estudios Moleculares Avanzados, Instituto de Ecología, A.C., Xalapa, Mexico.

² Red de Biología Evolutiva, Instituto de Ecología, A.C., Xalapa, Mexico.

³ Institut de Recherche pour le Développement, Montpellier, France.

*Corresponding authors: luis.ibarra@inecol.mx; juan.monribot@inecol.mx

Abstract

Background: Metabolomic studies in *Coffea* spp. have identified agronomically important metabolites in leaves. However, endosperm metabolomics has focused on cup quality using processed beans, where chemistry composition is altered during fermentation. Therefore, these studies provide limited insight into defense metabolites present in *planta*.

Hypotheses: *Coffea* spp. endosperms from different cultivars grown under similar conditions exhibit distinct chemical profiles according to its genotype, including differentially accumulated metabolites with potential biological activity against pathogens and pests.

Studied species: Four *Coffea arabica* genotypes (Red Bourbon, Gesha, Pink Bourbon, and SL28) and four *Coffea arabica* introgressed with *Coffea canephora* (Mundo Maya, Marsellesa, Costa Rica 95, and Yellow Catucaí).

Study site and dates: Demonstrative coffee plantation of Francisco Javier Clavijero Botanical Garden (Instituto de Ecología, A.C.), Xalapa, Veracruz and private conservation area Finca Metlapili, Coatepec, Veracruz on March 14-15, 2025.

Methods: Freeze-dried matured endosperms were extracted with methanol via Accelerated Solvent Extraction. Extracts were analyzed by Ultra High-Performance Liquid Chromatography coupled to a High-Resolution Quadrupole Time of Flight mass spectrometer. Compounds were annotated using Progenesis QI and manually curated. Statistical and functional analyses were conducted in MetaboAnalyst.

Results: Each genotype showed a characteristic chemical profile, grouping into Timor-related and Typica/Bourbon-related clusters. Sixty-nine compounds were identified, including polyphenols, alkaloids, terpenoids, lipids, amino acids, and organic acids. Most metabolites were differentially accumulated among genotypes, and several have documented biological activity or roles in plant defense.

Conclusions: All *Coffea* spp. genotypes were distinguishable by their endosperm chemical profile. Several over-accumulated metabolites have reported roles in defense.

Keywords: Chemotaxonomy, plant defense, diterpenoids, metabolomics, oxylipins, phenylpropanoids.

Resumen

Antecedentes: Estudios en *Coffea* spp. han identificado metabolitos relevantes agrónomicamente en hojas. Sin embargo, estos estudios se han enfocado en la calidad en taza usando granos procesados, cuya composición química cambia durante la fermentación, proporcionando información limitada de metabolitos de defensa.

Hipótesis: El endospermo de *Coffea* spp. de diferentes variedades cultivadas en condiciones similares presentan perfiles químicos distintivos, incluyendo metabolitos diferencialmente acumulados con potencial actividad biológica contra patógenos y plagas.

Especies de estudio: Cuatro genotipos de *Coffea arabica* (Borbón rojo, Gesha, Borbón rosado, y SL28) y cuatro genotipos *Coffea arabica* con introgresión de *Coffea canephora* (Mundo Maya, Marsellesa, Costa Rica 95 y Catucaí amarillo).

Sitio de estudio y fechas: Cafetal del Jardín Botánico Francisco Javier Clavijero (Instituto de Ecología, A.C.), Xalapa, Veracruz y el área privada de conservación Finca Metlapili, Coatepec, Veracruz, 14 y 15 de marzo del 2025

Métodos: Extractos metanólicos de endospermos maduros liofilizados obtenidos por Extracción Acelerada por Disolvente fueron analizados por Cromatografía de Líquidos acoplada a espectrometría de masas Cuadrupolo-Tiempo de Vuelo. Los compuestos fueron identificados utilizando Progenesis QI y revisados manualmente. Los análisis estadísticos y funcionales fueron realizados con MetaboAnalyst.

Resultados: Cada genotipo exhibió perfiles químicos característicos, agrupándose en aquellos relacionados a Timor y a Tipica/Borbón. Se identificaron 69 compuestos, incluyendo polifenoles, alcaloides, terpenos, lípidos, aminoácidos y ácidos orgánicos. La mayoría están diferencialmente acumulados y varios participan en mecanismos de defensa.

Conclusiones: Los genotipos se distinguieron por el perfil químico del endospermo. Varios de los metabolitos diferencialmente acumulados tienen funciones de defensa.

Palabras clave: Defensa de plantas, diterpenoides, metabolómica, oxilipinas, fenilpropanoides.

This is an open access article distributed under the terms of the Creative Commons Attribution License CCBY-NC (4.0) international.

<https://creativecommons.org/licenses/by-nc/4.0/>



Metabolomics is a powerful tool for plant breeding by linking metabolic traits to agronomic performance (Fernandez *et al.* 2016). This is particularly relevant for coffee (*Coffea* spp. L.), a globally important crop with limited genetic diversity. Most *Coffea arabica* lines derive from a few Typica and Bourbon founder plants and, due to high autogamy, plantations exhibit low genetic diversity and therefore are highly vulnerable to drought, diseases, and pests (Meyer 1965, Anthony *et al.* 2002).

In Mexico, coffee remains primarily produced in traditional agroforestry systems that preserve a fair share of the native biodiversity and provide crucial socioeconomic support for ancestral communities -including many Indigenous smallholders- making crop sustainability a priority for both conservation and livelihoods (Hernández-Martínez *et al.* 2019).

Breeding efforts have sought to broaden *C. arabica* resistance by introgressing genes from *C. canephora*, notably via the natural Timor hybrid (HDT, by its original Portuguese name) and artificial Arabusta hybrids (de Oliveira 2017, Cheserek *et al.* 2020). These hybrids improved resistance to coffee leaf rust (CLR), coffee berry disease (CBD), and nematodes (Lashermes 2002, Maldonado & Ángel-Giraldo 2020). Still, monogenic resistance can be overcome by pathogens (Eskes *et al.* 1990), and other threats -such as the coffee berry borer (CBB, *Hypothenemus hampei*)- remain insufficiently addressed even from fundamental knowledge such as the susceptibility of different hybrids (Guerreiro Filho & Mazzafera 2003, Oliva *et al.* 2023). Most coffee metabolomic research has concentrated on processed beans, specifically on sensory attributes, compounds associated with ripeness, origin, cup quality, species, and variety (Bertrand *et al.* 2008, Wei *et al.* 2012, Bicho *et al.* 2013, Setoyama *et al.* 2013, Kwon *et al.* 2015, Marie *et al.* 2024) -while agronomically focused studies are fewer (Subroto *et al.* 2019, Chekol *et al.* 2024).

Processing (fermentation, drying, roasting) substantially alters bean chemistry and therefore introduces variance desirable for marketing coffee beans and coffee in the cup but not relevant when the objective is to detect intrinsic, genotype-driven metabolites linked to agronomic traits (Koskei *et al.* 2020, Therdtatha *et al.* 2023, Bollen *et al.* 2025).

Agronomic metabolomic work on coffee leaves has revealed drought-associated accumulation of amino acids, sugars, and organic acids (Chekol *et al.* 2024) and relationships between flavonoid levels and rust severity (Subroto *et al.* 2019), but metabolic markers tied to bean targeting pests, such as CBB, remain underexplored. A few studies have examined berry volatiles or genotype antibiosis (Guerreiro Filho & Mazzafera 2003, Cruz *et al.* 2016, Molina 2022), yet specific metabolites that could be used as breeding targets for antibiosis or other bean-relevant resistance traits have not been identified.

Here, we profile the untreated endosperm (raw beans) of commercial *Coffea* genotypes grown in Central Veracruz state, Mexico, to (1) assess whether cultivars can be discriminated by their intrinsic chemical composition and (2) identify candidate chemical fingerprints potentially linked to agronomic performance or resistance to bean pests. By analyzing unprocessed endosperm, this study aims to reduce post-harvest processing variance and reveal genotype-driven metabolic signals that could support rapid phenotyping and targeted breeding for more resilient coffee cultivars.

Materials and methods

Coffea genotypes. Genotypes were selected based on their genetic background, consumer and producer preferences. All sampled plants were visually healthy. A summary of the origin and agronomic characteristics of the selected genotypes is in [Figure 1](#). Four *C. arabica* and four *C. arabica* introgressed with *C. canephora* genotypes were selected for the study. The worldwide extended *C. arabica* ‘Red Bourbon’ is in the Typica-Bourbon genotypic group (T/B) and is sensitive to most diseases and pests. This genotype was chosen to represent early cultivated genotypes in America, which are more susceptible to diseases and pests (Montagnon *et al.* 2021, WCR 2024). The cultivar *C. arabica* ‘SL28’ is drought-tolerant but susceptible to pests. This variety was selected from the Red Bourbon lineage and has been cultivated in Kenya for over 100 years and was recently introduced to America (Montagnon *et al.* 2021, WCR 2024). We also included two Ethiopian landraces of *C. arabica* in our study as they possess greater genetic diversity and retain some of the wild resistance traits (Labouisse *et al.* 2008). The Ethiopian landrace ‘Gesha’ was

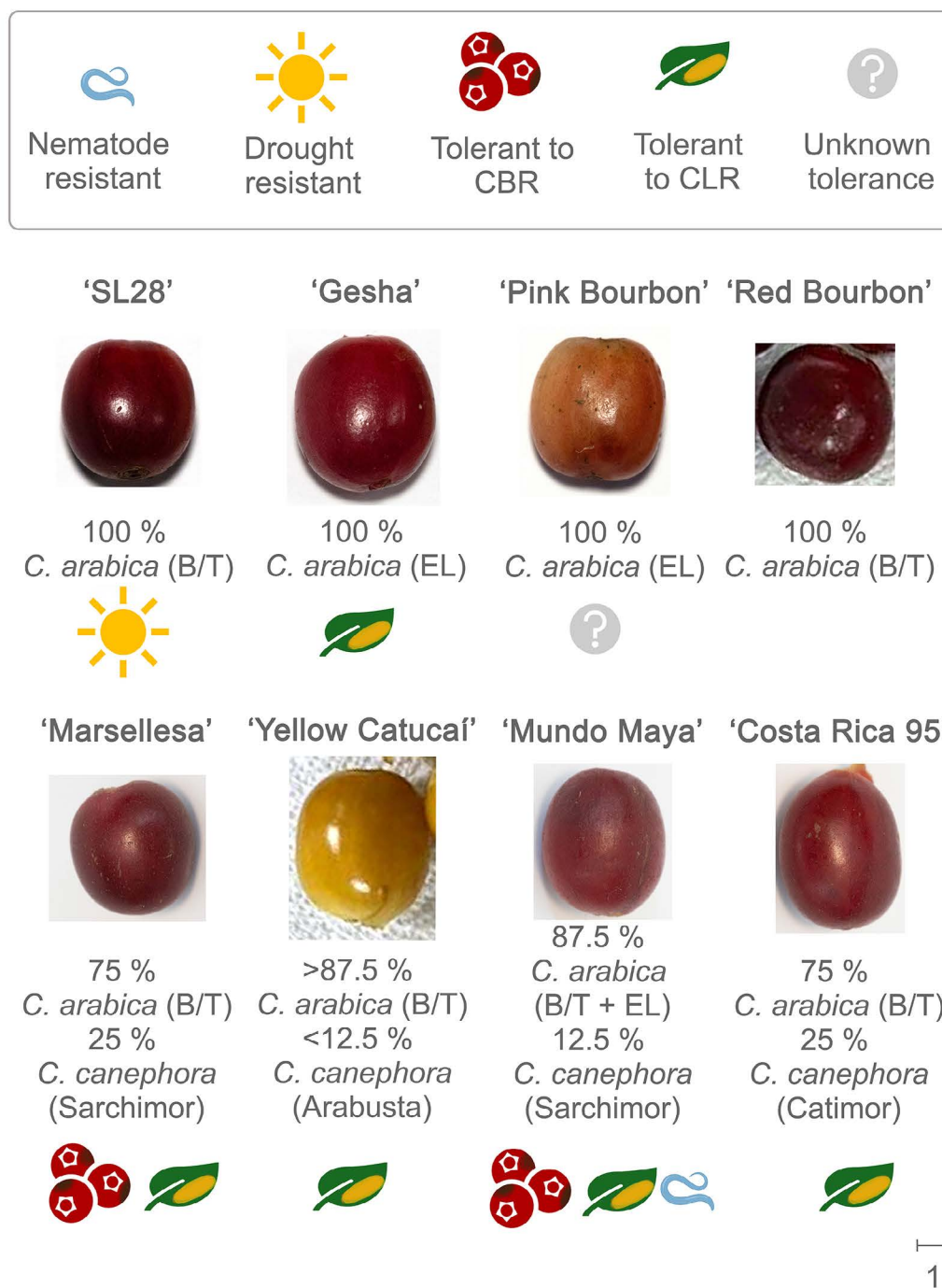


Figure 1. Coffee berries collected in the demonstrative coffee plantation of the *Francisco Javier Clavijero* Botanical Garden and the Private Conservation Area *Finca Metlapili*. Information about agronomic traits for each genotype is indicated. Pictures by Akeri Cruz Bonilla.

first introduced to the Americas in the mid-20th century. It has moderate tolerance to CLR and is renowned for its exceptional coffee cup quality (Montagnon *et al.* 2021, WCR 2024). The second Ethiopian landrace was “Pink Bourbon”, selected by coffee growers due to its favorable agricultural characteristics. This landrace was initially believed to be a cross between Red and Yellow Bourbon; however, it was later determined to be an Ethiopian landrace (Fretheim 2023). From the Sarchimor group we selected *C. arabica* ‘Marsellesa’ (Villa Sarchí CIFC 971/10 × Timor

832/2) and ‘Mundo Maya’ (T5296 × ET01) both resistant to CBD and CLR, and the latter tolerant to nematodes. There is also a representative of the Catimor group, *C. arabica* (Timor 832/1 × Caturra) ‘Costa Rica 95’, which is resistant to CLR. Though less common in America, we also included an Arabusta related genotype, *C. arabica* (Icatú × Catuai) ‘Yellow Catucaí.’ While Catuai is a cross between *C. arabica* cultivars of Caturra and Mundo Novo which in turn is a cross between Bourbon and Typica, Arabusta genes of Catucaí comes from Icatú, which was developed by a cross of Mundo Novo with the tetraploid *C. canephora*, followed by several backcrosses with Mundo Novo. Therefore, Yellow Catucaí are mostly T/B, however it retained moderate resistance to CLR from the Arabusta parents (Eskes *et al.* 1990, WCR 2024).

Sampling area. *Coffea* spp. berries from the genotypes of Marsellesa, Costa Rica 95, Mundo Maya and Yellow Catucaí were collected from the demonstrative coffee plantation of the *Francisco Javier Clavijero* Botanical Garden (INECOL), Xalapa, Veracruz (19° 29' N 97° 09' O, 1,400 m asl) while biological material corresponding to Red Bourbon, Gesha, Pink Bourbon, and SL28 were collected from the private conservation area *Finca Metlapili*, Coatepec, Veracruz (19° 29' N 96° 57' O, 1,300-1,400 m asl) on March 2025. Due to the geographical proximity and similar height of both coffee plantations, the climatic conditions were similar. Both sites were chosen due to their very similar agronomic management, where coffee plants grown in an agroforestry system under tree shade and adjacent to mist forest areas and water streams. Agronomic management include minimum intervention, avoiding the use of pesticides or herbicides, being weeds controlled manually. Fertilizing occurs only when necessary and only using compost. Coffee plants are not man-watered and are just irrigated by rainwater.

Plant material collection and processing. For each genotype at least 40 healthy berries at stage 88 according to the Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie’s scale for coffee development (Arcila-Pulgarín *et al.* 2002) were hand-picked from a minimum of five different plants. Pulp was removed immediately after harvest, and seeds were kept at -80 °C to avoid fermentation. The endosperms were lyophilized (FreeZone 2.5, Labconco, Kansas City, MO, USA) at -40 °C and < 0.1 mBar and milled using a food processor (NBR-0801, Nutribullet, Los Angeles, CA, USA).

Extract preparation. Methanolic extraction of each genotype was performed using 2.5 g of lyophilized endosperm powder mixed with one gram of diatomaceous earth as a dispersant agent in an Accelerated Solvent Extraction (ASE) system (Dionex ASE 350, Thermo Fischer Scientific, Sunnyvale, CA, USA). Extraction was performed using 34 mL cells with a 30 % rinse volume at 40 °C with two static cycles of five minutes each. The methanol solvent was completely removed using a rotary evaporator (R-100, Büchi, Flawil, Switzerland).

Untargeted metabolomics. Analyses were performed according to previous reports (Monribot-Villanueva *et al.* 2022). Briefly, dried extract was resuspended in methanol with 0.1 % (v/v) of formic acid to a concentration of 50 mg·mL⁻¹ and filtered with a 0.2 µm pore PVDF membrane. Five microliters were injected (15 °C) into an Ultra High-Performance Liquid Chromatography (UPLC) system (ACQUITY I-Class, Waters, Milford, MA, USA) coupled to a Quadrupole Time Of Flight (QTOF) high-resolution mass spectrometer (Synapt G2-Si, Waters, Milford, MA, USA). Samples run at 40 °C through an ACQUITY Bridged Ethylene Hybrid column (2.1 × 50 mm, particle size of 1.7 µm) with a linear elution gradient of water-acetonitrile with 0.1 % (v/v) of formic acid at a constant flux of 0.3 mL min⁻¹. Triplicates were run for each genotype and quality control (mix of equal volumes of each extract) samples were included.

Data processing, analysis, and identification of compounds. MassLynx (Waters version 4.3.1) software was used for the acquisition of raw data. Processing of the raw data was performed with Progenesis QI (Waters) including peak detection, retention time (RT) alignment, and analysis of variance (ANOVA) between genotypes. A good level of accuracy for the tentative annotation (confidence level two according to Schymanski *et al.* 2014) based on mass per charge (m/z)

error (≤ 10 ppm), mass spectra and isotopic pattern similarity was achieved. A quality filter based on taxonomy criterion was applied to the identifications to retain only those metabolites reported for *Coffea* spp. according to public databases included in Pubchem (Kim *et al.* 2025). The identity of caffeic acid, chlorogenic acid, *p*-coumaric acid and phenylalanine were confirmed by co-elution with authentic standards (level one according to Schymanski *et al.* 2014).

Peaks statistical analysis. Data analysis of the peak intensities was performed using MetaboAnalyst v. 6.0 software (Pang *et al.* 2024). Constant or single-value peaks were discarded. Raw intensities were quantile-normalized, $\log_{10}$ -transformed and Pareto-scaled. Unsupervised hierarchical clustering analysis (HCA) was performed using Euclidean distance and Ward's clustering method. Principal Component Analysis (PCA) was built using two components and the association between cultivars and the resulting structure was calculated through 999 permutations in a Permutational Multivariate Analysis of Variance (PERMANOVA). Functional analysis was performed including the 10 % most significant features (P value were all under 10^{-11}) for the ANOVA test performed by Progenesis Q1. *Arabidopsis thaliana* was chosen as a reference organism for metabolite identification and pathway enrichment analysis. Only pathways with at least three compounds identified by the algorithm, based on m/z values (level three according to Schymanski *et al.* 2014) were included in the analysis.

Chemical profile analysis. The annotated metabolites (accuracy levels 1 and 2) were analyzed. Peak intensities were transformed into $\log_2$ and Pareto scaled. To identify differentially accumulated metabolites, an HCA in a heatmap was built using Euclidean distance and Ward's clustering algorithm. To identify key metabolites for Coffee cultivars differentiation, it was performed a two component Spared Partial Least Squares Discriminant Analysis (sPLS-DA) model using ten metabolites by component.

Results

Untargeted metabolomics. The HCA dendrogram of the combined samples, along with a heatmap built using the feature-scaled, normalized peak intensities (Figure 2A), shows that the endosperms of each cultivar had characteristic metabolic profiles and were grouped based on their genetic background. In the dendrogram, the samples were first divided into two main groups: one included all Timor hybrids, and the other included all T/B-related cultivars. Yellow Catucaí, although it has introgression with Arabusta genes, was included in the T/B group, while the two Ethiopian landraces were each placed in their own group: Gesha was grouped within the Timor group, and SL28 was placed within the T/B group (Pink Bourbon). The Timor group showed more chemical variations among cultivars, and no clear subgroups were observed in the dendrogram. The first ramification was for the Catimor Costa Rica 95. The second and third ramifications were for the Sarchimor, Marsellesa, and Mundo Maya cultivars, respectively, and finally, the Gesha. The T/B-related group exhibited two subgroups: one formed by SL28 and the Yellow Catucaí, and the other by the Red Bourbon and the Pink Bourbon. On the other hand, the PCA (Figure 2B) was able to explain 35.4 % of the variability (the first component accounted for 23.1 % and the second component for 12.3 %), and the permutation analysis of the PCA was highly significant (P value = 0.001), discarding it as a chance occurrence. Cultivars were grouped by the first component's coefficients, similarly to the HCA (Figure 2A), into two groups: the Timor group, with negative coefficients, and the T/B group, with positive coefficients (Figure 2B). The first component coefficients of the three Timor-related cultivars were superposed; nevertheless, incorporating the second component allowed differentiation between the only Catimor (Costa Rica 95) and the two Sarchimor cultivars (Marsellesa and Mundo Maya). Gesha samples did not overlap with the Sarchimors but were close to Mundo Maya in both PCA dimensions. In the T/B related group, the first component mimics the first subdivision of the dendrogram, where Yellow Catucaí completely overlapped with SL28, and Red Bourbon partially overlapped with Pink Bourbon. The Ethiopian landrace Pink Bourbon had positive coefficient for the second component, contrary to the other T/B group genotypes which all had negative coefficients; even though Yellow Catucaí remain close to SL28 and Red Bourbon, the latter were the only ones which overlapped in the second component.

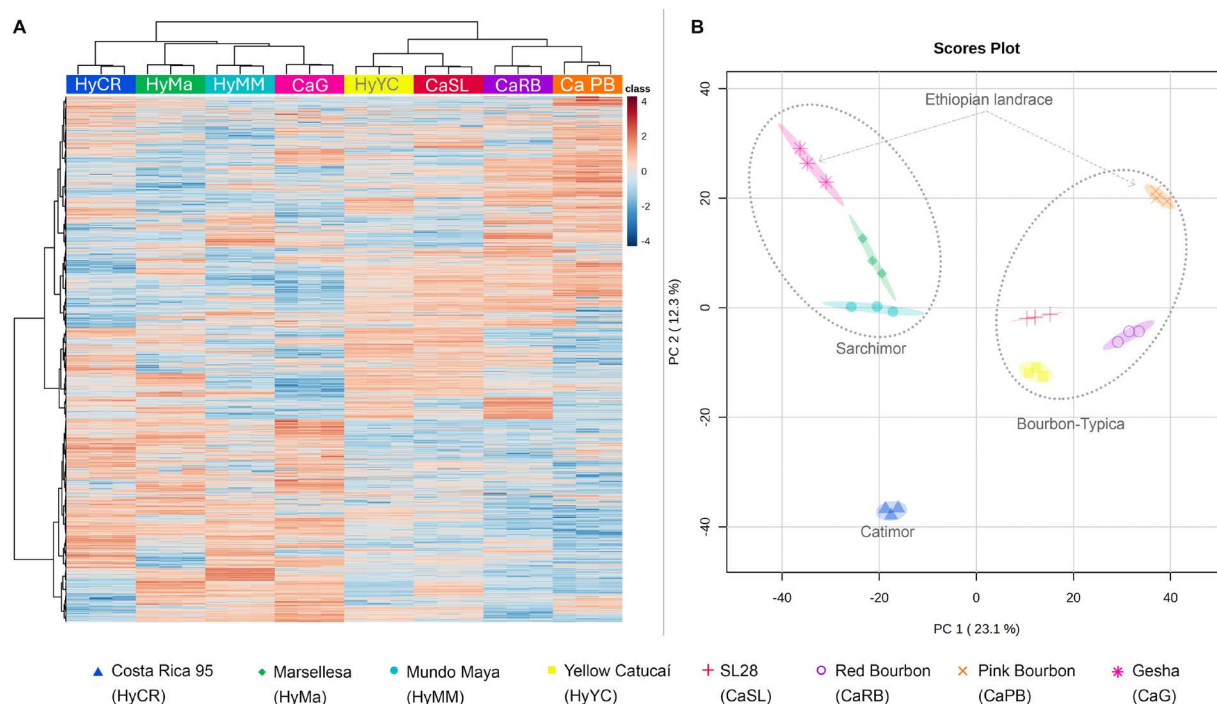


Figure 2. Statistical analysis on the full set of m/z-RT features from both positive and negative modes detected through UPLC-MS. A) Dendrogram using Euclidean distance and Ward algorithm combined with a heatmap showing color-coded z-scored normalized intensities of each m/z-RT features (rows) for each sample (columns). B) Unsupervised PCA (F value = 381.9 and $R^2 = 0.99405$) showing scores from the first PC in the X-axis and second PC in the Y-axis; percentage of explained variance by each PC is indicated in brackets.

Functional analysis. The Mummichog algorithm identified a total of 67 pathways that were shared by all genotypes, suggesting more quantitative than qualitative differences (Figure 3). The metabolic pathways with the greatest variance coefficient between genotypes in the enrichment factor are represented in Figure 3. The full list of pathways can be seen in the Supplementary material, Figure S1. Differences in enrichment factors between cultivars were small, as would be expected for individuals belonging to the same species. Arachidonic metabolism was the pathway with the most variation, and it was enhanced in Gesha, and least in SL28. Gesha also had enhanced glycine, serine, and threonine metabolism. Linoleic acid metabolism was more enriched in Marsellese, Red Bourbon, SL28, and Yellow Catucaí. Monoterpenoid biosynthesis was enriched in all genotypes compared to Costa Rica 95. Pink Bourbon stands out for its glycolysis-gluconeogenesis, inositol phosphate metabolism, and fructose metabolism enrichment. TCA cycle was enriched in all lines with introgressed genes from *C. canephora* and Red Bourbon.

Differential chemical profiles of Coffea cultivars. From the complete set of compounds identified with high confidence by Progenesis Q1, 69 fulfilled our stricter quality and taxonomic criteria (Table 1). Most compounds found were phenylpropanoids. We also identified alkaloids, terpenoids, organic acids and free amino acids. As expected in endosperm, auxin related compounds were identified. Peak intensities of the annotated compounds were used for an HCA (Figure 4) which gave a similar structure to the one produced by the full list of m/z-RT features (Figure 2A). The heatmap showed that the T/B-related group accumulated more metabolites than the Timor-related group. Most caffeoylquinic acids (CQAs), feruloylquinic acids (FQAs), amino acids and organic acids followed this pattern. Yellow Catucaí stood out because it accumulated theobromine, SL28 for the accumulation of 9-HOTrE, Red Bourbon for the accumulation of pipercolic acid, Pink Bourbon for the accumulation of 1-palmitoyl-2-hydroxyglycer-3-phosphoethanolamine, Marsellese for the accumulation of xanthosine, Mundo Maya for both linoleic, and 2,3-dihydroxypropyl palmitate and Gesha for its accumulation of cafestol.

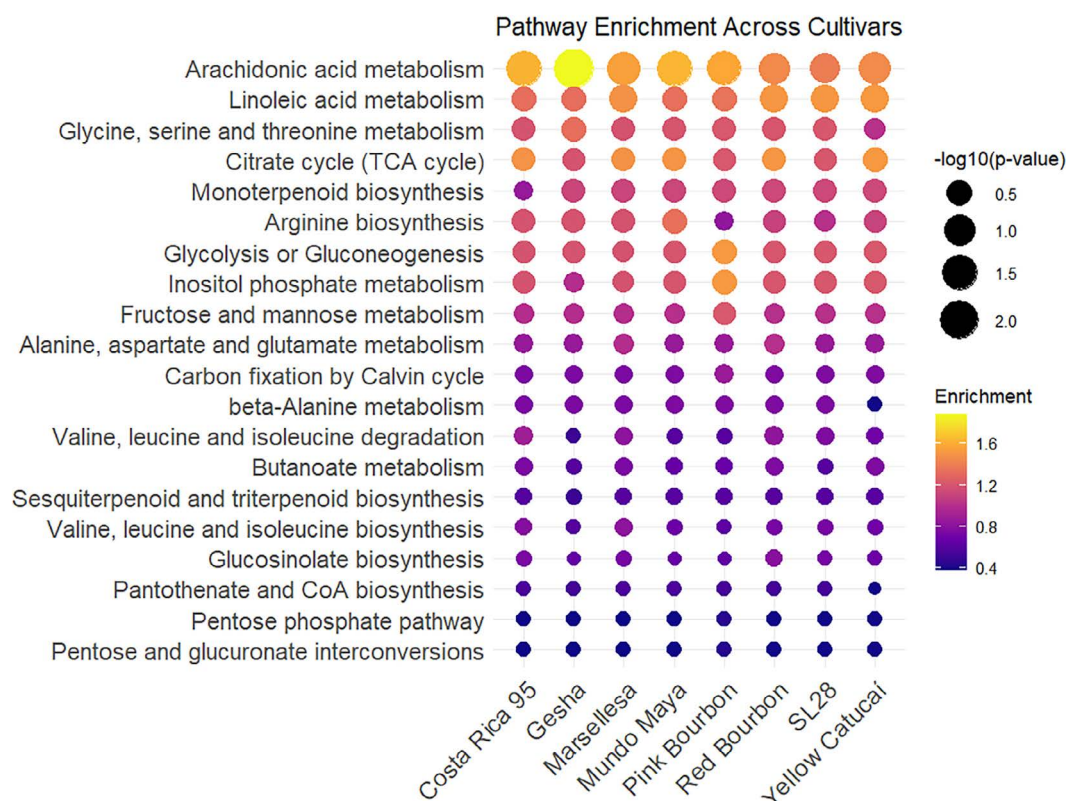


Figure 3. Pathway enrichment analysis using *Arabidopsis thaliana* as model organism. Color-coded enrichment factor and size of the dots representing the negative logarithm of the *P* value.

Table 1. Metabolites identified in the coffee endosperm.

Accepted Compound ID	RT (min)	m/z (Da)	Adducts	Mass Error (ppm)	Isotope Similarity (%)
Vanillic acid	5.92	165.0540	[M+H-H ₂ O] ⁺	-3.6	93.7
Methyl-chlorogenate	2.39	391.0975	[M+Na] ⁺	-6.7	89.7
3-O-Caffeoyl-4-O-sinapoylquinic acid	6.02	559.1442	[M-H] ⁻	-2.7	87.4
Sinapic acid	7.00	207.0645	[M+H-H ₂ O] ⁺	-2.9	92.9
Cinnamic acid	11.18	131.0491	[M+H-H ₂ O] ⁺	0.0	95.8
Caffeic acid hexoside	2.61	323.0754	[M-H ₂ O-H] ⁻	-5.3	89.2
Catechin gallate	3.87	407.0729	[M+H-2H ₂ O] ⁺	-7.24	86.7
5-Caffeoylshikimic acid	4.05	319.0796	[M+H-H ₂ O] ⁺	-4.7	88.9

Agronomic metabolites in eight *Coffea* spp. genotypes

Accepted Compound ID	RT (min)	m/z (Da)	Adducts	Mass Error (ppm)	Isotope Similarity (%)
Caffeic acid*	3.07	145.0289	[M+H-2H ₂ O] ⁺	2.8	92.2
4-Caffeoylquinic acid	2.94	353.0867	[M-H] ⁻	-3.2	98.0
3- <i>p</i> -Coumaroylquinic acid	5.59	337.0907	[M-H] ⁻	-6.3	91.7
Isoscopoletin	4.30	193.0486	[M+H] ⁺	-4.6	93.2
3,4-Dimethoxycinnamic acid	7.07	191.0706	[M+H-H ₂ O] ⁺	1.5	93.3
Coniferaldehyde	3.39	161.0588	[M+H-H ₂ O] ⁺	-4.9	95.9
1,3-Dicaffeoylquinic acid	4.96	517.1325	[M+H] ⁺	-2.9	89.0
4,5-Dicaffeoylquinic acid	5.41	515.1182	[M-H] ⁻	-2.5	96.0
3,4-Dicaffeoylquinic acid	4.93	515.1184	[M-H] ⁻	-2.2	96.0
3,5-Dicaffeoylquinic acid	5.06	539.1137	[M+Na] ⁺	-4.4	86.4
<i>p</i> -Coumaric acid*	3.47	147.0447	[M+H-H ₂ O] ⁺	4.0	93.4
3-Caffeoylquinic acid	2.26	353.0867	[M-H] ⁻	-3.1	98.1
Ferulic acid	5.87	177.0547	[M+H-H ₂ O] ⁺	0.5	93.1
Oxypeucedanin hydrate	7.78	305.0997	[M+H] ⁺	-7.4	90.8
4-Feruloylquinic acid	3.87	369.1167	[M+H] ⁺	-3.5	91.4
Chlorogenic acid*	2.85	355.1004	[M+H] ⁺	-5.4	90.1
Eugenol	11.84	165.0906	[M+H] ⁺	-2.3	98.6
LPC 18:3	12.80	562.3121	[M+FA-H] ⁻	-5.7	94.7
LPC 18:2	13.46	520.3397	[M+H] ⁺	-0.1	90.5

Accepted Compound ID	RT (min)	m/z (Da)	Adducts	Mass Error (ppm)	Isotope Similarity (%)
2-Palmitoylglycerol	13.83	363.3076	[M+CH ₃ OH+H]	-8.9	93.4
LPE 18:2	12.20	476.2770	[M-H] ⁻	-2.6	94.4
1-Palmitoyl-2-hydroxy-glycero-3-phosphoethanolamine	13.88	495.3169	[M+ACN+H] ⁺	-5.5	73.3
Linolenic acid	13.76	279.2317	[M+H] ⁺	-0.6	92.2
FA 18:2+3O	9.59	327.2151	[M-H] ⁻	-7.9	90.8
2,3-Dihydroxypropyl palmitate	7.56	331.2854	[M+H] ⁺	3.3	96.4
FA 18:3+2O	12.80	309.2070	[M-H] ⁻	-0.3	93.2
9-HOTrE	6.72	275.2024	[M-H ₂ O-H] ⁻	2.5	97.6
PC(16:0/0:0)	14.47	496.3399	[M+H] ⁺	0.2	96.7
Linoleic acid	12.22	279.2312	[M-H] ⁻	-6.4	96.8
Arachidic acid	13.15	312.3019	[M] ⁻	-4.8	96.8
Indolelactic acid	5.30	186.0541	[M-H ₂ O-H] ⁻	-9.6	93.3
Indole-3-ethanol	3.27	144.0801	[M+H-H ₂ O] ⁺	-4.4	84.3
Phenethylamine	3.52	122.0965	[M+H] ⁺	0.6	99.0
Dopamine	3.37	136.0752	[M+H-H ₂ O] ⁺	-3.5	94.7
Trigonelline	0.60	120.0448	[M+H-H ₂ O] ⁺	2.9	98.1
Indole-3-acetamide	1.43	175.0860	[M+H] ⁺	-3.6	93.4
Serotonin	1.53	177.1013	[M+H] ⁺	-5.5	95.7
Theobromine	2.37	181.0707	[M+H] ⁺	-7.2	90.9

Agronomic metabolites in eight *Coffea* spp. genotypes

Accepted Compound ID	RT (min)	m/z (Da)	Adducts	Mass Error (ppm)	Isotope Similarity (%)
Caffeine	2.99	195.0883	[M+H] ⁺	3.2	99.1
Mascaroside I	4.84	509.2383	[M+H] ⁺	0.4	91.2
Mascaroside II	6.19	513.2143	[M+Na] ⁺	9.7	85.1
Cafestol	11.34	299.2001	[M+H-H ₂ O] ⁺	-1.4	91.3
Kahweol	11.19	297.1851	[M+H-H ₂ O] ⁺	0.7	94.1
Geraniol	5.87	177.1250	[M+Na] ⁺	-0.1	90.2
Malic Acid	0.51	133.0139	[M-H] ⁻	-2.9	96.8
Isoleucine	0.62	132.1018	[M+H] ⁺	-0.6	97.2
Glutamic acid	0.64	130.0500	[M+H-H ₂ O] ⁺	1.1	93.5
Citric acid	0.82	191.0179	[M-H] ⁻	-9.3	96.1
Leucine	1.27	132.1019	[M+H] ⁺	0.0	95.1
3-methyl-2-oxovalerate	1.63	129.0547	[M-H] ⁻	-7.5	97.4
Phenylalanine*	1.8	166.0861	[M+H] ⁺	-0.8	93.9
Proline	3.25	116.0701	[M+H] ⁺	-4.1	97.7
Aspartic acid	3.29	134.0458	[M+H] ⁺	7.3	94.5
Pipecolic acid	0.77	130.0868	[M+H] ⁺	3.9	95.6
Quinic Acid	3.80	191.0551	[M-H] ⁻	-5.5	97.6
Xanthosine	1.50	283.0659	[M-H] ⁻	-8.7	94.2
Adenosine	1.32	268.1039	[M+H] ⁺	-0.4	92.7

Accepted Compound ID	RT (min)	m/z (Da)	Adducts	Mass Error (ppm)	Isotope Similarity (%)
Guanosine	1.36	284.0968	[M+H] ⁺	-7.6	94.9
9-(β-D-Ribofuranosyl)zeatin	6.12	393.1864	[M+ACN+H] ⁺	-4.8	97.6
Trehalose	0.51	377.0825	[M+Cl] ⁻	-9.0	90.6
Raffinose	0.72	549.1637	[M+FA-H] ⁻	-6.9	88.9

*Metabolites identity confirmed with standards.

Accumulation patterns of metabolites related to the caffeine biosynthetic pathway differed between genotypes. Trigonelline and caffeine followed the general pattern of accumulation in T/B related group and not in Timor related group. Gesha cultivar exhibited the lowest caffeine content, while the other Ethiopian landrace, Pink Bourbon, exhibited the highest level; nevertheless, both had similar levels of xanthosine and theobromine. Marsellesa presented elevated levels of xanthosine and low levels of theobromine and caffeine; Mundo Maya and Costa Rica 95 exhibited low levels of the three metabolites; SL28 and Red Bourbon had lower levels of xanthosine and theobromine than caffeine; and Yellow Catucaí greatly accumulated theobromine compared to all other genotypes. As well as caffeine, other alkaloids or alkaloids-related compounds were generally more accumulated in T/B related genotypes. Phenethylamine did not present great differences in accumulation between genotypes while dopamine was accumulated mostly in Pink Bourbon and lowest in Gesha. Pipecolic acid was mostly accumulated in Red Bourbon and exhibited the lowest content in Mundo Maya cultivar. In fact, Mundo Maya was the genotype with less accumulation of alkaloids being closely followed by Marsellesa.

Lipids were most accumulated in Mundo Maya and Gesha genotypes. Mundo Maya especially accumulated linolenic acid and 2,3-dihydroxypropyl palmitate and Gesha LPC 18:2, LPC 18:3 and 2-palmitoylglycerol. Amino acids and organic were mostly accumulated in the Timor-related group while being diminished in the T/B-related group. Diterpenoids exhibited different accumulation patterns. Only cafestol was differentially accumulated in Gesha while none of the other diterpenoids exhibited differences compared to the other genotypes. Pink Bourbon and Red Bourbon accumulated mascaroside II and cafestol, but not as much as in Gesha. The three varieties Yellow Catucaí, SL28 and Marsellesa follow a similar diterpenoid accumulation pattern and the monoterpenoid geraniol was mostly accumulated in SL28 genotype.

Modelling of metabolite profile differentiation. The sPLS-DA model built with the annotated compounds was able to differentiate all the genotypes using only two components with ten metabolites per component (Figure 5A). Together, the 20 metabolites included in both components explained 50.5 % of the total variability. Coffee cultivars were separated into four quadrants: Mundo Maya and Gesha clustered in the negative-negative quadrant; Costa Rica 95 and Marsellesa were clustered in the negative-positive quadrant; in the positive-positive quadrant the T/B related Yellow Catucaí and SL28 were grouped and in the positive-negative quadrant were Red Bourbon and Pink Bourbon. First component, built with primary metabolism compounds including amino acids and nucleosides as well as specialized metabolism, particularly CQAs and FQAs, was able to differentiate between the Timor related and the T/B genotypes. On the other hand, metabolites from the second component presented a more varied distribution of chemical groups, including phenolics, diterpenoids, and lipids. The diterpenoid cafestol, and the phenylpropanoid dimethoxycinnamic acid were the most important compounds in this component, followed by vanillic acid, mascaroside I and kahweol. This component was able to differentiate almost all Bourbon genotypes from Ethiopian landraces, excepting for Red Bourbon.

Agronomic metabolites in eight *Coffea* spp. genotypes

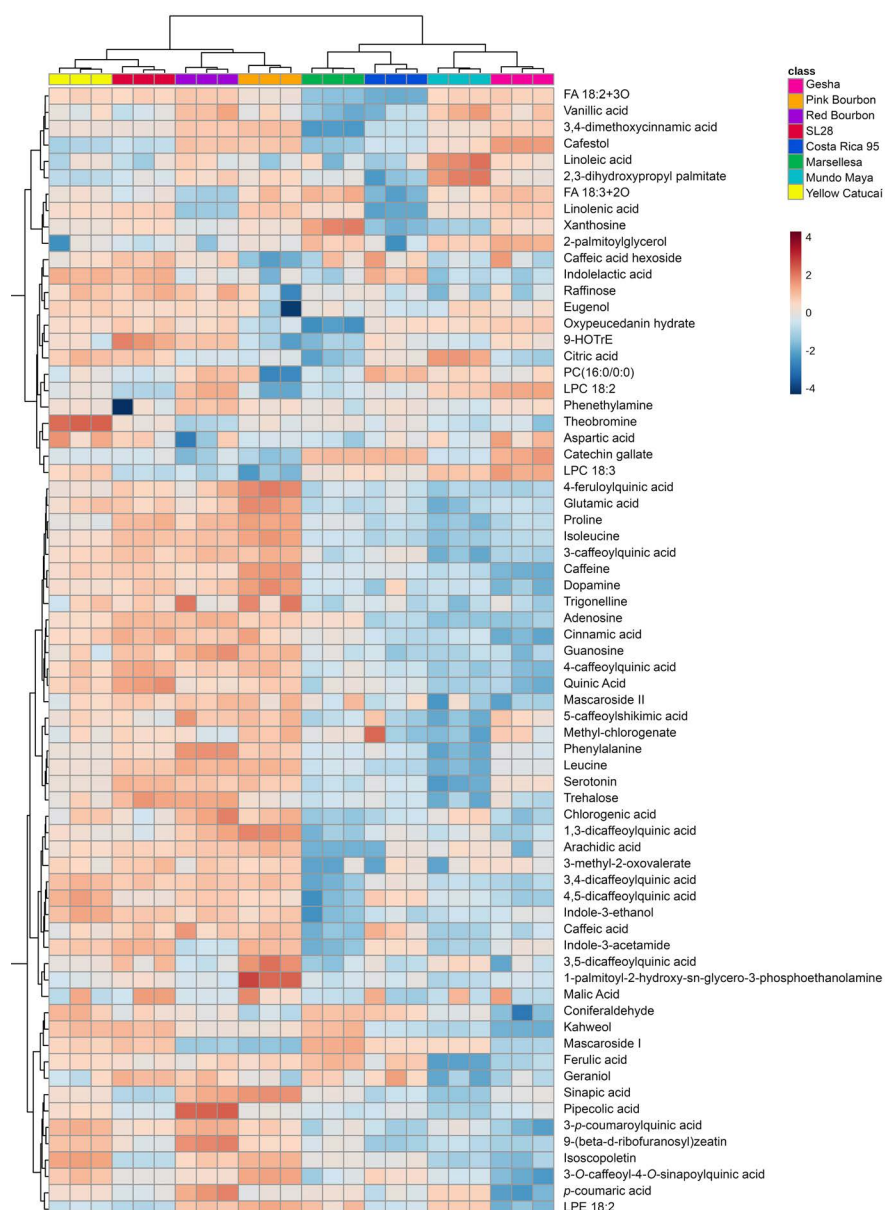


Figure 4. Dendrogram using Euclidean distance and Ward algorithm combined with a hierarchically ordered heatmap showing color-coded z-scored normalized intensities of each high confidence metabolite identification (rows) for each sample (columns).

Discussion

Discrimination of cultivars by their chemical profiles. Previous studies have successfully discriminated among cultivars based on endosperm chemical composition (Bertrand *et al.* 2008, Wei *et al.* 2012, Bicho *et al.* 2013). While some studies compared genotypes under the same conditions, they used processed green coffee beans, which introduced a source of variation. In this study, it was possible to correctly cluster samples in consideration of their genotype using unprocessed lyophilized endosperm by their differential chemical profile. Importantly, genotypes did not cluster according to the coffee plantation, whereas other studies have shown geographic origin to be a crucial factor in determining metabolic profiles (Bertrand *et al.* 2008). It might be due to our effort to minimize environmental variability by choosing similar coffee plantations. However, the studies with processed green coffee beans highlight

the contribution of the process to variability, which is known to have a great effect on physicochemical characteristics and cup quality (Therdatha *et al.* 2023, Freire Muñoz *et al.* 2024).

Genotypes in the HCA and PCA were grouped into cultivars according to their genetic background. Timor cultivars were all grouped, as were T/B genotypes. It is important to note that Ethiopian landraces exhibit high genetic variability; therefore, it is not surprising that they are not clustered together (Silvestrini *et al.* 2007, Labouisse *et al.* 2008). In fact, Pink Bourbon's genetic studies conducted by two private laboratories showed that its closest relatives were Ethiopian landraces; however, none of them were Gesha, while exhibiting some similarity with Bourbon and other Yemen-dispersed varieties (Fretheim 2023). In fact, even in academic writings it was wrongly thought to be a natural cross between Red and Yellow Bourbon due to its phenotypic characteristics, therefore the name given to this variety (Osso-Bolaño & Perez-Santofimio 2021). The association between the Ethiopian landrace Gesha, famous for its excellent cup aroma, and the Sarchimor varieties was unsuspected. However, they share important agronomic characteristics. For instance, Timor hybrids were engineered to be resistant to the CLR, and Gesha is known to be naturally resistant to type II CLR, the most common type found in Latin America (WCR 2024). Interestingly, Gesha is chemically closer to Mundo Maya in the HCA; the latter is a cross of a Sarchimor and an Ethiopian landrace collected by CATIE, and the Gesha genotype was also collected by CATIE. It is known that they come from different original Ethiopian plants, with distinct accessions -Gesha original accession ID was 'T2722' and ancestor of Mundo Maya 'ET01' - however, close genetic relation between both ancestors might be possible (WCR 2024). The presence of the backcross hybrid Yellow Catucaí in the T/B related group is not surprising as both parentals Icatú and Catucaí varieties have mostly T/B genetics. The parental *C. arabica* Catucaí is 100 % T/B ('Caturra' and 'Mundo Novo') while Icatú was developed through several backcrosses of Arabusta with 'Bourbon' and 'Mundo Novo' ('Bourbon' × 'Typica') and it is mostly considered *C. arabica* (Eskes *et al.* 1990, Velásquez Orozco 2019, WCR 2024).

Caffeine, CQAs, and FQAs, although they do not belong to the same biosynthetic pathway, are usually correlated and have been used for chemotaxonomic classification of *Coffea* spp. Both *C. arabica* and *C. canephora* belonged to the group of *Coffea* species that accumulated high concentrations of all caffeine, CQAs and FQAs in comparison with other *Coffea* spp. (Anthony *et al.* 1993). Studies in commercial green coffee beans generally found higher content of

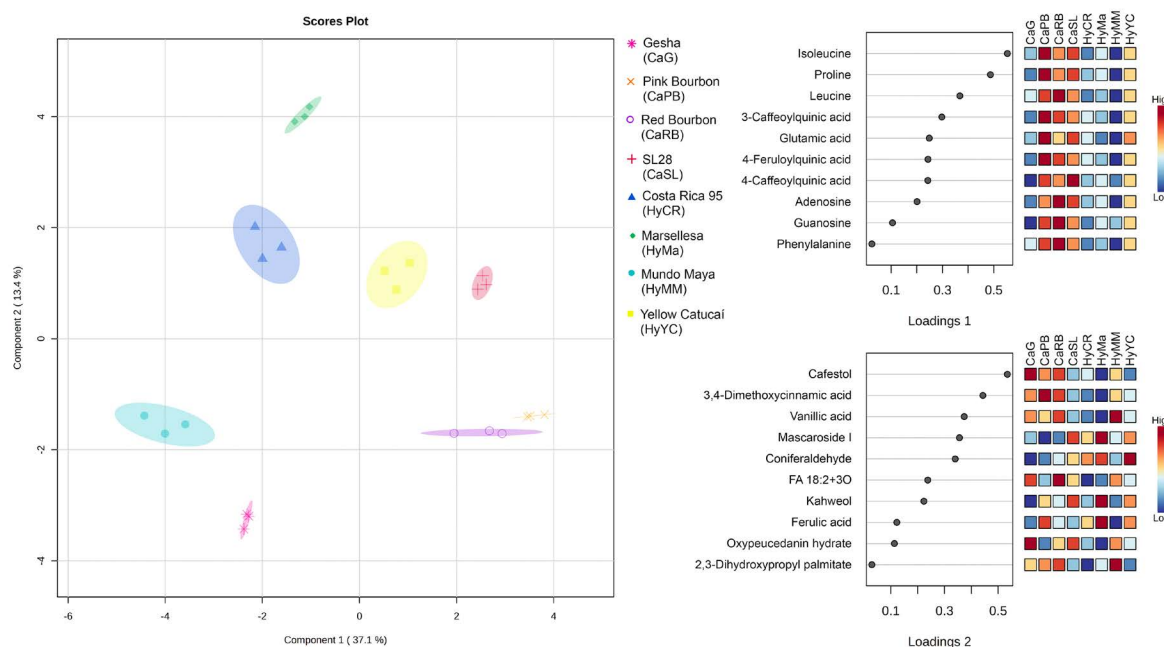


Figure 5. Score plot (left) of the first two components of the sPLS-DA built with the high confidence identifications (error rate ~ 0 calculated by Leave-One-Out-Cross-Validation). The percentage of explained variation by each component is written in parenthesis. The loading values for metabolites of each component is presented together with the z-score for each genotype.

caffeine and total CQAs and FQAs in *C. canephora* than in *C. arabica* (Wei *et al.* 2012, Bicho *et al.* 2013). Nevertheless, different genotypes are usually grown in different conditions, *C. canephora* and hybrids, are usually grown at lower altitudes, in warmer climates and in plain sun while *C. arabica* is usually grown at higher altitudes, cooler climate and under tree shade and caffeine production has been reported to varied enormously according to environmental conditions (Ramalho *et al.* 2018, Chekol *et al.* 2024). Other studies have shown that there are complex metabolic variations in hybrids, backcrosses and introgressed lines showing overlapping and smaller values of caffeine, CQAs, and FQAs than in *C. arabica* (Guerrero *et al.* 2001, Bertrand *et al.* 2008, Cheserek *et al.* 2021). In the present study all *C. arabica* genotypes but for Gesha exhibited higher accumulation of caffeine and CQAs than the genotypes introgressed with *C. canephora*. A previous comparative analysis, including three of the genotypes included in the present study, the introgressed Costa Rica 95 and Marsellesa and the pure *C. arabica* Gesha, showed that there were no significant differences in total CQAs and FQAs content between these varieties (Jeon *et al.* 2024). While there are discrepancies in the CQAs and FQAs accumulation pattern between this previous study and ours, for example, 4-CQA in the present study was accumulated in Marsellesa than in Gesha, while the pattern was opposite in the Jeon *et al.* (2024) study. Previous studies also showed that location not only affects the content of these metabolites, but the effect it produces differs depending on the genotype (Bertrand *et al.* 2008, Cheserek *et al.* 2021). Despite the generalized use of CQAs as coffee genotype identification criteria, due to the complex relation between genotype and phenotype for these metabolites it might not be possible to generalize this criterion and might only apply for specific climatic conditions.

Identified compounds with possible agronomic traits. Many of the identified compounds play a role in plant defense. Caffeine has been known to be an insecticide for decades (Nathanson 1984), however the most important insect pest that attacks coffee bean, CBB, is known to be well adapted to this compound (Ceja-Navarro *et al.* 2015). Similarly to previous studies, a great diversity of CQAs, FQAs, and other cinnamic acid derivatives were identified which seems to be useful in genotypes discrimination due to their presence in the first component of the sPLS-DA. These compounds belong to a very active group with reported antimicrobial and insecticide activity (Almeida *et al.* 2006, Martínez *et al.* 2017, Lin *et al.* 2022). This extends to its constituents, cinnamic acid, *p*-coumaric acid, ferulic acid, and caffeic acid some of which have shown biocidal activity against *Collecotrichum*, the causal agent of the CBD (Cheng *et al.* 2008, Roy *et al.* 2018, Kaushik *et al.* 2019, Hu *et al.* 2023b). Ferulic acid, caffeic acid, and chlorogenic acid have been reported as antifeedant or/and insecticidal and have been related with plant resistance mechanisms to insects (Classen *et al.* 1990, Joshi *et al.* 2014, Lin *et al.* 2022). Moreover, in the potential control of CBB, a member of the Curculionidae family, ferulic acid exhibited antibiosis (affecting Dobie index) and antixenosis effect against *Sitophilus zeamais*, another endosperm borer member of this family (Classen *et al.* 1990). Nevertheless, transcriptomic studies of CBB have shown high expression levels of cytochrome P450, which is known to detoxify these metabolites (Noriega *et al.* 2019, Lin *et al.* 2022).

Vanillic acid is a well-known antifungal compound and has also shown modest insecticidal effect (Cheng *et al.* 2008, Roy *et al.* 2018, Kaushik *et al.* 2019, Hu *et al.* 2023a). In this group, eugenol might be an interesting compound as it is a potent insecticide to some insects (Huang *et al.* 2002, Cheng *et al.* 2008, Adhikari *et al.* 2022). Moreover, eugenol also affects *S. zeamais* through antifeedant effect (Huang *et al.* 2002). Terpenoids are another set of plant metabolites with agronomic potential, they are important in differentiating genotypes based on the chemical profiles due to their inclusion in the second component of the sPLS-DA. They have been reported with antibacterial, antifungal and insecticidal properties, particularly geraniol and cafestol (Zanetti *et al.* 2015, Tang *et al.* 2018, Antoine *et al.* 2023). Lipids, particularly linolenic acid, both as a free fatty acid and oxidated forms, are signaling and precursors of jasmonates (Wasternack & Hause 2013). Some exogenously applied hydroxy unsaturated fatty acids have direct antifungal properties (Yasari *et al.* 2021).

Even though cultivars have been developed against some diseases and pests important for coffee, such as nematodes, CBD, and CLR, there are few genotypes described to have a negative effect on the worst plant pest for this crop, CBB (Molina 2022), none of which are commercially grown. The antixenosis or antibiosis effect of these genotypes is not characterized and is not being related yet to any metabolite, though it seems to respond to a quantitative trait locus

(Molina *et al.* 2022). Caffeine does not seem to be involved, however, in this study some interesting compounds have been identified with potential insecticidal or antifeedant properties such as ferulic acid, eugenol, coniferaldehyde, geraniol and cafestol (Huang *et al.* 2002, Zanetti *et al.* 2015, AlJabr *et al.* 2017, Novaes *et al.* 2019, Martins *et al.* 2025). These compounds and *Coffea* genotypes accumulating them should be evaluated against the CBB.

Altogether, this study highlights the presence of metabolites with potential agronomic importance and proves that they are differentially accumulated under the same conditions in different genotypes. These results pave the way to assessing the antixenosis or antibiosis effect of these genotypes and correlate them with the chemical composition.

Supplementary material

Supplemental data for this article can be accessed here: <https://doi.org/10.17129/botsci.3812>

Acknowledgements

We are grateful to Dr. E. Couttolenc Brenis for his advice about *Coffea* genotypes. We thank the Francisco Javier Clavijero Botanical Garden and the Private Conservation Area Finca Metlapili for donating the coffee berries. We also thank Akeri Cruz Bonilla for taking pictures of the collected coffee berries.

Literature cited

- Adhikari K, Khanikor B, Sarma R. 2022. Persistent susceptibility of *Aedes aegypti* to eugenol. *Scientific Reports* **12**: 2277. DOI: <https://doi.org/10.1038/s41598-022-06302-8>
- AlJabr AM, Hussain A, Rizwan-ul-Haq M, Al-Ayedh H. 2017. Toxicity of Plant Secondary Metabolites Modulating Detoxification Genes Expression for Natural Red Palm Weevil Pesticide Development. *Molecules: A Journal of Synthetic Chemistry and Natural Product Chemistry* **22**: 169. DOI: <https://doi.org/10.3390/molecules22010169>
- Almeida AAP, Farah A, Silva DAM, Nunan EA, Glória MBA. 2006. Antibacterial activity of coffee extracts and selected coffee chemical compounds against enterobacteria. *Journal of Agricultural and Food Chemistry* **54**: 8738-8743. DOI: <https://doi.org/10.1021/jf0617317>
- Anthony F, Clifford MN, Noirot M. 1993. Biochemical diversity in the genus *Coffea* L.: chlorogenic acids, caffeine and mozambioside contents. *Genetic Resources and Crop Evolution* **40**: 61-70. DOI: <https://doi.org/10.1007/BF00052636>
- Anthony F, Combes M, Astorga C, Bertrand B, Graziosi G, Lashermes P. 2002. The origin of cultivated *Coffea arabica* L. varieties revealed by AFLP and SSR markers. *Theoretical and Applied Genetics* **104**: 894-900. DOI: <https://doi.org/10.1007/s00122-001-0798-8>
- Antoine G, Vaissayre V, Meile J-C, Payet J, Conéjéro G, Costet L, Fock-Bastide I, Joët T, Dussert S. 2023. Diterpenes of *Coffea* seeds show antifungal and anti-insect activities and are transferred from the endosperm to the seedling after germination. *Plant Physiology and Biochemistry* **194**: 627-637. DOI: <https://doi.org/10.1016/j.plaphy.2022.12.013>
- Arcila-Pulgarín J, Buhr L, Bleiholder H, Hack H, Meier U, Wicke H. 2002. Application of the extended BBCH scale for the description of the growth stages of coffee (*Coffea* spp.). *Annals of Applied Biology* **141**: 19-27. DOI: <https://doi.org/10.1111/j.1744-7348.2002.tb00191.x>
- Bertrand B, Villarreal D, Laffargue A, Posada H, Lashermes P, Dussert S. 2008. Comparison of the effectiveness of fatty acids, chlorogenic acids, and elements for the chemometric discrimination of coffee (*Coffea arabica* L.) varieties and growing origins. *Journal of agricultural and food chemistry* **56**: 2273-80. DOI: <https://doi.org/10.1021/jf073314f>
- Bicho N, Leitão A, Ramalho J, Alvarenga N, Lidon F. 2013. Identification of chemical clusters discriminators of

- Arabica and Robusta green coffee. *International Journal of Food Properties* **16**: 895-904. DOI: <https://doi.org/10.1080/10942912.2011.573114>
- Bollen R, Rojo-Poveda O, Katshela B, Tshimi E, Caroline S, Delporte C, Vandeloock F, Stoffelen P, Honnay O. 2025. The impact of coffee genotype and processing method on the sensory and metabolite profiles of Robusta coffee (*Coffea canephora*). *European Food Research and Technology*: **251**: 1171-11901-20. DOI: <https://doi.org/10.1007/s00217-025-04696-z>
- Ceja-Navarro JA, Vega FE, Karaoz U, Hao Z, Jenkins S, Lim H C, Kosina P, Infante F, Northen TR, Brodie EL. 2015. Gut microbiota mediate caffeine detoxification in the primary insect pest of coffee. *Nature Communications* **6**: 7618. DOI: <https://doi.org/10.1038/ncomms8618>
- Chekol H, Warkineh B, Shimber T, Mierek-Adamska A, Dąbrowska G, Degu A. 2024. Drought Stress Responses in Arabica Coffee Genotypes: Physiological and Metabolic Insights. *Plants* **13**. DOI: <https://doi.org/10.3390/plants13060828>
- Cheng S-S, Liu J-Y, Chang E-H, Chang S-T. 2008. Antifungal activity of cinnamaldehyde and eugenol congeners against wood-rot fungi. *Bioresource Technology* **99**: 5145-5149. DOI: <https://doi.org/10.1016/j.biortech.2007.09.013>
- Cheserek JJ, Ngugi K, Muthomi JW, Omondi CO. 2020. Genetic variability, heritability and correlation of quantitative traits for Arabusta coffee (*C. arabica* L. X Tetraploid *C. canephora* Pierre). *Journal of Plant Breeding and Crop Science* **12**: 50-57. DOI: <https://doi.org/10.5897/JPBCS2019.0831>
- Cheserek JJ, Ngugi K, Muthomi JW, Omondi CO, Ezekiel NK. 2021. Green bean biochemical attributes of Arabusta coffee hybrids from Kenya using HPLC and soxhlet extraction methods. *Australian Journal of Crop Science* **15**: 201-208. DOI: <https://doi.org/10.3316/informit.709519244977316>
- Classen D, Arnason JT, Serratos JA, Lambert JDH, Nozzolillo C, Philogène BJR. 1990. Correlation of phenolic acid content of maize to resistance to *Sitophilus zeamais*, the maize weevil, in CIMMYT'S collections. *Journal of Chemical Ecology* **16**: 301-315. DOI: <https://doi.org/10.1007/bf01021766>
- Cruz L, Díaz-Díaz B, Rojas J. 2016. Coffee volatiles induced after mechanical injury and beetle herbivory attract the coffee berry borer and two of its parasitoids. *Arthropod-Plant Interactions* **10**. DOI: <https://doi.org/10.1007/s11829-016-9417-0>
- de Oliveira SC. 2017. *Origin of the allotriploid "Híbrido de Timor" through a karyotype comparison with its Coffea ancestors*. PhD Thesis. Universidade Federal do Espírito Santo.
- Eskes AB, Hoogstraten JGJ, Toma-Braghini M, Carvalho A. 1990. Race-specificity and inheritance of incomplete resistance to coffee leaf rust in some Icatu coffee progenies and derivatives of Híbrido de Timor. *Euphytica* **47**: 11-19. DOI: <https://doi.org/10.1007/BF00040356>
- Fernandez O, Urrutia M, Bernillon S, Giauffret C, Tardieu F, Le Gouis J, Langlade N, Charcosset A, Moing A, Gibon Y. 2016. Fortune telling: metabolic markers of plant performance. *Metabolomics* **12**: 158. DOI: <https://doi.org/10.1007/s11306-016-1099-1>
- Freire Muñoz DA, Franco Pérez AF, Tello Velastegui AI, Llundu Michelena BJ. 2024. Analysis of the Fermentation Processes of Arabica Coffee and its Impact on the Cup. *Health Leadership and Quality of Life* **3**: 1-7 DOI: <https://doi.org/10.56294/hl2024.500>
- Fretheim I. 2023. Pink... Bourbon?: Cryptozoology and Genetics in Specialty Coffee. *Cafe Imports*. Website <https://www.cafeimports.com/north-america/blog/2023/09/26/pink-bourbon-cryptozoology-and-genetics-in-specialty-coffee/> (accessed 22 May 2025)
- Guerreiro Filho O, Mazzafera P. 2003. Caffeine and resistance of coffee to the berry borer *Hypothenemus hampei* (Coleoptera: Scolytidae). *Journal of Agricultural and Food Chemistry* **51**: 6987-6991. DOI: <https://doi.org/10.1021/jf0347968>
- Guerrero G, Suárez M, Moreno G. 2001. Chlorogenic acids as a Potential criterion in coffee genotype selections. *Journal of Agricultural and Food Chemistry* **49**: 2454-2458. DOI: <https://doi.org/10.1021/jf001286u>
- Hernández-Martínez G, Vera Espíndola R, De la Rosa Portilla A. 2019. *The coffee of Mexico. The richness of its regions*. México: Cafecol AC. ISBN: 978-607-98601-0-3

- Hu F, Bi X, Fu X, Li Y, Li G, Li Y, Liu D, Yang Y, Shi R, Dong W. 2023a. Comparative metabolome profiles and antioxidant potential of four *Coffea arabica* L. varieties differing in fruit color. *Diversity* **15**: 724. DOI: <https://doi.org/10.3390/d15060724>
- Hu J, Wu J, Wang M, Jiao W, Chen Q, Du Y, Chen X, Yang X, Fu M. 2023b. p-coumaric acid prevents *Colletotrichum gloeosporioides* by inhibiting membrane targeting and organic acid metabolism. *Postharvest Biology and Technology* **204**: 112447. DOI: <https://doi.org/10.1016/j.postharvbio.2023.112447>
- Huang Y, Ho S-H, Lee H-C, Yap Y-L. 2002. Insecticidal properties of eugenol, isoeugenol and methyleugenol and their effects on nutrition of *Sitophilus zeamais* Motsch. (Coleoptera: Curculionidae) and *Tribolium castaneum* (Herbst) (Coleoptera: Tenebrionidae). *Journal of Stored Products Research* **38**: 403-412. DOI: [https://doi.org/10.1016/S0022-474X\(01\)00042-X](https://doi.org/10.1016/S0022-474X(01)00042-X)
- Jeon YA, Natraj P, Kim SC, Moon J-K, Lee YJ. 2024. Comparative analysis of phytochemical and functional profiles of Arabica coffee leaves and green beans across different cultivars. *Foods* **13**: 3744. DOI: <https://doi.org/10.3390/foods13233744>
- Joshi RS, Wagh TP, Sharma N, Mulani FA, Sonavane U, Thulasiram HV, Joshi R, Gupta VS, Giri AP. 2014. Way toward “Dietary Pesticides”: Molecular investigation of insecticidal action of caffeic acid against *Helicoverpa armigera*. *Journal of Agricultural and Food Chemistry* **62**: 10847-10854. DOI: <https://doi.org/10.1021/jf503437r>
- Kaushik P, Sarkar D J, Chander S, Rana V, Shakil N. 2019. Insecticidal activity of phenolic acid amides against brown planthopper (BPH), *Nilaparvata lugens* (Stål) and their QSAR analysis. *Journal of Environmental Science and Health, Part B* **54**: 1-9. DOI: <https://doi.org/10.1080/03601234.2019.1574174>
- Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L, Zhang J, Bolton EE. 2025. PubChem 2025 update. *Nucleic Acids Research* **53**: D1516-D1525. DOI: <https://doi.org/10.1093/nar/gkaf1059>
- Koskei RK, Mugendi B, Muliro P. 2020. Effects of processing methods on fatty acid profiles and biochemical compounds of Arabica coffee cultivars. *African Journal of Food Science* **14**: 92-97. DOI: <https://doi.org/10.5897/AJFS2020.1917>
- Kwon D-J, Jeong H-J, Moon H, Kim H-N, Cho J-H, Lee J-E, Hong K S, Hong Y-S. 2015. Assessment of green coffee bean metabolites dependent on coffee quality using a 1H NMR-based metabolomics approach. *Food Research International* **67**: 175-182. DOI: <https://doi.org/10.1016/j.foodres.2014.11.010>
- Labouisse J-P, Bellachew B, Kotecha S, Bertrand B. 2008. Current status of coffee (*Coffea arabica* L.) genetic resources in Ethiopia: Implications for conservation. *Genetic Resources and Crop Evolution* **55**: 1079-1093. DOI: <https://doi.org/10.1007/s10722-008-9361-7>
- Lashermes P. 2002. Breeding tools for durable resistance to nematodes (*Meloidogyne* spp.) of coffee varieties. *Plant Protection Science* **38**: 717-720. DOI: <https://doi.org/10.17221/10598-PPS>
- Lin D, Fang Y, Li L, Zhang L, Gao S, Wang R, Wang J. 2022. The insecticidal effect of the botanical insecticide chlorogenic acid on *Mythimna separata*(Walker) is related to changes in MsCYP450 gene expression. *Frontiers in Plant Science* **13**. DOI: <https://doi.org/10.3389/fpls.2022.1015095>
- Maldonado-Londoño CE, Ángel-Giraldo L. 2020. Resistencia genética a la enfermedad de la cereza del café en variedades cultivadas en Colombia. *Revista Cenicafé* **71**: 68-90. DOI: <https://doi.org/10.38141/10778/1121>
- Marie L, Breitler J-C, Bamogo PKA, Bordeaux M, Lacombe S, Rios M, Lebrun M, Boulanger R, Lefort E, Nakamura S, Motoyoshi Y, Mieulet D, Campa C, Legendre L, Bertrand B. 2024. Combined sensory, volatilome and transcriptome analyses identify a limonene terpene synthase as a major contributor to the characteristic aroma of a *Coffea arabica* L. specialty coffee. *BMC Plant Biology* **24**: 238. DOI: <https://doi.org/10.1186/s12870-024-04890-3>
- Martínez G, Regente M, Jacobi S, Del Rio M, Pinedo M, de la Canal L. 2017. Chlorogenic acid is a fungicide active against phytopathogenic fungi. *Pesticide Biochemistry and Physiology* **140**: 30-35. DOI: <https://doi.org/10.1016/j.pestbp.2017.05.012>
- Martins V de C, da Silva MAE, da Veiga VF, Pereira HMG, de Rezende CM. 2025. Ent-Kaurane diterpenoids

- from *Coffea* genus: An update of chemical diversity and biological aspects. *Molecules* **30**: 59. DOI: <https://doi.org/10.3390/molecules30010059>
- Meyer FG. 1965. Notes on wild *Coffea arabica* from Southwestern Ethiopia, with some historical considerations. *Economic Botany* **19**: 136-151. DOI: <https://doi.org/10.1007/BF02862825>
- Molina D. 2022. Revisión sobre la broca del café, *Hypothenemus hampei* (Coleoptera: Curculionidae: Scolytinae) con énfasis en la resistencia mediante antibiosis y antixenosis. *Revista Colombiana de Entomología* **48**. DOI: <https://doi.org/10.25100/socolen.v48i2.11172>
- Molina D, Moncada-Botero M-P, Cortina-Guerrero HA, Benavides P. 2022. Searching for a Coffee variety with antibiosis effect to *Hypothenemus hampei* Ferrari (Coleoptera: Curculionidae). *Euphytica* **218**: 1-22. DOI: <https://doi.org/10.1007/s10681-022-03047-3>
- Monribot-Villanueva JL, Altúzar-Molina A, Aluja M, Zamora-Briseño JA, Elizalde-Contreras JM, Bautista-Valle MV, Arellano de los Santos J, Sánchez-Martínez DE, Rivera-Reséndiz FJ, Vázquez-Rosas-Landa M, Camacho-Vázquez C, Guerrero-Analco JA, Ruiz-May E. 2022. Integrating proteomics and metabolomics approaches to elucidate the ripening process in white *Psidium guajava*. *Food Chemistry* **367**: 130656. DOI: <https://doi.org/10.1016/j.foodchem.2021.130656>
- Montagnon C, Mahyoub A, Solano W, Sheibani F. 2021. Unveiling a unique genetic diversity of cultivated *Coffea arabica* L. in its main domestication center: Yemen. *Genetic Resources and Crop Evolution* **68**: 2411-2422. DOI: <https://doi.org/10.1007/s10722-021-01139-y>
- Nathanson JA. 1984. Caffeine and related methylxanthines: Possible naturally occurring pesticides. *Science* **226**: 184-187. DOI: <https://doi.org/10.1126/science.6207592>
- Noriega DD, Arias PL, Barbosa HR, Arraes FBM, Ossa GA, Villegas B, Coelho RR, Albuquerque EVS, Togawa RC, Grynberg P, Wang H, Vélez AM, Arboleda JW, Grossi-de-Sa, MF, Silva MCM, Valencia-Jiménez A. 2019. Transcriptome and gene expression analysis of three developmental stages of the coffee berry borer, *Hypothenemus hampei*. *Scientific Reports* **9**: 12804. DOI: <https://doi.org/10.1038/s41598-019-49178-x>
- Novaes F, Bayan F, Radler de Aquino Neto F, Resende C. 2019. The occurrence of Cafestol and Kahweol diterpenes in different coffee brews. *Coffee Science* **14**: 265. DOI: <https://doi.org/10.25186/cs.v14i2.1571>
- Oliva M, Rubio KB, Chinguel D, Carranza J, Bobadilla LG, Leiva S. 2023. Coffee berry borer infestation and population per fruit relationship with coffee variety, shade level, and altitude on specialty coffee farms in Peru. *International Journal of Agronomy* **2023**: 6782173. DOI: <https://doi.org/10.1155/2023/6782173>
- Osso-Bolaños JB, Perez-Santofimio EA. 2021. *Análisis de la calidad en taza y el factor de rendimiento de las variedades de café (Coffea arábica L.) Castillo y Bourbon Rosado, bajo las mismas condiciones de fermentación en la finca la primavera de la vereda Betania del municipio Pitalito-Huila*. BSc Thesis. Universidad Nacional Abierta y a Distancia.
- Pang Z, Lu Y, Zhou G, Hui F, Xu L, Viau C, Spigelman AF, MacDonald PE, Wishart DS, Xia J. 2024. MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation. *Nucleic Acids Research* **2024**: 1-9. DOI: <https://doi.org/10.1093/nar/gkae253>
- Ramvalho JC, Pais IP, Leitão AE, Guerra M, Reboredo FH, Máguas CM, Carvalho ML, Scotti-Campos P, Ribeiro-Barros AI, Lidon FJC, DaMatta FM. 2018. Can elevated Air [CO₂] conditions mitigate the predicted warming impact on the quality of Coffee bean? *Frontiers in Plant Science* **9**. DOI: <https://doi.org/10.3389/fpls.2018.00287>
- Roy S, Nuckles E, Archbold DD. 2018. Effects of phenolic compounds on growth of *Colletotrichum* spp. In Vitro. *Current Microbiology* **75**: 550-556. DOI: <https://doi.org/10.1007/s00284-017-1415-7>
- Schymanski E L, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, Hollender J. 2014. Identifying small molecules via high resolution mass spectrometry: Communicating confidence. *Environmental Science & Technology* **48**: 2097-2098. DOI: <https://doi.org/10.1021/es5002105>
- Setoyama D, Iwasa K, Seta H, Shimizu H, Fujimura Y, Miura D, Wariishi H, Nagai C, Nakahara K. 2013. High-Throughput metabolic profiling of diverse green *Coffea arabica* beans identified tryptophan as a universal discrimination factor for immature beans. *Plos One* **8**: e70098. DOI: <https://doi.org/10.1371/journal.pone.0070098>

- Silvestrini M, Junqueira MG, Favarin AC, Guerreiro-Filho O, Maluf MP, Silvarolla MB, Colombo CA. 2007. Genetic diversity and structure of Ethiopian, Yemen and Brazilian *Coffea arabica* L. accessions using microsatellites markers. *Genetic Resources and Crop Evolution* **54**: 1367-1379. DOI: <https://doi.org/10.1007/s10722-006-9122-4>
- Subroto G, Kusianto D, Avivi S, Slameto S, Setiyono S. 2019. Correlation between secondary metabolites of leaf and the resistance to leaf rust (*Hemileia vastatrix*) on several Arabica coffee clones. *Ilmu Pertanian (Agricultural Science)* **4**: 71. DOI: <https://doi.org/10.22146/ipas.42124>
- Tang X, Shao Y-L, Tang Y-J, Zhou W-W. 2018. Antifungal Activity of Essential Oil Compounds (Geraniol and Citral) and inhibitory mechanisms on grain pathogens (*Aspergillus flavus* and *Aspergillus ochraceus*). *Molecules* **23**: 2108. DOI: <https://doi.org/10.3390/molecules23092108>
- Therdthatha P, Jareontanahun N, Chaisuwan W, Yakul K, Paemane A, Manassa A, Moukamnerd C, Phimolsiripol Y, Sommano SR, Seesuriyachan P. 2023. Production of functional Arabica and Robusta green coffee beans: Optimization of fermentation with microbial cocktails to improve antioxidant activity and metabolomic profiles. *Biocatalysis and Agricultural Biotechnology* **53**: 102869. DOI: <https://doi.org/10.1016/j.bcab.2023.102869>
- Velásquez-Orozco RA. 2019. *Guía de variedades de café Guatemala*. Guatemala. Asociación Nacional del Café. <https://www.anacafe.org/uploads/file/9a4f9434577a433aad6c123d321e25f9/Gu%C3%A1-de-variedades-Anacaf%C3%A9.pdf> (accessed August 12, 2025).
- Wasternack C, Hause B. 2013. Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development. An update to the 2007 review in *Annals of Botany*. *Annals of Botany* **111**: 1021-1058. DOI: <https://doi.org/10.1093/aob/mct067>
- WCR [World Coffee Research]. 2024. Coffee varieties catalog. *World Coffee Research*. Website <https://worldcoffeeresearch.org/resources/coffee-varieties-catalog> (accessed 5 March 2025).
- Wei F, Furihata K, Koda M, Hu F, Kato R, Miyakawa T, Tanokura M. 2012. ¹³C NMR-Based metabolomics for the classification of green Coffee beans according to variety and origin. *Journal of Agricultural and Food Chemistry* **60**: 10118-10125. DOI: <https://doi.org/10.1021/jf3033057>
- Yasari A, Liang N, Foroutan A, Gänzle MG, Strelkov SE, Kav NNV. 2021. Investigating the potential of unsaturated fatty acids as antifungal crop protective agents, *Canadian Journal of Plant Science* **101**: 73-85. DOI: <https://doi.org/10.1139/cjps-2020-0113>
- Zanetti M, Ternus Z, Dalcanto F, de Mello M, de Oliveira D, Araujo P, Riella H, Fiori M. 2015. Microbiological characterization of pure Geraniol and comparison with bactericidal activity of the Cinnamic Acid in Gram-Positive and Gram-Negative bacteria. *Journal of Microbial & Biochemical Technology* **07**. DOI: <https://doi.org/10.4172/1948-5948.1000203>

Associate editor: Juan Rodrigo Salazar

Author contributions: IMP, material collection and processing, data processing, analysis of results and writing; RG, advise, writing and correction of the manuscript; OB, conceptual and experimental design, advise, writing and correction of the manuscript; JAGA, advise, writing and correction of the manuscript; LAIJ, conceptual and experimental design, advise, writing and correction of the manuscript; JLMV, conceptual and experimental design, LC-MS processing, data processing, analysis of results, writing and correction of the manuscript.

Supporting Agencies: IMP acknowledges SECIHTI for her PhD scholarship (CVU 1205430). We thank the authorities of the Instituto de Ecología, A.C. (INECOL) for providing financial support to maintain the UPLC-MS-QTOF platform.

Conflict of interests: The authors declare that there is no conflict of interest, financial or personal, in the information, presentation of data and results of this article.