



CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL

UNIDAD IRAPUATO
DEPARTAMENTO DE BIOTECNOLOGÍA Y BIOQUÍMICA

“Estudio integral metabolómico en plantas de *Coffea canephora* para el mejoramiento sensorial del café”

Tesis que presenta:
M.C. Roberto Gamboa Becerra

Para obtener el grado de:
Doctorado en Ciencias

En la especialidad en:
Biología de Plantas

Director de Tesis
Dr. Robert Winkler

El presente trabajo de tesis doctoral titulado “Estudio integral metabolómico en plantas de *Coffea canephora* para el mejoramiento sensorial del café” se realizó en el Laboratorio de Análisis Bioquímico e Instrumental del Departamento de Biotecnología y Bioquímica del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Irapuato, bajo la dirección del Dr. Robert Winkler y en colaboración con Agroindustrias Unidas de México (AMSA). Este proyecto se desarrolló gracias al apoyo del Consejo Nacional de Ciencia y Tecnología (CONACYT) y la beca 346955 otorgada para estudios de doctorado.

Este trabajo se realizó del 1 de septiembre de al 15 de noviembre de 2016

Miembros del comité de asesores:

Dra. Nayelli Marsch Martínez

Dr. Agustino Martínez Antonio

Dr. Charles Stewart Gillmor III

Dr. Jorge Molina Torres

Dr. Óscar González Ríos

Fecha de examen de grado: 24 de enero de 2017

Agradecimientos

Al Consejo Nacional de Ciencia y Tecnología (CONACyT), por la beca otorgada para la realización de mis estudios de doctorado (Beca No.346955).

Al centro de Investigación y de Estudios Avanzados (CINVESTAV) Unidad Irapuato, por el uso de su infraestructura para realizar mi proyecto de doctorado.

A la empresa Agroindustrias Unidas de México (AMSA) por permitirme realizar los muestreos en la Finca Badenia de Tapachula, Chiapas. Especialmente a Jenner Hernández quien me apoyó durante todas las visitas que realicé a la finca.

Al Dr. Robert por aceptarme y otorgarme un lugar en su laboratorio, y por permitir el desarrollo de este trabajo bajo su dirección.

A los integrantes de mi comité de asesores: Dra. Nayelli Marsch Martínez, Dr. Agustino Martínez Antonio, Dr. Charles Stewart Gillmor III, Dr. Jorge Molina Torres y Dr. Óscar González Ríos, por la contribución con sus opiniones y sugerencias durante el desarrollo de mi proyecto doctoral.

Al Dr. Óscar González Ríos y a la M.C. Ma. Cecilia Hernández Hernández por su ayuda y asesoría en parte del trabajo experimental de mi proyecto que realicé en su Laboratorio de Tecnología del café y el área de evaluación sensorial del Instituto Tecnológico de Veracruz.

A mis compañeros LABIs que forman o formaron parte del Laboratorio de Análisis Bioquímico e Instrumental: Abigail, Cristina, Daniel, Héctor, Josaphat, Janet, Mauricio, Sandra, Teresa, Janet, laboratorio con quienes compartí estos años de mi formación profesional.

A Dora E. Anguiano por su amable atención y la disponibilidad que siempre mostró para solucionar cualquier situación académica desde mi ingreso a CINVESTAV hasta el último trámite requerido para mi titulación.

A JuanPa, Sarita, Don Felipe y Don Juan, que se salen un momento de su rutina y con quienes compartimos pláticas amenas.

A mis amigos, Oli, Uriel, Hugo, Gus, Damaris, Nelly, Luis Jorge, Citlali, Crista, Carlos, Edith. Gracias por los momentos tan agradables que compartí con ustedes, en el ámbito personal, profesional y académico.

A mis amigos-paísanos: por dejarnos compartir momentos muy agradables con sus familias y que con su amistad me permitieron sentirme más cerca de mi querido Chiapas: Juan, Alice, Quique (Guanajuatense casi Chiapaneco¡¡), Susy, María y Marthita, muchas gracias.

A mi familia Gamboa y Ley, que estando lejos siempre transmiten su apoyo y proporcionan la energía para salir adelante de cualquier obstáculo.

Dedicatorias

A dos de los principales motores de mi vida: mi Abue y mi mamita. Les agradezco toda la felicidad que me han dado por amarme tanto durante toda mi vida. Su apoyo ha sido incondicional e interminable para todos los objetivos profesionales que me he propuesto, y éste no menos importante se los ofrezco como un logro suyo que se minimiza ante todo lo que ustedes se merecen. Las amo muchísimo;;

A la luz que aparece cuando la noche es tan oscura, a la mujercita que día a día vive una batalla conmigo para animarme a enfrentar todo tipo de obstáculos que se presentan en nuestras vidas, a la que muchas veces literalmente me soluciona la vida;;. A la mujer que amo y que llena día a día mis deseos de vivir, y alimenta la esperanza de mi alma con la dicha de tenerla siempre a mi lado. A ti Dianita hermosa te dedico este logro tuyo y mío. Gracias por impulsarme a que este día llegara. TE AMO;;

A quien algunas ocasiones pienso desaparece de mi vida, pero siempre me demuestra que ha estado ahí para mostrarme el camino, y que sobre todo me ha dado experiencias buenas y malas para aprender de ellas, te agradezco enormemente DIOS;;

ÍNDICE GENERAL

ÍNDICE FE FIGURAS.....	iv
ÍNDICE DE TABLAS.....	vi
RESUMEN	vii
ABSTRACT	ix
1. ANTECEDENTES.....	1
1.1 El cultivo de café.....	1
1.1.1 Importancia del café	1
1.1.2 Problemática del café en México	2
1.1.3 Características del cultivo	3
1.2 Desarrollo del fruto de café	5
1.3 Composición química del grano verde	7
1.4 Composición química del grano tostado	8
1.4.1 Reacciones que ocurren durante el tostado	9
1.5 Café: la bebida	12
1.6 Calidad en taza del café.....	14
1.6.1 Calidad en taza y composición química.....	15
1.7 Mejoramiento genético tradicional de café.....	17
1.7.1 Heredabilidad genética	17
1.7.2 Métodos y técnicas de mejoramiento genético.....	19
1.7.3 Propagación de plantas de café	22
1.8 Fuentes genómicas del café	23
1.8.1 Genoma del café	23
1.9 Mejoramiento tradicional de la calidad en taza del café.....	24
1.9.1 Mejoramiento tradicional de <i>C. arabica</i>	24
1.9.2 Mejoramiento de calidad en <i>C. canephora</i>	26
1.10 Mejoramiento asistido por marcadores moleculares	28
1.10.1 Marcadores moleculares	28
1.10.2 Análisis de identificación de QTLs y mapeo genético.....	29
1.11 Selección de plantas asistida por metabolómica: herramienta emergente	30
2. JUSTIFICACIÓN.....	34
3. HIPÓTESIS.....	35
4. OBJETIVOS.....	35
4.1 Objetivo general	35
4.2 Objetivos específicos	35

5. METODOLOGÍA	36
5.1 Estrategia general	36
5.2 Lugar de muestreo	37
5.3 Análisis de desarrollo del fruto	37
5.3.1 Muestreo.....	37
5.3.2 Procesamiento de las muestras.....	38
5.3.3 Extracción de metabolitos.....	38
5.3.4 Análisis metabólico	39
5.3.5 Condiciones del espectrómetro de masas	39
5.4 Análisis de heredabilidad en hojas y frutos	40
5.4.1 Diseño experimental y muestreo	40
5.4.2 Procesamiento de las muestras y análisis metabólico.....	41
5.4.3 Estimación de heredabilidad.....	41
5.5 Análisis de correlación de calidad sensorial del café y perfiles químicos.....	42
5.5.1 Muestras.....	42
5.5.2 Evaluación sensorial de la bebida	42
5.5.4.1 Proceso de tostado del café.....	42
5.5.4.2 Preparación de la bebida	43
5.5.4.3 Análisis sensorial descriptivo de la bebida de café	43
5.5.3 Extracción de metabolitos.....	44
5.5.4 Análisis metabólico	44
6. RESULTADOS	45
6.1 Desarrollo del fruto	45
6.1.1 Evaluación de perfiles metabólicos durante el desarrollo del fruto	46
6.1.2 Selección de metabolitos predictivos de estadio de desarrollo.....	48
6.1.3 Cianidin-3,5-diglucósido y cianidin-3-O-(6-O-p-coumaroil) como marcadores predictivos de estadio de desarrollo del fruto de café	64
6.2 Análisis de Heredabilidad metabólica	65
6.2.1.1 Variabilidad de fenotipos metabólicos y parentesco entre plantas madre	67
6.2.1.2 Estimación de heredabilidad en frutos y hojas	70
6.3 Correlación de perfiles sensoriales de la bebida-metabolitos altamente heredables de frutos y hojas.....	72
6.3.1 Evaluación sensorial.....	72
6.3.2 Análisis metabólico e identificación de metabolitos marcadores de calidad sensorial en granos tostados de café	76
6.4 Selección de metabolitos marcadores de selección temprana	80

7. DISCUSIÓN.....	86
8. CONCLUSIONES.....	90
9. PERSPECTIVAS.....	91
10. BIBLIOGRAFÍA.....	92
11. ANEXOS.....	106
11.1 Formato de evaluación sensorial descriptivo de la bebida de café.....	106
11.2 Escala de 6 puntos para evaluación de atributos sensoriales.....	107
11.3 Presentación de bebidas de café para su degustación.....	107
11.4 Metabolitos de frutos con niveles altos de heredabilidad (≥ 0.4).....	108
11.5 Metabolitos de granos tostados correlacionados con atributos sensoriales de la bebida de café.....	119
11.6 Lista de publicaciones y participaciones en congresos.....	127

ÍNDICE FE FIGURAS

Figura 1. Principales productores de café a nivel mundial.....	1
Figura 2. Producción mundial de café Arábica y Robusta.....	3
Figura 3. Esquema de las etapas de desarrollo del fruto de café.....	5
Figura 4. Principales precursores del aroma en el grano de café.....	10
Figura 5. Composición química típica de 100 mL de bebida.....	13
Figura 6. Diagrama esquemático de la generación de la variedad Oro Azteca.....	19
Figura 7. Cronograma de muestreo de etapas del desarrollo del fruto.....	38
Figura 8. Diseño de medios hermanos maternos para análisis de heredabilidad.....	40
Figura 9. Estadios muestreados para análisis de desarrollo del fruto.....	45
Figura 10. PCA de perfiles metabólicos durante el desarrollo del fruto.....	47
Figura 11. Ranking de importancia de metabolitos de frutos que presentan mayor diferencia durante el desarrollo del fruto.....	50
Figura 12. Ranking de importancia de metabolitos de hojas que presentan mayor diferencia durante el desarrollo del fruto.....	51
Figura 13. Clasificación de metabolitos seleccionados basada en KEGG.....	53
Figura 14. Heatmap de los niveles de metabolitos detectados en frutos durante el desarrollo del fruto.....	54
Figura 15. Heatmap de los niveles de metabolitos detectados en hojas durante el desarrollo del fruto.....	55
Figura 16. Cambios de los niveles de 5-O-p-coumaroilquinato durante el desarrollo del fruto de café.....	56
Figura 17. Cambios de los niveles de galactosa durante el desarrollo del fruto de café	57
Figura 18. Cambios en los niveles de ácidos grasos durante el desarrollo del fruto de café.....	58
Figura 19. Cambios de los niveles de sirohidroclorina durante el desarrollo del fruto de café.....	59
Figura 20. Cambios de los niveles de D-mio-inositol-1,2,5,6 tetrakisfosfato durante el desarrollo del fruto de café.....	60
Figura 21. Cambios en los niveles de carotenoides durante el desarrollo del fruto de	

café.....	61
Figura 22. Cambios en los niveles de compuestos involucrados en metabolismo y biosíntesis de terpenoides durante el desarrollo del fruto.....	62
Figura 23. Cambios en los niveles de compuestos involucrados en la biosíntesis de fenilpropanoides, flavonoides y antocianinas durante el desarrollo del fruto.....	63
Figura 24. Cambios en los niveles de compuestos marcadores de desarrollo del fruto...	64
Figura 25. Dendograma de 40 plantas madre.....	68
Figura 26. Dendograma de progenie F1.....	69
Figura 27. Histograma de heredabilidad de metabolitos estimada en frutos y hojas.....	71
Figura 28. Gráfica de radar de perfiles sensoriales.....	74
Figura 29. Análisis de componentes principales (PCA) de perfiles sensoriales.....	75
Figura 30. Contribución de variables a los componentes 1 y 2.....	75
Figura 31. Gráfico de correlación entre caracteres sensoriales.	77
Figura 32. Metabolitos que mostraron alta correlación atributos sensoriales de la bebida.....	78
Figura 33. Gráfico de correlación entre m/z 727.9535 y los atributos sensoriales de la bebida.....	79
Figura 34. Metabolitos altamente heredables ($H^2 > 0.4$) correlacionados con atributos sensoriales de la bebida.....	80
Figura 35. Heat map de correlaciones de metabolitos altamente heredables en hojas vs atributos sensoriales de la bebida.....	83

ÍNDICE DE TABLAS

Tabla 1. Características y requerimientos de la planta de café.....	4
Tabla 2. Estudios reportados del desarrollo y maduración del fruto de café.....	6
Tabla 3. Composición química de café verde expresada en porcentaje de materia seca...8	
Tabla 4. Composición química de café tostado.....	11
Tabla 5. Compuestos relacionados con caracteres sensoriales de la bebida.....	15
Tabla 6. Tipo de herencia y heredabilidad de caracteres determinantes de calidad en taza.....	27
Tabla 7. Genes reguladores de compuestos determinantes de calidad en taza.....	29
Tabla 8. Análisis de QTLs relacionados con calidad en taza del café.....	30
Tabla 9. Gradiente de solventes para frutos.....	39
Tabla 10. Gradiente de solventes para hojas.....	39
Tabla 11. Matriz de error de modelo predictivo de bosques aleatorios generado con perfiles metabólicos de frutos.....	48
Tabla 12. Matriz de error de modelo predictivo de bosques aleatorios generado con perfiles metabólicos de hojas.....	49
Tabla 13. Familias incluidas en el análisis de heredabilidad.....	66
Tabla 14. Perfil sensorial descriptivo de las bebidas.....	73
Tabla 15. Metabolitos detectados en frutos para selección de plantas con características sensoriales deseables.....	81
Tabla 16. Metabolitos detectados en hojas para selección de plantas con características sensoriales deseables.....	84

RESUMEN

El café es el segundo producto más comercializado a nivel mundial, *Coffea arabica* y *C. canephora* son las dos especies más cultivadas. México es uno de los principales exportadores y el líder mundial en producción de café orgánico. La producción de plantas con características de calidad mejoradas es uno de los objetivos principales de los programas de mejoramiento genético del café. Debido a que el desarrollo de los árboles de café es lento, los métodos convencionales toman varias generaciones y tiempos largos. Por esta razón, existe la necesidad de nuevas estrategias que permitan la selección de variedades de café con características agronómicas y de calidad deseables. Por ello, desarrollamos un enfoque de metabolómica no dirigida para identificar metabolitos altamente heredables asociados con la calidad en taza del café en estadios tempranos del desarrollo de plantas de *C. canephora*. Primero, considerando que el nivel de madurez de los granos de café influencia la calidad en taza del café y desempeña un papel importante en el control de calidad de los productos de café, determinamos el perfil metabólico de seis estadios de desarrollo de frutos de café. Mediante un modelo predictivo de random forest fue posible identificar en frutos el compuesto cianidin-3-O- (6-Op-coumaroil) glucósido como un marcador predictivo de estadio de desarrollo del fruto. El cianidin-3,5-diglucósido detectado en hojas fue un factor discriminante entre frutos de estadios tempranos del desarrollo y frutos de café semimaduros-maduros. Estos compuestos podrían ser empleados en análisis de rutina para el control de calidad de productos de café. Por otra parte, se determinó la heredabilidad de metabolitos empleando un enfoque de medios hermanos maternos en el que se incluyeron cuarenta plantas madre y ochenta medios hermanos de plantas *C. canephora*. En el análisis se incluyeron hojas de estadios tempranos del desarrollo de la planta y frutos. Se identificaron 450 y 116 metabolitos altamente heredables ($H^2 > 0.4$) en hojas y frutos, respectivamente. Estos metabolitos se encontraron principalmente influenciados por el factor genético y correlacionaron con diez atributos sensoriales de la bebida: acidez, resabio, intensidad aromática, calidad aromática, astringencia, cuerpo, amargor, sabor, agrio y preferencia. 164 metabolitos altamente heredables presentes en frutos de café predicen más de un atributo sensorial. En hojas, se detectaron 234 asociaciones entre 45 metabolitos y los atributos sensoriales de la bebida. Por lo tanto, nuestra estrategia propuesta representa una alternativa para la selección temprana de plantas de café con

características deseables. El enfoque de selección basado en metaboloma completo (metabolome-wide selection) permite acortar el tiempo de selección comparado con las estrategias convencionales de mejoramiento de café. Esta nueva estrategia basada en marcadores metabólicos tiene potencial para ser empleada en el mejoramiento de otros cultivos agronómicamente importantes.

ABSTRACT

Coffee is the second most valuable traded commodity around the world, with *Coffea arabica* and *C. canephora* as the two main cultivated species. Mexico is one of the main exporters and the globally leading producer of organic coffee. Production of plants with improved coffee quality attributes is a main target of coffee genetic improvement programs. Since the development of the tree is slow, conventional breeding with several generations is time consuming. For this reason, there is a need for new strategies to enable the selection of coffee plants varieties with desirable agronomic and sensory traits. Thus, we developed an un-targeted metabolomic approach to identify highly heritable metabolites linked to coffee cup quality in early development stages of *C. canephora* plants. First, considering that maturity level of coffee bean is influencing the coffee cup quality and plays a role in quality control of coffee products, we determined the metabolic profile of six development stages of coffee fruits. Employing a predictive random forest model it was possible to identify in fruits the compound cyanidin-3-O- (6-O-p-coumaroyl) glucoside as a predictive marker of fruit development stage. Cyanidin-3,5-diglucoside found in leaves was a discriminant factor between fruits at early stages of development and semimature-ripe coffee fruits. These compounds could be employed in routine analysis for quality control of coffee products. For another hand, we determined the metabolites heritability employing a maternal half sibs approach through forty mother and eighty half sibs of *C. canephora* plants. Leaves at early plant development stage and fruits were included in the analysis. We identified 450 and 116 highly heritable metabolites ($H^2 > 0.4$) in fruits and leaves, respectively. These metabolites were mainly influenced by genetic factors and correlated with ten sensory attributes of coffee beverage: acidity, aftertaste, aromatic intensity, aromatic quality, astringency, body, bitterness, flavor, sourness and preference. 164 highly heritable metabolites in coffee fruits predict more than one sensory trait. 234 associations between 45 metabolites and the sensory traits of beverage were detected in leaves. Therefore, our proposed strategy in this project represents an alternative for the early selection of coffee plants with desirable characteristics. The “metabolome-wide” selection approach allows to shorten the time compared to conventional breeding strategies for coffee. Further this new metabolic-marker based strategy has the potential to be employed in the improvement of other agronomically relevant cultivars.

1. ANTECEDENTES

1.1 El cultivo de café

1.1.1 Importancia del café

El café es uno de los productos agrícolas más importantes a nivel mundial. La infusión preparada a partir de sus granos tostados y molidos es una de las bebidas más populares consumidas en el mundo. La planta de café probablemente tuvo su origen en África, los reportes más antiguos acerca de este cultivo provienen de Etiopía (Oestreich-Janzen, 2010).

El café se produce en todos los continentes y es el segundo producto más comercializado, únicamente superado por el petróleo (Gielessen & Graafland, 2009; Davis *et al.*, 2012). En 2015-2016, se produjeron más de 8.6 millones de toneladas en todo el mundo, siendo el continente americano el que aporta el mayor porcentaje (55.12 %). Brasil encabeza la lista de los 10 principales productores y exportadores de café (34.15 % de la producción total) (Fig. 1), México se encuentra dentro de esta lista produciendo alrededor de 3 %, lo cual corresponde más de 234,000 toneladas de café por año (Organización internacional del café, 2016).

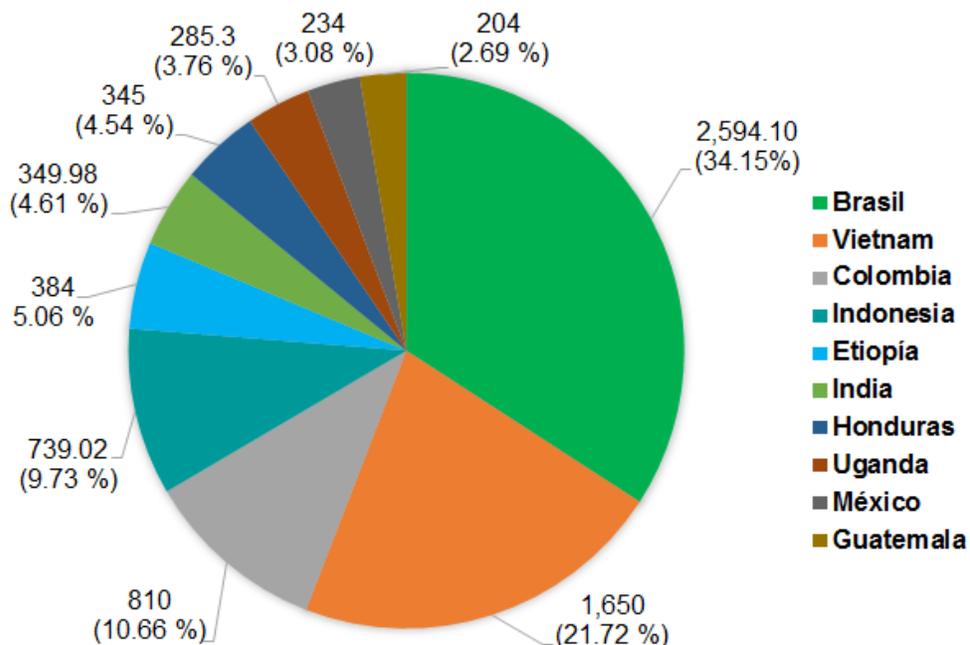


Figura 1. Principales productores de café a nivel mundial. Los datos corresponden a la producción 2015/2016 expresada en miles de toneladas. Elaborada con datos de (Organización internacional del café, 2016).

La producción de café en México se encuentra distribuida en 13 estados, sin embargo Chiapas, Veracruz, Puebla y Oaxaca destacan como los principales productores nacionales y aportan más del 88 % de la producción. México ocupa el segundo lugar en producción de café orgánico a nivel mundial (SAGARPA, 2016).

Se estima que en México, la actividad cafetalera provee de empleo (directo e indirecto) a más de tres millones de personas. Actualmente en México existen 511,669 productores que en total cuentan con una superficie de producción de 697,366 hectáreas (SAGARPA, 2016).

1.1.2 Problemática del café en México

La producción de café ha disminuido de 4.5 millones de sacos (60 Kg c/u) hasta 2.3 millones de sacos en 2015/2016. Esta cifra ha superado la demanda de café en México, que corresponde a aproximadamente 5.2 millones de sacos, 2.7 para consumo interno y 2.5 millones para exportación. Se espera que para la cosecha 2016/2017 el déficit de café aumente debido a que la demanda interna y el valor de las exportaciones ha crecido 30 % en los últimos años (CAFÉS DE MÉXICO, 2016; SAGARPA, 2016).

La problemática del sector cafetalero en México involucra aspectos políticos, tecnológicos, económicos, socioculturales y ecológicos. Las razones por las que la producción del café en México ha sido afectada incluyen: un sector productivo sin integración de mecanismos de mercado, carencia de redes de investigación, falta de capacitación y asistencia técnica, falta de ampliación de nichos de exportación de cafés diferenciados. En términos ecológicos, los periodos de lluvias a destiempo y exceso de humedad son condiciones que han favorecido la expansión de la roya de café. Otros factores han sido la falta de buenas prácticas agrícolas por parte de los agricultores, falta de fertilización de suelos, predominancia de plantaciones viejas, así como el aumento en el costo de producción (López Morgado *et al.*, 2013; CAFÉS DE MÉXICO, 2016; GAIN Report, 2016).

Aunado a la crisis que se vive en México, se ha presentado una fuerte disminución de los precios en la bolsa de café de Nueva York (Dirección de Investigación y Evaluación Económica y Sectorial, 2015; CAFÉS DE MÉXICO, 2016; USDA, 2016)

1.1.3 Características del cultivo

El café pertenece a la familia Rubiaceae y al género *Coffea*, más de 100 especies pertenecen a este género. La producción comercial de café a nivel mundial está liderada por dos especies: *Coffea arabica* y *Coffea canephora* (58.79 % y 41.21 %, respectivamente). En los últimos años la producción de *C. canephora* ha recobrado importancia ya que, para el año 2005 la producción fue 30 % del total de producción de café a nivel mundial, sin embargo para la cosecha 2015-2016 se produjeron 3.546 millones de toneladas correspondientes a 41.21 % de la producción mundial de café (Fig. 2) (Organización internacional del café, 2016). En México de la superficie total de producción (697,366 hectáreas), en el 97 % se cultiva *C. arabica* mientras que el restante 3 % corresponde a *C. canephora* (variedad robusta), principalmente destinada a la industria de café soluble (SAGARPA, 2016). Estos datos contrastan con el consumo nacional, ya que de acuerdo a la USDA en 2014/2015 el 78 % del consumo en México correspondió a café soluble mientras que únicamente el 22 % fue de café tostado y molido (USDA, 2016).

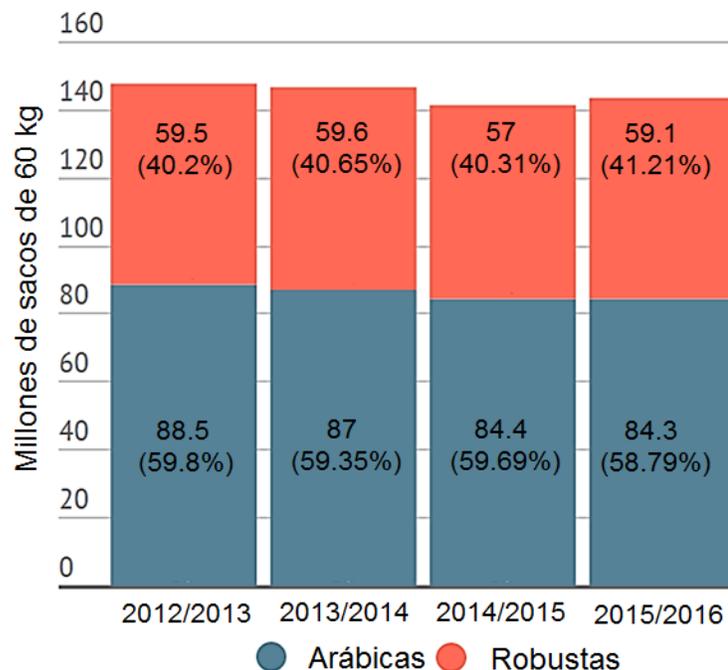


Figura 2. Producción mundial de café Arábica y Robusta. Los datos corresponden a la producción de las cosechas 2012/2013 - 2015/2016. Modificada de Organización internacional del café 2016.

C. arabica y *C. canephora* presentan diferencias importantes en las características y requerimientos para su cultivo (Tabla 1). Todas las especies de café conocidas son diploides y por lo general auto-incompatibles, *C. arabica* es la excepción, ésta es tetraploide auto-fértil y se considera producto de una hibridación espontánea entre *C. canephora* y *C. eugenioides* (Lashermes *et al.*, 1999; Charrier & Eskes, 2008).

Tabla 1.- Características y requerimientos de la planta de café.

Característica	<i>C. arabica</i>	<i>C. canephora</i>
Genoma	Tetraploide 44 cromosomas	Diploide 22 cromosomas
Ecología	Bosque húmedo tropical Altitud: 1250 – 2200 m Temperatura: 15 – 24 °C Precipitación: 1200 – 2200 mm/año	Bosque húmedo tropical 250 – 1500 m 18 – 36 °C 2200 – 3000 mm /año
Altura del árbol	4 – 6 m	8 – 12 m
Hojas	Longitud: 10 – 15 cm; Ancho: 4 – 6 cm	22 cm 17 cm
Raíz	Profunda	Poco profunda
Inflorescencias	16 – 48 flores por nodo	30 – 100 flores por nodo
Reproducción	Autógama (autopolinización)	Alógama (polinización cruzada)
Maduración del fruto	7 – 9 meses post-antesis	9 – 11 meses post-antesis

Tomada y modificada de (Oestreich-Janzen, 2010).

Se ha considerado que *C. arabica* produce una bebida de mejor calidad en taza comparada con *C. canephora* y debido a la importancia económica de ambas especies, éstas han sido el objeto principal de la mayoría de los programas de mejoramiento genético del café (DaMatta *et al.*, 2007; P Tshilenge, 2009; Organización internacional del café, 2016).

1.2 Desarrollo del fruto de café

La planta de café toma aproximadamente 3 años para desarrollarse desde la germinación de la semilla hasta su primera floración y producción de fruto. El desarrollo del fruto de café comprende desde el registro de la antesis (floración) hasta la maduración completa del fruto. En *C. arabica* este proceso toma de 7 a 9 meses, mientras que para *C. canephora* este dura entre 9 y 11 meses. El crecimiento del fruto no es sincronizado y es posible observar diferentes tamaños de frutos en la misma planta (De Castro & Marraccini, 2006; Eira *et al.*, 2006).

El proceso inicia con el desarrollo del perispermo y la división de las células del endospermo (Fig. 3). Los estadios de fruto en forma de “cabeza de alfiler” y expansión son los siguientes en esta fase inicial de desarrollo. En el estadio de fruto verde, se observa elongación de células del endospermo y la pérdida gradual de tejido del perispermo.

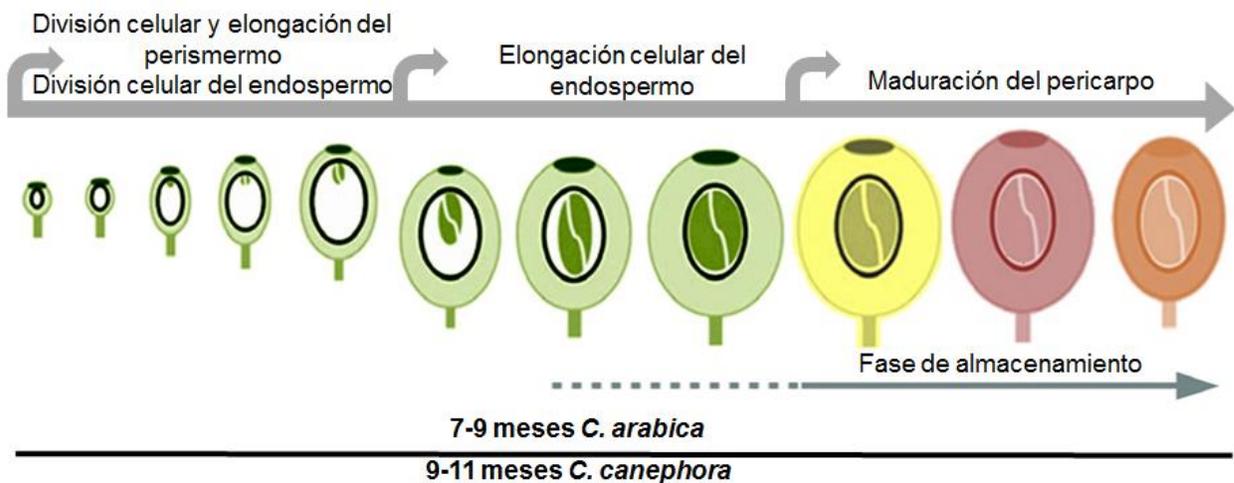


Figura 3. Esquema de las etapas de desarrollo del fruto de café. Tomada y modificada de (De Castro & Marraccini, 2006).

Durante la fase final, correspondiente al estadio “amarillo-verde” a cereza, se lleva a cabo la maduración del pericarpio, caracterizado por el endurecimiento del endospermo, acumulación gradual de proteínas de almacenamiento, síntesis de sacarosa y carbohidratos complejos, pigmentos, y degradación de clorofila. El último paso de la maduración se caracteriza por la deshidratación del endospermo y el cambio de coloración del pericarpio. Importantes cambios tanto cuantitativos como cualitativos son

observados durante el crecimiento del fruto, los cuales se ven reflejados en las características y la composición química final del grano de café (De Castro & Marraccini, 2006; Eira *et al.*, 2006; DaMatta *et al.*, 2007).

A la fecha, existen múltiples estudios enfocados en la caracterización de los cambios a nivel genómico, metabolómico y proteómico que explican la fenología y los procesos de regulación que ocurren durante el desarrollo del fruto de café. Con el desarrollo de tecnologías de alto rendimiento sobre todo en el área de la genómica puede ser posible conocer cientos de genes y las complejas redes transcripcionales que orquestan los procesos fisiológicos durante el desarrollo del fruto.

Tabla 2. Estudios reportados del desarrollo y maduración del fruto de café.

Tipo de abordaje	Especie	Compuestos estudiados	Referencia
Proteómico	<i>C. arabica</i>	Múltiples ^a	(Marraccini <i>et al.</i> , 2001)
Genómico	<i>C. canephora</i>	ÁCGs	(Lepelley <i>et al.</i> , 2007)
	<i>C. arabica</i> , <i>C. canephora</i>	ÁCGs	(Koshiro <i>et al.</i> , 2007)
	<i>C. arabica</i>	Múltiples ^b	(Salmona <i>et al.</i> , 2008)
	<i>C. arabica</i> , <i>C. canephora</i>	Sacarosa	(Privat <i>et al.</i> , 2008)
	<i>C. arabica</i>	Múltiples ^d	(Joet <i>et al.</i> , 2009)
	<i>C. arabica</i> , <i>C. canephora</i>	Múltiples ^e	(Mondego <i>et al.</i> , 2011)
	<i>C. arabica</i>	Múltiples ^f	(Gaspari-Pezzopane <i>et al.</i> , 2012)
Metabólico y genómico	<i>C. arabica</i>	Sacarosa	(Geromel <i>et al.</i> , 2006)
	<i>C. arabica</i> <i>C. canephora</i>	Trigonelina y cafeína	(Koshiro <i>et al.</i> , 2006)
	<i>C. arabica</i> <i>C. canephora</i>	Carotenoides	(Simkin <i>et al.</i> 2008)
Metabólico	<i>C. arabica</i>	Monosacáridos y galactomananos	(Redgwell <i>et al.</i> , 2003)
	<i>C. arabica</i>	Cafeína	(Ashihara 2006)
	<i>C. arabica</i>	Múltiples ^c	(Amorim <i>et al.</i> , 2009)
	<i>C. arabica</i> <i>C. canephora</i>	Azúcares y ACGs	(Koshiro <i>et al.</i> , 2015)
	<i>C. arabica</i>	Cafeína, sacarosa, ACGs	(Smrke <i>et al.</i> , 2015)

ACGs: ácidos clorogénicos; a: Análisis de proteínas solubles de perispermio y endospermio; b: Análisis de 266 genes candidato de múltiples vías; c: Análisis global de iones obtenidos por ESI(-)-MS; d: Análisis de 137 genes seleccionados; e: Análisis de 187,142 ESTs de *C. arabica* y 78,470 de *C. canephora*; f: Análisis de 28 genes de vías metabólicas diferentes.

Los reportes que hasta la fecha existen se resumen en la tabla 2, principalmente se han enfocado en el estudio de metabolitos específicos y sus rutas de biosíntesis. Dentro de éstos se incluyen ácidos clorogénicos, azúcares (sacarosa, monosacáridos), trigonelina, cafeína, monosacáridos y galactomananos.

1.3 Composición química del grano verde

Se denomina grano verde a los frutos de café tratados post-cosecha que están listos para ser tostados. La composición química básica del grano verde que determina las propiedades y la calidad en taza del café depende primordialmente de aspectos genéticos (especie, variedad) y de aspectos fisiológicos tales como el proceso y el grado de maduración. Sin embargo es bien sabido que estas características intrínsecas son influenciadas por otros factores extrínsecos como condiciones de cultivo (composición del suelo, condiciones climáticas, altitud), edad de la planta, prácticas agrícolas (sombra, fertilización e irrigación), condiciones de cosecha, procesamiento post-cosecha y condiciones de almacenamiento que finalmente afectan la fisiología del grano (Farah, 2012).

La composición química del café ha sido estudiada extensamente en granos verdes (Clifford, 2000; Bertrand *et al.*, 2006; Arya & Rao, 2007; Clarke & Vitzthum, 2008; Fujioka & Shibamoto, 2008; Farah, 2012), los principales componentes se muestran en la tabla 3.

La fracción no volátil del café verde está principalmente constituida por agua, carbohidratos y fibra, proteínas y aminoácidos libres, lípidos, minerales, ácidos orgánicos, ácidos clorogénicos (ACGs), trigonelina y cafeína (Tabla 3) (Farah & Donangelo, 2006). En cantidades mucho menores o como trazas, se han identificado compuestos fenólicos menores como antocianinas y lignanos, así como metabolitos de cafeína como teofilina y teobromina (Mazzafera *et al.*, 2009).

Por otra parte, aproximadamente 200 compuestos volátiles han sido identificados en café verde (Holscher & Steinhart, 1995; Cantergiani *et al.*, 2001; Flament & Bessièrre-Thomas, 2002; Gonzalez-Rios *et al.*, 2007). Algunos de estos compuestos permiten diferenciar entre café Arabica y Robusta (Mathieu *et al.*, 1998).

Las clases más abundantes de compuestos volátiles en grano verde son alcoholes, ésteres, hidrocarburos, aldehídos, cetonas, pirazinas, furanos, y compuestos azufrados

(Gonzalez-Rios *et al.*, 2007; Toci & Farah, 2008).

Tabla 3.- Composición química de café verde expresada en porcentaje de materia seca.

Componente	Concentración (g/100g)	
	<i>C. arabica</i>	<i>C. canephora</i>
Carbohidratos	6-9	0.9-4
Sacarosa	0.1	0.4
Azúcares reductores	34-44	48-55
Polisacáridos	3	3
Lignina	3	3
Pectina	2	2
Proteína/péptidos	10-11	11-15
Aminoácidos libres	0.5	0.8-1
Cafeína	0.9-1.3	1.5-2.5
Trigonelina	0.6-2	0.6-0.7
Lípidos (triglicéridos/esteroles/tocoferoles)	15-17	7-10
Diterpenos (libres y esterificados)	0.5-1.2	0.2-0.8
Minerales	3-4.2	4.4-4.5
Ácidos clorogénicos	4.1-7.9	6.1-11.3
Ácidos alifáticos	1	1
Ácido quínico	0.4	0.4
Volátiles	Trazas	Trazas

Tomada de (Farah, 2012).

1.4 Composición química del grano tostado

Los granos verdes de café carecen de color y de las propiedades aromáticas que observamos en el grano tostado, ya que estas características se adquieren mediante reacciones químicas complejas que ocurren durante el proceso de tostado (Baggenstoss *et al.*, 2008).

Se sabe que los volátiles principalmente contribuyen al aroma mientras que los no volátiles están relacionados con amargor, acidez y astringencia (Buffo & Cardelli-Freire, 2004). El proceso de tostado se divide en tres fases:

A) Fase inicial de evaporación: es un proceso endotérmico en el cual se elimina la humedad, dan inicio los cambios de olor y el color de los granos se torna amarillento, los cambios químicos son relativamente pocos comparado con la siguiente fase.

B) Fase real de tostado: se caracteriza por un conjunto de reacciones pirolíticas, como las reacciones de Maillard y Strecker que involucran carbohidratos (azúcares reductores),

proteínas, compuestos de bajo y alto peso molecular tales como las melanoidinas, que son simultáneamente degradadas y producidas (Buffo & Cardelli-Freire, 2004; Farah, 2012). Por lo tanto, la composición química de los granos se modifica drásticamente, se liberan grandes cantidades de dióxido de carbono y se forman cientos de compuestos asociados con aroma y sabor. Los granos cambian su coloración a café oscuro. Este proceso es exotérmico con liberación de compuestos volátiles a temperaturas entre 190 y 210 °C (Buffo & Cardelli-Freire, 2004; Clarke & Vitzthum, 2008).

C) Fase rápida de enfriamiento: para detener las reacciones exotérmicas del proceso de tostado (Clarke & Vitzthum, 2008).

Todo el proceso es completamente dependiente de la temperatura y el tiempo de tostado, el color final del grano está directamente correlacionado con el grado de tostado, normalmente descrito como claro, medio u oscuro. El tiempo de tostado influye en las reacciones que ocurren dentro del grano, por ejemplo periodos largos producen un café más amargo pero éste carece de aroma satisfactorio. Periodos muy cortos resultan insuficientes tanto para que las reacciones pirolíticas se lleven a cabo, así como para que las características organolépticas sean completamente desarrolladas (Flament & Bessièrre-Thomas, 2002; Sunarharum *et al.*, 2014).

1.4.1 Reacciones que ocurren durante el tostado

Los mecanismos que ocurren durante el proceso de tostado, así como la formación de aroma del café a partir de compuestos precursores presentes en el grano de café son muy complejos (Ribeiro *et al.*, 2009). Durante este proceso se lleva a cabo un amplio rango de reacciones (Fig. 4) que incluyen:

A) Reacciones de Maillard: son reacciones de condensación entre compuestos que contienen nitrógeno (proteínas, péptidos, aminoácidos, serotonina y trigonelina) y azúcares reductores, hidroxiaácidos y fenoles, para la formación de aminoaldosas y aminocetonas (Baggenstoss *et al.*, 2008).

B) Degradación de Strecker: se refiere a reacciones entre un aminoácido y un α -dicarbonil con la formación de una aminocetona que se condensa para formar compuestos heterocíclicos nitrogenados o reaccionan con formaldehído para formar oxazoles (Viani & Illy, 2005).

C) Rompimiento de aminoácidos sulfurados, los cuales son transformados en

mercaptanos, tiofenos o tiazoles, después de reaccionar con azúcares reductores o productos intermediarios de reacciones de Maillard.

D) Rompimiento de hidroxiaminoácidos capaces de reaccionar con sacarosa y formar alkilpirazinas.

E) Rompimiento de prolina e hidroxiprolina para reaccionar con intermediarios de productos de Maillard, dando lugar a la formación de piridinas, pirroles, pirrolizinas así como alquil, acil y furfurylpirroles (Farah, 2009).

F) Degradación de trigonelina para formar alquil-piridinas y pirroles.

G) Degradación de ácido quínico para la formación de fenoles.

H) Degradación de pigmentos, principalmente carotenoides.

I) Degradación de lípidos, principalmente diterpenos (Buffo & Cardelli-Freire, 2004).

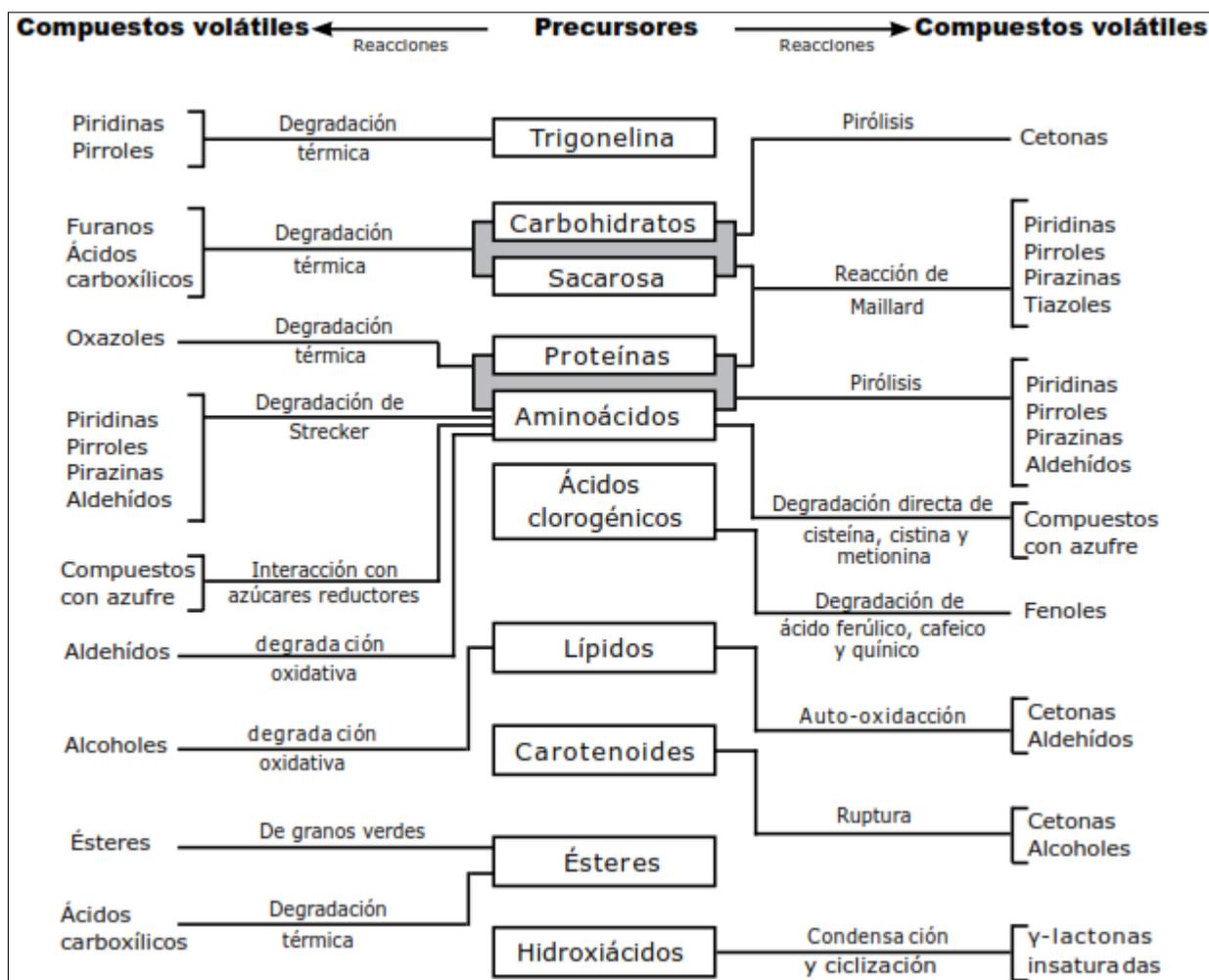


Figura 4. Principales precursores del aroma en el grano de café. Tomada y modificada de (Ribeiro *et al.*, 2009).

Como anteriormente se mencionó, los compuestos no volátiles sufren modificaciones para dar lugar a la formación de la fracción volátil durante el proceso de tostado, sin embargo no son totalmente consumidos durante las reacciones químicas que se presentan y en los granos tostados encontramos presentes: carbohidratos (sacarosa, azúcares reductores, polisacáridos de pared celular, lignina), compuestos nitrogenados (proteínas, aminoácidos libres, cafeína, trigonelina), lípidos, minerales (potasio y fósforo), ácidos y ésteres (ácidos clorogénicos, ácidos alifáticos y ácido quínico) (Tabla 4). Este grupo de compuestos son considerados como precursores de las características sensoriales que los granos adquieren durante el proceso de tostado (Flament & Bessièrre-Thomas, 2002; Farah, 2009).

Tabla 4.- Composición química de café tostado.

Componente	Concentración (g/100g)	
	<i>C. arabica</i>	<i>C. canephora</i>
Carbohidratos/fibra		
Sacarosa	4.2	1.6
Azúcares reductores	0.3	0.3
Polisacáridos (arabinolactano, manano y glucano)	31-33	37
Lignina	3	3
Pectina	2	2
Compuestos nitrogenados		
Proteína	7.5-10	7.5-10
Aminoácidos libres	ND	ND
Cafeína	1.1-1.3	2.4-2.5
Trigonelina	1.2-0.2	0.7-0.3
Ácido nicotínico	0.016-0.026	0.014-0.025
Lípidos		
Triglicéridos con insaponificables	17	11
Diterpenos (libres y esterificados)	0.9	0.2
Minerales	4.5	4.7
Ácidos y ésteres		
Ácidos clorogénicos	1.9-2.5	3.3-3.8
Ácidos alifáticos	1.6	1.6
Ácido quínico	0.8	1.0
Melanoidinas	25	25

ND: no detectado. Fuente: Tomada de (Farah, 2012).

La composición química de los granos verdes cambia drásticamente durante el proceso de tostado como consecuencia de las reacciones de pirólisis ya mencionadas, caramelización y Maillard.

Debido a esto, algunos compuestos se destruyen mientras otros se forman. El contenido de humedad de granos tostados (1.5 – 5 %) es mucho menor que en granos verdes, éstos valores son variables dependiendo del grado de tostado (Farah *et al.*, 2006).

Parte de la porción proteica se degrada, mientras que los aminoácidos libres y péptidos son consumidos en las reacciones de Strecker. Algunos de los aminoácidos reaccionan con azúcares reductores (mediante reacciones de Maillard) para formar compuestos de bajo peso molecular y melanoidinas que incorporan en sus estructuras compuestos tales como ácidos clorogénicos, galactomananos y arabinogalactan (Bekedam *et al.*, 2008).

1.5 Café: la bebida

El café es una de las bebidas más populares en el mundo, consumida por millones de personas en el mundo cada día. La importancia económica del café se debe precisamente a la bebida, que hace referencia a la infusión preparada a partir de los granos molidos y tostados de café (Esquivel & Jiménez, 2012).

La bebida puede ser preparada empleando varios métodos, los más comunes son: mediante simple filtración, café turco, preparación en cafetera eléctrica o máquina de espresso, cafetera italiana y prensa francesa (Farah, 2012). En general, estos métodos se basan en llevar a ebullición en agua los granos tostados molidos, verter, gotear o asperjar agua caliente sobre los granos tostados y molidos, para finalmente realizar el proceso de filtración. Independientemente del método de preparación empleado, la bebida y el café tostado comparten la mayoría de los compuestos, sin embargo se presentan ciertas diferencias organolépticas debido a los cambios en la concentración de sustancias durante la preparación de la bebida (Belitz *et al.*, 2009; Farah, 2012).

La preparación de la bebida involucra la extracción de compuestos solubles en agua, la mayor parte de la fracción lipofílica permanece en el filtro junto con el material sólido. Sin embargo esto se ve modificado ya que existen métodos de preparación como la elaboración de expreso que involucran un proceso de extracción con pequeñas cantidades de agua a altas presiones en un tiempo corto (Viani & Illy, 2005).

De la fracción lipídica total presente en la bebida, el 75 % (peso seco) corresponde a

triacilgliceroles mientras que únicamente el 1% son ácidos grasos libres (Trugo, 2003). Al igual que en granos tostados en la bebida se encuentran presentes galactomananos y arabinolactanos tipo II, estos polisacáridos representan entre 15-25 % en peso seco (Trugo, 2003).

La bebida contiene además, compuestos funcionales como flavonoides (catequinas y antocianinas), ácido cafeico y ferúlico. Se encuentran presentes otros compuestos biológicamente activos como ácido nicotínico, trigonelina, ácido quinolínico, ácido tánico, ácido pirogálico y cafeína. La bebida posee también actividad antioxidante atribuida principalmente a compuestos como cafeína, ácidos clorogénicos ácidos hidroxicinámicos y melanoidinas. Aproximadamente 80-100 % de los ácidos clorogénicos son extraídos mediante los métodos tradicionales de preparación de la bebida. Por su parte, las melanoidinas del café presentan mucho mayor actividad antioxidante que las obtenidas de otras fuentes, por ejemplo a cerveza (Minamisawa *et al.*, 2004; Morales & Jiménez-Pérez, 2004; Meletis, 2006; Vignoli *et al.*, 2011). La actividad antioxidante de la bebida se asocia a la presencia tanto de compuestos encontrados propiamente en el grano verde así como de compuestos que se forman durante el proceso de tostado (Esquivel & Jiménez, 2012). La figura 5 muestra algunos compuestos no volátiles que se encuentran presentes en la bebida de café.

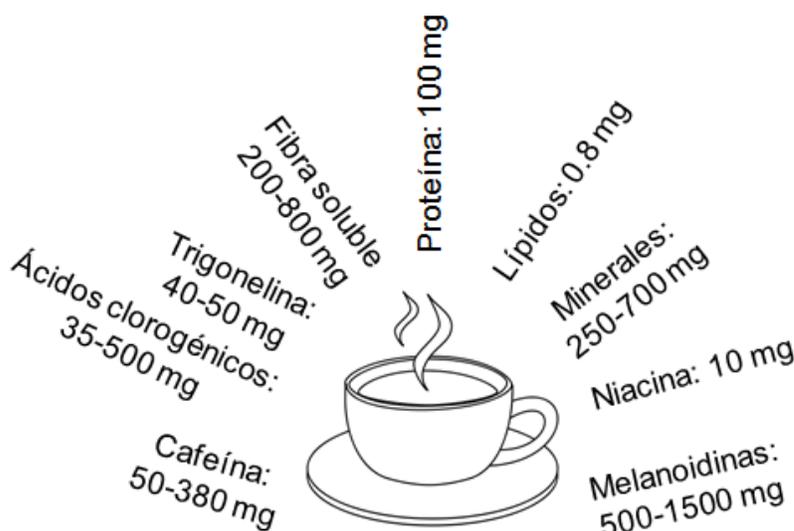


Figura 5. Composición química típica de 100 mL de bebida. La figura muestra las concentraciones de los principales compuestos no volátiles en bebidas preparadas a partir de granos con tostado medio. Modificada de (Farah, 2012).

1.6 Calidad en taza del café

El término “calidad” del café resulta un poco complejo, ya que su definición es completamente dependiente del nivel de la cadena producto-consumidor: a) a nivel del productor, la calidad resulta de la evaluación de características tales como el nivel de producción, precio y facilidades para el manejo del cultivo; b) a nivel exportador o importador, principalmente se refiere a características físicas y ausencia de defectos del grano, regularidad de abastecimiento y precio; c) nivel tostador: depende de las características o proporciones de las mezclas, estabilidad de las características, origen, precio, composición bioquímica y calidad organoléptica (calidad en taza); y finalmente, d) nivel consumidor: la determinación de calidad es en base a precio, sabor, efectos en la salud o en el estado de alerta, origen geográfico, así como aspectos ambientales y sociológicos (Leroy *et al.*, 2006; Bertrand *et al.*, 2012).

La calidad en taza del café es determinada en base a las características organolépticas de la bebida, las cuales incluyen aroma, sabor, cuerpo, acidez, calidad aromática y demás características que son percibidas principalmente por los sentidos del gusto y del olfato. La valoración de la calidad en taza por parte del consumidor es variable, por ejemplo, dependiendo de la nacionalidad la percepción de dichas características cambia. Se sabe que alemanes y suecos prefieren bebidas con mayor grado de acidez, contrario a las preferencias de los italianos (Leroy *et al.*, 2006).

La calidad en taza del café está definida por diversos factores: el genotipo de la planta, factores ambientales (origen geográfico, clima, altitud, temperatura, porcentaje de sombra del cultivo, capacidad de nutrientes del suelo, fertilización), prácticas agrícolas, procesamiento post-cosecha, almacenamiento del grano, así como la composición química básica del grano (Costa Freitas & Mosca, 1999; Bertrand *et al.*, 2006, 2012; Bosselmann *et al.*, 2009; Bhumiratana *et al.*, 2011; Poltronieri *et al.*, 2011).

En este contexto, un amplio rango de compuestos tanto volátiles como no volátiles presentes en los granos verdes de café contribuyen a la adquisición de los atributos sensoriales de la bebida. Por lo tanto uno de los factores que de manera directa influyen la calidad en taza del café es la composición química del grano verde (Tran *et al.*, 2016).

1.6.1 Calidad en taza y composición química

Un grupo considerable de compuestos tanto no volátiles como volátiles presentes en los granos de café son los determinantes de la calidad organoléptica de la bebida (Tabla 5). Dentro de los compuestos no volátiles presentes en granos tostados se encuentran carbohidratos y fibra (sacarosa, azúcares reductores, polisacáridos de pared celular, lignina), compuestos nitrogenados (proteínas, aminoácidos libres, cafeína, trigonelina), lípidos, minerales como potasio y fósforo, ácidos y ésteres (ácidos clorogénicos, ácidos alifáticos, ácido quínico) y melanoidinas (Sunarharum *et al.*, 2014; Tran *et al.*, 2016).

Tabla 5.- Compuestos relacionados con caracteres sensoriales de la bebida.

Compuesto / clase de compuesto	Impacto en características organolépticas
Cafeína	Intensidad del sabor, cuerpo y amargor
Trigonelina	Percepción global del aroma, calidad en taza, amargor.
Ácidos clorogénicos	Astringencia, amargor, acidez, calidad en taza, niveles altos de ACGs se relacionan con pobre calidad en taza y aparición de sabores indeseables
Ácidos: Acético, glicólico y láctico	Acidez
Polisacáridos	Retención de volátiles y del sabor, cuerpo (viscosidad)
Carbohidratos	Dulzor, notas a caramelo (mediante reacciones de Maillard)
Lípidos	Textura de la bebida, formación de crema en café espreso, acarrean volátiles y vitaminas liposolubles
Proteínas	Participan en formación de melanoidinas: responsables de la coloración del grano

(Spadone *et al.*, 1990; Urgert *et al.*, 1995; Clifford, 2000; Ginz *et al.*, 2000; Flament & Bessière-Thomas, 2002; Trugo, 2003; Buffo & Cardelli-Freire, 2004; Viani & Illy, 2005; Farah *et al.*, 2006; Koshiro *et al.*, 2006; Oliveira *et al.*, 2006; Redgwell & Fischer, 2006; Ashihara *et al.*, 2008; de Azevedo *et al.*, 2008; Sunarharum *et al.*, 2014).

Por otra parte, los compuestos volátiles contribuyen de manera importante a la formación del complejo aroma del café. Más de 1000 volátiles se han identificado en café, éstos se

pueden clasificar en diferentes clases (en orden decreciente de abundancia) que incluyen furanos, pirazinas, cetonas, pirroles, fenoles, hidrocarburos, ácidos y anhídridos, aldehídos, ésteres, alcoholes, compuestos azufrados, entre otros.

Sin embargo, se sugiere que un número reducido de estos volátiles (20-30) se encuentran relacionados de manera importante con características de sabor y aroma del café (Blank *et al.*, 1992; Semmelroch *et al.*, 1995; Czerny & Grosch, 2000; Ribeiro *et al.*, 2012; Toledo *et al.*, 2016).

En términos cuantitativos, las dos principales clases de volátiles que impactan las características organolépticas del café son los furanos y las pirazinas, y cualitativamente, los compuestos azufrados y las pirazinas son los más significativos. La amplia variación de la concentración de estos compuestos, es lo que conlleva a que el sabor y aroma del café sean caracteres extremadamente complejos (Risticovic *et al.*, 2008; Sunarharum *et al.*, 2014).

Los furanos son compuestos asociados con aromas dulces y salados, se forman durante el proceso de tostado a partir de la degradación de carbohidratos, ácido ascórbico y ácidos grasos insaturados (Sunarharum *et al.*, 2014).

Los compuestos que contienen azufre, a pesar encontrarse en muy bajas concentraciones poseen un potencial sensorial muy alto, por lo que son considerados dentro de los más importantes determinantes del aroma del café (Toledo *et al.*, 2016).

Ejemplos de este grupo de compuestos son el 2-furfuriltiol que se asocia fuertemente con notas tostadas del café (Blank *et al.*, 1992; Semmelroch *et al.*, 1995), 2-metil-3-furantiol y 3-metil-2-buten-1-tiol asociados a aroma carnoso (Czerny & Grosch, 2000).

Por su parte, se ha descrito la asociación entre pirazinas y aromas a nuez, terroso, notas de tostado y aromas verdes (Blank *et al.*, 1992; Holscher & Steinhart, 1995; Semmelroch & Grosch, 1996; Czerny & Grosch, 2000; Akiyama *et al.*, 2008). Dentro de este grupo de compuestos, los que tienen un impacto más significativo en el aroma del café son 3-isobutil-2-metoxipirazina, etilpirazinas, etenilpirazinas, 2-etil-3,5-dimetilpirazina y 2,3-dietil-5-metilpirazina, todos ellos con muy bajos límites sensoriales de detección (Mayer *et al.*, 2000; Mayer & Grosch, 2001; Akiyama *et al.*, 2003).

Las furanonas, son otro grupo de compuestos de alta abundancia y potencia aromática en el café, y contribuyen a al aroma dulce acaramelado del café tostado. Las principales furanonas son 4-hidroxi-2,5-dimetil-3(2H)-furanona, 2(5)-etil-4-hidroxi-5(2)-metil-3(2H)-

furanona, 3-hidroxi-4,5-dimetil-2(5H)-furanona y 4-etil-3hidroxi-5-metil-2(5H)-furanona. (Blank *et al.*, 1992; Akiyama *et al.*, 2003, 2008)

Durante el tostado también se generan y liberan algunos compuestos fenólicos que se asocian con aroma a especias como guayacol, 4-etilguayacol, 4-vinilguayacol y vainillina (Mayer *et al.*, 2000; Akiyama *et al.*, 2008; Ribeiro *et al.*, 2009).

Desde el punto de vista de la composición química del café, los compuestos volátiles y no volátiles influyen enormemente tanto la percepción como la aceptación del sabor y aroma del café por parte del consumidor. Por ello, es necesario contar con información confiable de mediciones de los caracteres sensoriales. Esto permitiría el desarrollo de investigaciones que conduzcan a encontrar una asociación verdadera entre dichos caracteres y la composición química de la bebida (Sunarharum *et al.*, 2014).

A pesar de contar con suficiente información, aún sigue siendo una limitante el desarrollo de modelos predictivos de características sensoriales de la bebida a través de perfiles químicos.

1.7 Mejoramiento genético tradicional de café

El mejoramiento genético de café es un factor de gran relevancia en el desarrollo tecnológico de la cafecultura, ya que representa un mecanismo de obtención de nuevas variedades con características incorporadas como: alta eficiencia productiva, resistencia o tolerancia a enfermedades, apariencia física del grano, cualidades organolépticas de la bebida, adaptabilidad a condiciones adversas de clima y suelo, entre otros caracteres (Zamarripa Colmenero & Escamilla Prado, 2002; López Morgado *et al.*, 2013).

Debido a que el café es un cultivo perenne y que para la obtención de la primera cosecha se requieren de 3.5 – 5 años, los programas de mejoramiento y selección de nuevas variedades requieren mínimo 25 años para producir resultados y por lo tanto requieren tanto de una buena inversión de fondos así como de recursos humanos.

1.7.1 Heredabilidad genética

Los primeros mejoradores de plantas no poseían ningún conocimiento acerca de la herencia de caracteres, por lo que la selección se basaba en principios tales como que los parentales más productivos producen progenies de plantas de igual o más alta

productividad. Sin embargo, es bien sabido que en los esquemas de mejoramiento moderno se obtiene mayor eficiencia teniendo conocimiento básico de la genética o la heredabilidad de los caracteres que son sujetos a selección (Brown & Caligari, 2008).

La heredabilidad de un carácter métrico o cuantitativo es una de sus propiedades más importantes. Expresa la proporción de la varianza total que es atribuible a la variación genotípica entre individuos, y que determina el grado de semejanza entre parientes. La principal función de la heredabilidad de un carácter cuantitativo es su rol predictivo, es decir expresa la confiabilidad de un valor fenotípico como una guía para la selección (Falconer & Mackay, 1996).

La heredabilidad en sentido amplio (H^2) se refiere al cociente de la varianza genotípica y la varianza fenotípica, mientras que en la heredabilidad en sentido estricto (h^2) es el cociente de la varianza genética aditiva y la varianza fenotípica. Es decir, la heredabilidad en sentido estricto explica qué proporción de los caracteres observables en los progenitores son transmitidos a la progenie.

Cuando el objetivo es mejorar los valores de un carácter en una progenie, conocer la heredabilidad en sentido estricto (h^2) permite predecir en qué medida aumentará la media poblacional en la próxima generación. Esto se determina dependiendo de qué tanto la media poblacional de la progenie difiere con respecto a la media de la población de la cual fueron seleccionados los parentales. La “respuesta a la selección” observada conduce a una estimación de la heredabilidad en sentido estricto conocida como “heredabilidad realizada” (Maldonado *et al.*, 2002; Visscher *et al.*, 2008).

Para la estimación de los componentes de varianza genética, destacan dos metodologías: el cálculo de la covarianza entre parientes y la regresión lineal media de la progenie sobre la media de los progenitores (medición directa de grado de semejanza entre la progenie y uno o ambos parentales) (Visscher *et al.*, 2008).

Los valores de heredabilidad de un carácter cuantitativo se encuentran en un rango de 0 a 1. Mientras el valor es más cercano a uno, existe mayor posibilidad de que mediante programa de mejoramiento potencialmente pueda alterarse la media de los valores de un carácter en futuras generaciones (Brown & Caligari, 2008).

1.7.2 Métodos y técnicas de mejoramiento genético

Las técnicas de mejoramiento son herramientas usadas por el mejorador para lograr la manipulación de la variación genética de las plantas que son objeto de este tipo de programas (López Morgado *et al.*, 2013).

Los programas de mejoramiento de café se han apoyado de la selección tanto de plantas provenientes de poblaciones existentes (variedades de agricultores y colecciones de germoplasma) así como de selección de cruzas creadas artificialmente entre variedades o especies. La elección del método de mejoramiento a ser empleado depende del comportamiento de la especie de café, de los objetivos del mejoramiento (resistencia a enfermedades, calidad, rendimiento, etc.) y del tipo de variedad a seleccionar, es decir si se trata de clones, líneas puras o variedades puras (Eskes & Leroy, 2008; Montagnon *et al.*, 2012).

Las técnicas de mejoramiento tradicional incluyen:

A) Selección genealógica: la selección genealógica es aplicada después de combinar características deseables de dos progenitores, permite la obtención de variedades suficientemente homogéneas para su propagación por semilla y multiplicación de genotipos para su distribución a los productores. En el caso de *C. arabica* este método ha resultado eficaz para elevar rendimientos y para obtener resistencia a enfermedades. La creación de las variedades Oro azteca en México (Fig. 6) e IAPAR en Brasil, son dos ejemplos claros de la aplicación de este tipo de selección (Leroy *et al.*, 2006; López Morgado *et al.*, 2013).

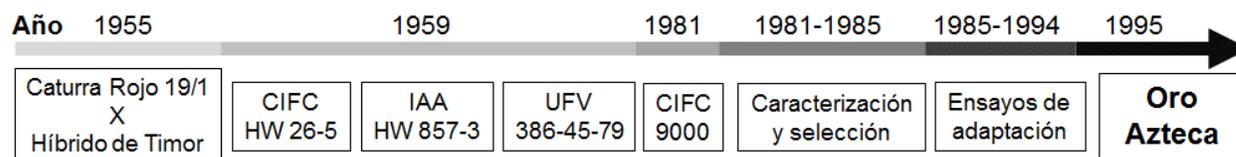


Figura 6. Diagrama esquemático de la generación de la variedad Oro Azteca. La variedad azteca fue generada en INIFAP a partir de la cruce entre la variedad Caturra Rojo y el Híbrido de Timor, después de varios ciclos de selección generacional y ensayos de adaptabilidad, se liberó en el año de 1995 como una variedad resistente a la roya anaranjada. Modificada de (López Morgado *et al.*, 2013).

B) Selección tradicional de clonas: la selección de clonas se basa inicialmente en la generación de un banco de germoplasma grande formado por plantas seleccionadas por agricultores locales o introducidas de otros países. Estas colecciones permiten una primera evaluación de las características de interés de una clona, tales como rendimiento, vigor, tamaño de grano, hábitos de crecimiento, resistencia a sequía, contenido de cafeína, resistencia a enfermedades, calidad en taza, etc. (Zamarripa Colmenero & Escamilla Prado, 2002).

Las plantas de interés son multiplicadas mediante esquejes (tallos enraizados) y se evalúan en ensayos de clonas junto con variedades estándar (semillas o clonas). Si más de un ciclo de selección es aplicado, la estrategia se conoce como selección recurrente (Eskes & Leroy, 2008).

C) Selección de líneas puras seguida de selección genealógica: dentro del proceso de mejoramiento se conoce como “líneas” a progenies avanzadas obtenidas de ciclos sucesivos de autopolinización, pero éstas no son necesariamente homocigotas, contrario a las “líneas puras”, las cuales sí son homocigotas.

La estrategia de selección de líneas puras a partir de una población de plantas fue aplicada en los años iniciales del mejoramiento del café, los híbridos resultantes fueron suficientemente homogéneos para permitir propagación por semilla y su posterior selección genealógica (Leroy *et al.*, 2006).

Mediante esta estrategia se seleccionaban plantas élite con determinadas características agronómicas, las cuales son auto fecundadas y multiplicadas para derivar variedades comerciales. Variedades como Caturra de Brasil, Kent y S.288 de India, SL.28 de Kenia y Java de Camerún son resultado de la aplicación de este método de selección (J. Singh, 2012).

D) Formación de variedades híbridas: los híbridos F1 son plantas obtenidas mediante la cruce entre variedades tradicionales de características conocidas y variedades de colecciones de germoplasma que pueden aportar características de resistencia a enfermedades, mejor rendimiento, etc. En los híbridos F1 se obtiene una ventaja: La heterosis o vigor híbrido, que se define como la capacidad de los híbridos de superar los valores de las propiedades deseables de sus progenitores (Eskes & Leroy, 2008).

En Camerún y Madagascar, en cruzamientos de líneas etíopes con variedades cultivadas, se logró elevar la productividad en 200 % con relación a la media de las variedades progenitoras. En un caso más reciente, en Costa Rica se ha permitido la distribución de los híbridos Centroamericano, Milenio y Casiopea, desarrollados durante 20 años. Éstos fueron seleccionados en base a sus buenas características de arquitectura de la planta, precocidad, rendimiento, calidad en taza y tolerancia e enfermedades (Noirot *et al.*, 2015).

E) Hibridación inter-específica: la finalidad es aprovechar el potencial de especies diploides como fuente importante de variabilidad. Como se sabe, la especie tetraploide *C. arabica* se reproduce principalmente por autopolinización mientras que *C. canephora* es diploide estrictamente alógama, debido a la presencia de alelos de autoincompatibilidad en su genoma. La generación de cruza inter-específicas entre estas dos especies es posible, aunque las plantas resultantes son menos fértiles (Philippe *et al.*, 2009).

La historia de mejoramiento de *C. arabica* está muy relacionada con la introgresión de genes de resistencia de *C. canephora* a través de hibridación inter-específica. El híbrido más importante hasta ahora de esta clase es el Híbrido de Timor (HdT), generado mediante hibridación espontánea entre *C. arabica* y *C. canephora* (Sobreira *et al.*, 2015). Dentro de las especies diploides, *C. canephora* provee la principal fuente de resistencia a roya del café y nemátodos de nodos de raíces (*Meloidogyne* spp) (Cabos *et al.*, 2010), sin embargo *C. liberica* así como *C. racemosa* han sido empleadas también como fuentes de resistencia a roya y resistencia a minador de la hoja de café causada por *Perileucoptera coffeella* (Filho *et al.*, 1999; Silvestrini *et al.*, 2007).

F) Hibridación intra-específica seguida por selección genealógica: esta estrategia ha sido aplicada principalmente para el mejoramiento de características como rendimiento, tolerancia a enfermedades y mejoramiento de calidad.

Este método consiste en la realización de cruza entre un genotipo ya conocido y un parental donador, la progenie híbrida es sujeta a selección genealógica. Algunos ejemplos exitosos de la aplicación de esta metodología son las variedades conocidas: Catuai, Tupi, líneas Catimor, Castillo, Costa Rica y Costa Rica 95 (Eskes & Leroy, 2008; J. Singh, 2012).

G) Selección por retrocruzamiento: esta estrategia se basa en la cruce de un híbrido con uno de sus parentales o con un individuo genéticamente similar a sus parentales, el objetivo es generar una progenie de identidad genética similar a la de los parentales. El retrocruzamiento se ha empleado para transferir características particulares a una variedad (Eskes & Leroy, 2008). Un ejemplo es la variedad Icatu, cuyo objetivo fue la formación de una variedad resistente a la roya anaranjada y a nemátodos. Icatu es el resultado de un cruzamiento entre *C. arabica* y *C. canephora*, cuyo genoma fue duplicado con colchicina para tener la condición tetraploide. El híbrido F1 obtenido se sometió de 4 a 5 ciclos de retrocruzamiento con las variedades Mundo Novo y Bourbón Catuai, creando el Icatu rojo y el icatu amarillo (López Morgado *et al.*, 2013).

En conclusión, debido a la naturaleza de *C. arabica* el método de selección comúnmente usado para esta especie es la selección generacional en poblaciones segregadas naturales o artificialmente creadas, la hibridación y selección por retrocruzamientos. Mientras que el mejoramiento de *C. canephora* involucra la selección de clonas o de híbridos entre clonas.

1.7.3 Propagación de plantas de café

Una vez seleccionada una variedad con buenas características agronómicas, éstas tienen que ser propagadas para su posterior plantación.

Durante muchos años, la propagación por semilla ha sido el método de elección, ya que resulta más fácil y rápido. Sin embargo, el café puede ser propagado vegetativamente mediante injertos, corte de tallo (estaca) o mediante propagación *in vitro* (Wintgens & Zamarripa Colmenero, 2008).

Las semillas de café Arábica producen plantas idénticas a la planta madre, excepto en el caso de híbridos, los cuales no se encuentran genéticamente estables. Por su parte, las semillas de café Robusta son resultado de cruces entre la planta madre de la cual se toman las semillas y una planta padre desconocida. Estas semillas, por lo tanto producen plantas genéticamente diferentes a la planta madre. La población generada por estas semillas es muy heterogénea (Vega *et al.*, 2007).

Es por ello que la selección del método de selección depende principalmente de dos factores: la especie de café a propagar y por otra parte, el objetivo planteado

(investigación, experimentación o fines comerciales) (Wintgens & Zamarripa Colmenero, 2008).

1.8 Fuentes genómicas del café

Numerosas investigaciones se han realizado en torno a la diversidad genética de *C. arabica*, enfocándose principalmente en tres tipos de material: A) diferentes cultivares / variedades; B) múltiples accesiones, líneas de introgresión e híbridos; C) líneas espontáneas o sub-espontáneas.

Los sistemas de marcadores empleados en estos trabajos incluyen ADN polimórfico amplificado al azar (RAPD), polimorfismos de longitud de fragmentos amplificados (AFLP, por sus siglas en inglés), repetidos de secuencia simple (SSR), repetidos de intersecuencia simple (SRAP) y polimorfismos de amplificación de regiones blanco (TRAP) (Steiger *et al.*, 2002; Masumbuko *et al.*, 2003; Maluf *et al.*, 2005; Tran, 2005; Vieira *et al.*, 2010; Mishra *et al.*, 2012; Al-Murish *et al.*, 2013; Motta *et al.*, 2014).

En general, los estudios demuestran que existe baja variación genética entre los cultivares de *C. arabica*. En el segundo grupo evaluado (accesiones/líneas de introgresión/híbridos) se demuestran bajos niveles de variación genética, siendo más diversas las líneas de introgresión. Por su parte, el grupo de líneas espontáneas/sub-espontáneas presenta diversidad desde baja hasta alta (Silvestrini *et al.*, 2007; Dessalegn *et al.*, 2008; Teressa *et al.*, 2010; Geleta *et al.*, 2012; Aerts *et al.*, 2013).

Al contrario de *C. arabica*, *C. canephora* posee alta diversidad genética debido a su origen, método de reproducción y disseminación, y se encuentra estructurada en 2 grupos principales (Guinean y Congolese) con 6 sub-grupos. El desequilibrio de ligamiento observado en poblaciones *C. canephora* seleccionadas para mejoramiento ha sido variable, lo cual significa que es posible la aplicación de estudios de asociación en búsqueda de mejorar genéticamente caracteres agronómicos y/o de calidad del café (Cubry *et al.*, 2012, 2013; Ferrão *et al.*, 2012; Leroy *et al.*, 2014).

1.8.1 Genoma del café

Se ha estimado que el genoma de café en el caso de *C. arabica* es de 1.3 Gb (Kochko *et al.*, 2010a), y 710 Mb en *C. canephora* (Denoëud *et al.*, 2014). La fuente genómica más importante hasta estas fechas es la secuenciación y reporte del genoma de *C. canephora*

en el año 2014. Los 11 cromosomas secuenciados de *C. canephora* presentan el orden de genes en los cromosomas más conservado entre las plantas angiospermas *Asteridae*. Los elementos transponibles (TEs) representan 49.2 % de la longitud del genoma de referencia. El análisis de la secuencia genómica muestra que existe transferencia de genoma organelos-núcleo con una típica cantidad de fragmentos derivados de cloroplasto (2014 inserciones, 0.16 % del genoma) y un tamaño inusual de 750 Kb de fragmentos provenientes de mitocondrias, este valor representa 100 Kb más de lo que se ha encontrado hasta ahora en genomas de otras plantas (Denoëud *et al.*, 2014).

En total el genoma de *C. canephora* posee 25,574 genes codificantes de proteínas. El genoma muestra la expansión de varias familias de genes específicos de especie como *N*-metiltransferasas involucradas en la producción de cafeína, genes relacionados con defensa, y de alcaloides y flavonoides involucrados en las síntesis de compuestos secundarios de las plantas de café (Denoëud *et al.*, 2014).

Adicionalmente al genoma de café, se han desarrollado otras fuentes genómicas como librerías de marcadores de secuencias expresadas (ESTs) (Fernandez *et al.*, 2004; Aggarwal *et al.*, 2007; Kochko *et al.*, 2010b; Privat *et al.*, 2011), cromosomas artificiales bacterianos (BACs) (Noir *et al.*, 2004; Leroy *et al.*, 2005; Cação *et al.*, 2013; Dereeper *et al.*, 2013) que podrían facilitar el desarrollo de estudios completos de transcriptómica así como la identificación de rutas biosintéticas relacionadas con la calidad del café. Éstas no han sido explotadas y las estrategias se han enfocado principalmente en: la asociación de marcadores moleculares con la calidad del café, identificación de genes reguladores de caracteres de calidad y análisis de loci asociados a caracteres cuantitativos (QTLs).

1.9 Mejoramiento tradicional de la calidad en taza del café

1.9.1 Mejoramiento tradicional de *C. arabica*

Los programas de mejoramiento genético de café, en los últimos años le han dado prioridad a la selección y mejoramiento de plantas de alta calidad de *C. canephora*, esto debido a su baja calidad comparada con *C. arabica*. Sin embargo, en el caso de *C. arabica* los esfuerzos se han dirigido a la transferencia de resistencia a enfermedades o adaptación al cambio climático (Davis *et al.* 2012, Leroy *et al.* 2006, 2011).

Entre los genotipos de *C. arabica*, tres grupos principales pueden identificarse: los genotipos silvestres de la región Sudán-Etiopía, las líneas cultivadas sin introgresión (Typica y Bourbon) y las variedades de introgresión, principalmente generadas a partir de genotipos de híbrido Timor (Leroy *et al.*, 2006).

Para la especie *C. arabica*, se ha sugerido la selección genealógica para la transferencia de genes de resistencia del híbrido Timor. A partir de este híbrido y mediante múltiples programas implementados en varios países se ha transferido resistencia a roya (*Hemileia vastatrix*), nemátodos en nodos de raíces (*Meloidogyne* sp.) y la enfermedad del grano de café (*Colletotrichum kahawae* sensu Hindorf) (Silva *et al.*, 2006; Cabos *et al.*, 2010).

Aunque la selección de *C. arabica* esté más enfocada en la adquisición de características de resistencia, estos procesos involucran la fijación de genes relacionados con otros caracteres, como la calidad en taza del café. En este contexto, los análisis organolépticos de líneas de *C. arabica* generadas por introgresión demuestran que éstas producen una bebida de buena calidad, similar a la de líneas sin introgresión empleadas como referencia (Puerta Q., 1998; Alvarado *et al.*, 2005; Gichimu *et al.*, 2014). Sin embargo existen otros estudios que demuestran que la calidad de la bebida se ve disminuida, tal es el caso de la línea T17924 derivada del híbrido Timor que presentó diferencias significativas comparada con las líneas sin introgresión control. Las diferencias se reflejaron tanto en los atributos de acidez y preferencia de la bebida, así como en el contenido de algunos compuestos tales como trigonelina, sacarosa y ácidos (Bertrand *et al.*, 2003).

Otra de las estrategias propuestas para generar diversidad genética en *C. arabica* es la generación de híbridos F1, mediante las cruzas entre variedades tradicionales de Etiopía con línea (Gimase *et al.*, 2014). En el centro de investigación francés CIRAD (Centro de cooperación internacional en investigación agronómica para el desarrollo), se han desarrollado variedades F1 mediante la cruce de variedades tradicionales que muestran vigor híbrido con más del 50 % de rendimiento comparado con su mejor parental (Bertrand *et al.*, 2005), la calidad en taza de varias de estas variedades resultó ser igual o superior a las variedades silvestres (Bertrand *et al.*, 2006).

Se ha propuesto que el uso de híbridos F1 podría ayudar a reducir variaciones cuantitativas de algunos caracteres bioquímicos así como la calidad en taza de la bebida (Leroy *et al.*, 2006).

1.9.2 Mejoramiento de calidad en *C. canephora*

Dentro de la especie *C. canephora*, se pueden diferenciar dos grandes grupos genéticos y geográficos, que incluyen Guinea (en África occidental) y Congolese de África Central. El mejoramiento de *C. canephora* se ha enfocado en características como tamaño de grano, cantidad de sólidos solubles, calidad en taza, así como en caracteres bioquímicos considerados determinantes de calidad en taza como contenido de azúcares, lípidos, cafeína, trigonelina y ácidos clorogénicos (Leroy *et al.*, 2006).

Tomando en cuenta lo anterior, el entendimiento de la herencia genética de caracteres asociados con calidad en taza podría optimizar los programas de mejoramiento, sin embargo en el caso del café pocos estudios se han enfocado en la identificación de compuestos bioquímicos o metabolitos relacionados con la calidad de la bebida, esto se ve reflejado en la Tabla 6 en la que se enlista el tipo de herencia así como los niveles de heredabilidad de los únicos 5 caracteres bioquímicos asociados a calidad en taza del café que han sido analizados (Mazzafera & Carvalho, 1991; Barre *et al.*, 1998; Montagnon *et al.*, 1998; Ky *et al.*, 1999; Leroy *et al.*, 2011; Akaffou *et al.*, 2012).

Montagnon y colaboradores (Montagnon *et al.*, 1998) demostraron en plantas de *C. canephora* que es posible mejorar caracteres relacionados con la calidad en taza del café sin afectar el rendimiento, debido a que estos caracteres se encuentran positivamente correlacionados.

Se han realizado también estudios tanto en híbridos intra-específicos como en híbridos inter-específicos con la finalidad de conocer el modo y nivel de herencia de sacarosa y trigonelina. El contraste entre estos dos grupos muestra que la heredabilidad medida a estos dos niveles (intra e inter-específico) es diferente, esto puede ser explicado por las diferencias entre especies en relación a la fijación de alelos específicos de genes que controlan la variabilidad de caracteres de calidad (Leroy *et al.*, 2011; Mérot-L'Anthoëne *et al.*, 2014a).

La alta heredabilidad en sentido estricto y la correlación positiva entre cafeína y contenido de lípidos sugiere que estos caracteres pueden ser mejorados juntos en híbridos intra-específicos, no es así para el caso de contenido de lípidos y sacarosa que muestran correlación negativa entre ellos, y deberían tomarse precauciones para seleccionar esta

combinación (Montagnon *et al.*, 1998). En el caso de cafeína, tanto genes mayores así como genes con efecto aditivo parecen estar involucrados en la variabilidad observada en el contenido de cafeína en semillas, lo que indica la complejidad de este carácter (Priolli *et al.*, 2008a).

Tabla 6. Tipo de herencia y heredabilidad de caracteres determinantes de calidad en taza.

Carácter	Híbridos intra-específicos	Híbridos inter-específicos
Contenido de sacarosa	Herencia poligénica NSH= 0.11	Herencia poligénica y aditiva Valores de heredabilidad más alta
Cafeína	Herencia aditiva NSH= 0.33; BSH= 0.76 NSH= 0.80 Efecto aditivo en hembras	Herencia poligénica y aditiva Un gen mayor con dos alelos No cafeína = gen recesivo (caf1) High BSH
Trigonelina	Efecto dominante y aditivo en machos y hembras NSH= 0.38	No aditivo, herencia materna Heredabilidad= 0.71
Contenido de lípidos	NSH= 0.74 Aditivo en hembras	- -
Ácidos clorogénicos	NSH= 0.36 Efectos aditivos en hembras	Herencia nuclear Isómero 3-FQA controlado por un gen mayor

Fuente:(Mazzafera & Carvalho, 1991; Barre *et al.*, 1998; Montagnon *et al.*, 1998; Ky *et al.*, 1999, 2001; Leroy *et al.*, 2011; Akaffou *et al.*, 2012; Mérot-L'Anthoène *et al.*, 2014b)

La herencia de la calidad en taza del café *per se* también ha sido objeto de estudio, Leroy y colaboradores reportaron que existe efecto aditivo materno para amargor en plantas de café robusta, en el caso de acidez observaron efecto aditivo paterno. Esta información podría ser de gran utilidad para la selección de parentales para el mejoramiento de la calidad en taza de café Robusta (Leroy *et al.*, 2011).

En conclusión, caracteres altamente heredables como lípidos o cafeína podrían ser mejorados con éxito mediante cruza entre parentales con valores favorables, sin embargo para caracteres con bajos niveles de heredabilidad como ácidos clorogénicos,

trigonelina o sacarosa, sería necesario implementar estrategias de selección asistidas por marcadores (Davis *et al.*, 2007).

1.10 Mejoramiento asistido por marcadores moleculares

1.10.1 Marcadores moleculares

La determinación de predictores moleculares de caracteres de calidad en taza del café podría ayudar a reducir los tiempos que toman los ciclos de selección en los programas de mejoramiento, así como los costos de evaluación fenotípica de las plantas de interés. Sin embargo, el uso de tecnologías basadas en DNA aún está en etapas iniciales y por lo tanto se considera con limitaciones (Tran *et al.*, 2016).

Pot y colaboradores, analizaron la diversidad de genes involucrados en el metabolismo de la sacarosa con el objetivo de identificar genes candidato que controlan la variabilidad de los niveles de este metabolito en *C. canephora* y *Coffea spp.* Para este fin emplearon polimorfismos tipo SNPs, INDELS y SSRs, lo que permitió la identificación de polimorfismos a nivel intraespecífico e interespecífico. Este grupo de variantes podría representar un enfoque para la identificación de marcadores moleculares útiles en el mejoramiento de café (Pot *et al.*, 2007).

Priolli y colaboradores, mediante el uso de variantes AFLPs y SSRs identificaron 48 marcadores asociados con caracteres de calidad en población F2 interespecífica (*C. canephora* X *C. arabica*). De éstos, 19 se asociaron con contenido de azúcares, 8 para cafeína, 8 para ácidos clorogénicos y uno compartido para cafeína y ácidos clorogénicos. Esta información podría ser usada para la selección asistida por marcadores en programas de mejoramiento de la calidad del café (Priolli *et al.*, 2008b).

De esta manera, se han identificado genes candidato que se piensa se encuentran regulando compuestos involucrados en la determinación de los caracteres de calidad del café (Tabla 7).

Tabla 7. Genes reguladores de compuestos determinantes de calidad en taza.

Especie	Sacarosa	Cafeína	Trigonelina	Ácidos grasos	CGAs
<i>C. arabica</i>	<i>CaSUS1</i>	<i>CmXRS1</i>	<i>CTgS1</i>	<i>OLE2</i>	<i>4CL8</i>
	<i>CASUS2</i>	<i>CTS2</i>	<i>CTgS2</i>	<i>DGAT</i>	<i>HQT</i>
	<i>CalnV3</i>	<i>CCS1</i>			<i>F5H1</i>
		<i>CaXMT1-2</i>			<i>POD</i>
		<i>CaMXMT1-2</i> <i>CaDXMT1-2</i>			
<i>C. canephora</i>	<i>CcSUS1</i>	<i>CmXRS1</i>		<i>CcOLE-5</i>	Genes de la vía fenilpropanoides
	<i>CcSUS1</i>	<i>CTS2</i>		<i>CcSTO1</i>	
	8 nuevos genes	<i>CCS1</i>			
		23 genes de alcaloides			

Referencias: (Paillard *et al.*, 1996; Ky *et al.*, 2000; Lashermes *et al.*, 2001, 2014; Coulibaly *et al.*, 2003; Noir *et al.*, 2004; Teixeira-Cabral *et al.*, 2004; López-Gartner *et al.*, 2009; Davis *et al.*, 2011; Gartner *et al.*, 2013; De Kochko *et al.*, 2015; Moncada *et al.*, 2015)

Resende y colaboradores reportaron recientemente una plataforma de 33,239 polimorfismos (SNPs) específicos de *C. arabica* y 87,271 específicos de *C. canephora*. Esta plataforma representa una nueva herramienta de utilidad en la aceleración del mejoramiento genético, el conocimiento de las bases genéticas de caracteres de interés y el manejo de la diversidad genética en especies de café (Resende *et al.*, 2016).

1.10.2 Análisis de identificación de QTLs y mapeo genético

La construcción de mapas genéticos es uno de los primeros pasos para el análisis de caracteres complejos tales como rendimiento, calidad y enfermedades. Esta estrategia permite la identificación de loci asociados a caracteres cuantitativos (QTLs) que se encuentran controlando dichos caracteres, consecuentemente es posible identificar genes candidato para el mejoramiento de caracteres como la calidad el café (Vega *et al.*, 2007). En algunos casos es posible que los marcadores que se encuentran asociados a QTLs sean usados directamente en selección asistida por marcadores (MAS), sin antes haber estudiado la funcionalidad del gen en cuestión (Tran *et al.*, 2016).

Paillard y colaboradores reportaron el primer mapa genético de *C. canephora* (Paillard *et al.*, 1996), a la fecha varios análisis de QTLs se han realizado en *C. canephora* y otras especies (Tabla 8), en búsqueda de regiones del genoma asociadas con la calidad del café.

Tabla 8.- Análisis de QTLs relacionados con calidad en taza del café.

Carácter	Número de QTLs y grupos de ligamiento	
	<i>C. canephora</i>	Otras especies
Sacarosa	9 QTLs/A,E,I,J	-
Cafeína	14 QTLs(A,C,E,I,K)	2 QTLs/A,G
Trogonelina	5 QTLs/F,G,I,K	1 QTL/G
Lípidos	7 QTLs/B,C,E,G,I	-
Ácidos clorogénicos	23 QTLs/A,B,D,E,F,I,J,K	1 QTL/A
Calidad en taza	14 QTLs(A,B,D,E,G,H,I)	-

Fuente:(Paillard *et al.*, 1996; Ky *et al.*, 1999; Pearl *et al.*, 2004; Landey *et al.*, 2013; Lashermes *et al.*, 2014)

En conjunto, la disponibilidad de las fuentes de información genómica del café resulta de gran relevancia para el desarrollo de métodos más eficientes de selección de plantas de café con características deseadas.

La complementación entre los nuevos avances en las tecnologías de secuenciación y el fenotipo de caracteres (características físicas, bioquímicas, niveles de metabolitos) podrían facilitar la identificación de marcadores asociados con calidad en taza para asistir la selección de plantas de café.

La aplicación de avances científicos y tecnológicos, principalmente en el área de la secuenciación de genomas y áreas emergentes como la metabolómica, podrían facilitar el entendimiento del genoma del café, por ende, la herencia de caracteres o rasgos cuantitativos que determinan la calidad en taza (Tran *et al.*, 2016).

1.11 Selección de plantas asistida por metabolómica: herramienta emergente

Actualmente, la selección asistida por marcadores (MAS, por sus siglas en inglés) se ha realizado mediante el monitoreo y la selección de caracteres empleando marcadores moleculares o SNPs. La aplicación de esta herramienta se considera que es exitosa cuando los caracteres son monogénicos o están regulados por genes dominantes (Herrmann & Schauer, 2013).

Sin embargo, en el caso de caracteres poligénicos, poliploidías, y caracteres influenciados fuertemente por efectos epistáticos o por el medio ambiente, como sabor, color, tolerancia a estrés, etc. este enfoque presenta limitaciones. Durante la aplicación de estas herramientas, aún sigue siendo preponderante la evaluación del fenotipo de la

planta, por lo que no es posible que el mejoramiento se base únicamente en la tecnología de MAS (Herrmann & Schauer, 2013).

Los fenotipos observados en una planta están estrechamente relacionados con los metabolitos. Es por ello que el empleo de metabolitos como marcadores predictivos de caracteres complejos podría ser de alto potencial.

El metaboloma se define como el conjunto completo de metabolitos presentes en un organismo en un momento determinado bajo condiciones específicas. La metabolómica es una herramienta que permite medir, cuantificar e identificar compuestos consumidos, sintetizados y excretados por un organismo. Una de las ventajas de la metabolómica es que esta tecnología es independiente de la disponibilidad de información genómica del organismo estudiado (Ausubel, 2010; Herrmann & Schauer, 2013).

En el perfil metabólico de una muestra biológica se ven reflejados varios factores como: efectos epistáticos, epigenéticos o post-traduccionales, los cuales finalmente determinan la ausencia o presencia de un carácter específico. Los avances en la metabolómica de alto rendimiento permiten actualmente el monitoreo de un gran número de metabolitos. La espectrometría de masas de alta resolución así como los patrones de fragmentación de los compuestos químicos permiten una identificación más precisa de los metabolitos (Hollywood *et al.*, 2006; Saxena & Cramer, 2013; Zabolina, 2013).

Recientes estudios sugieren que la metabolómica puede tener potencial para el descubrimiento de metabolitos marcadores predictivos de caracteres fenotípicos. Meyer y colaboradores estudiaron la relación entre la biomasa y los perfiles metabólicos de una población de líneas endogámicas recombinantes de *Arabidopsis* bajo condiciones altamente controladas. Encontraron una alta correlación (0.73) entre contenido de biomasa y una combinación específica de metabolitos, estos datos demuestran un alto valor predictivo de la composición química para contenido de biomasa (Meyer *et al.*, 2007).

En un estudio similar, empleando una planta modelo (varios ecotipos de *Arabidopsis*), se investigó la relación entre metabolitos, enzimas, productos finales como almidón, medidos al final de la noche y la biomasa de las plantas. Los resultados muestran que niveles bajos de almidón están ligados con alto contenido de biomasa. Esto podría explicarse por la acumulación de almidón durante el día y su degradación durante la noche para ser usado como combustible por el metabolismo de la planta, de esta manera

se estimula su crecimiento (Sulpice *et al.*, 2010). Este estudio representa otro ejemplo del potencial que podría tener la aplicación de la metabolómica en el mejoramiento de plantas.

En líneas de *Solanum lycopersicum*, Schauer y colaboradores realizaron un estudio de correlación de QTLs metabólicos y caracteres fenotípicos tales como rendimiento, índice de cosecha, número de semillas y contenido total de sólidos solubles. Sus hallazgos, basados en datos de tres cosechas independientes, mostraron correlación entre los sólidos solubles totales y los azúcares y ácidos orgánicos. También encontraron fuerte correlación entre el índice de cosecha y el contenido de un grupo de metabolitos (Schauer *et al.*, 2006). Este estudio se realizó con líneas de introgresión inter-específicas homocigotas. Sin embargo, en un estudio posterior de este mismo grupo de investigación (Schauer *et al.*, 2008) empleando líneas de tomate con fondo genético heterocigoto, encontraron resultados altamente promisorios demostrando que algunas de las fuertes limitaciones entre el perfil metabólico y caracteres morfológicos podrían ser superadas.

El primer estudio enfocado en la identificación de marcadores predictivos de caracteres agronómicos de importancia, fue el realizado por Steinfath y colaboradores. Cultivos de papa fueron estudiados para predecir a partir de su perfil metabólico las características de calidad de dorado y color de papas fritas (Steinfath *et al.*, 2010). La calidad de la papa a nivel industrial es determinada principalmente en base a dos parámetros: el grado de oscurecimiento que ocurre durante el freimiento y la aparición de puntos negros cuando fuerza física es aplicada sobre el tubérculo. Estas dos, son consideradas características indeseables, por lo que una herramienta de análisis que pudiera predecir potencialmente la aparición de puntos negros y el grado de oscurecimiento podría ser de gran beneficio para la industria. Por ello, en este estudio analizaron los perfiles metabólicos de varios cultivos comerciales de papa que mostraban un amplio rango de variación de estas dos características y establecieron un método que permitió la selección de un subgrupo de metabolitos que mostraran correlación con los caracteres evaluados. El modelo matemático fue validado en una segunda cosecha correspondiente a otra estación del año y de un área completamente diferente. A pesar de las grandes diferencias en las condiciones ambientales, los biomarcadores seleccionados predijeron exitosamente el fenotipo del material (Steinfath *et al.*, 2010).

La selección asistida por metabolitos marcadores representa una opción potencial para

el mejoramiento de plantas independientemente de la existencia de información genómica, sobre todo en casos en que los marcadores moleculares presentan ciertas limitaciones. En este sentido, es importante recalcar que un metabolito con alto valor predictivo de un carácter sensorial que adicionalmente es altamente heredable, representa una herramienta de selección que puede ser aún más robusta.

El uso de este tipo de enfoque metabolómico en combinación con el desarrollo de modelos de análisis matemáticos podrían ser implementados tanto en análisis de control de calidad, pero principalmente en programas de mejoramiento de plantas de café y la obtención de cultivos con características deseables.

2. JUSTIFICACIÓN

El café es una planta perenne, debido a esto los programas de selección y mejoramiento genético tradicional toman como mínimo 25 años para la obtención de variedades con caracteres agronómicos y de calidad deseables.

Importantes avances se han logrado en la selección de variedades asistida por marcadores moleculares. Sin embargo, esta herramienta presenta limitaciones cuando se trata de especies poliploides, caracteres poligénicos, así como para caracteres influenciados fuertemente por efectos epistáticos o medio ambientales.

Actualmente, la metabolómica representa una herramienta emergente altamente promisorio para el descubrimiento de firmas metabólicas o marcadores capaces de predecir fenotipos agronómicamente importantes en diferentes cultivos. Esto se debe a que los niveles de metabolitos están estrechamente relacionados con el fenotipo de la planta.

Por lo tanto, el empleo de metabolitos como marcadores predictivos de caracteres complejos, como en el caso de las características organolépticas de la bebida de café, podría ser de gran potencial.

Es por ello que en este trabajo se propone el desarrollo de una novedosa estrategia asistida por metabolómica, para la selección y mejoramiento de plantas de café con características sensoriales deseadas.

El uso de esta herramienta, de manera independiente o complementaria al mejoramiento genético tradicional, podría contribuir al acortamiento de los tiempos de selección de variedades de café con caracteres deseables.

3. HIPÓTESIS

A partir de los perfiles metabólicos de plantas de café en etapas tempranas de su crecimiento se pueden predecir características sensoriales de la bebida.

4. OBJETIVOS

4.1 Objetivo general

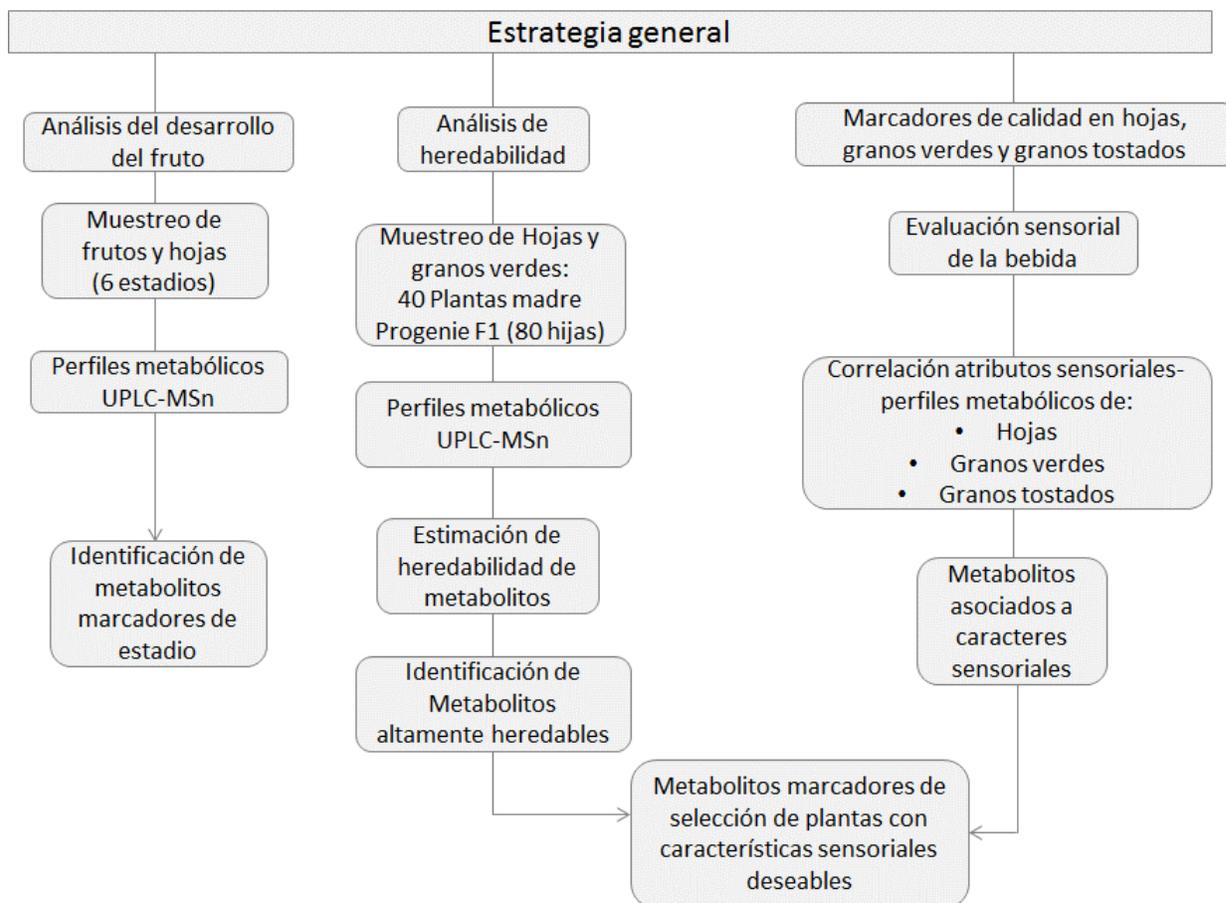
Identificar metabolitos altamente heredables predictivos de calidad en taza en *Coffea canephora* para la selección de plantas con características sensoriales deseables.

4.2 Objetivos específicos

- Identificar metabolitos marcadores de estadio de desarrollo de fruto a partir de perfiles metabólicos de frutos y de hojas café (*C. canephora*).
- Estimar la heredabilidad de metabolitos en familias de *C. canephora*, mediante un abordaje de medios hermanos maternos a partir de granos verdes y hojas.
- Realizar evaluación sensorial de la bebida y establecer correlación con los perfiles metabólicos de hojas, granos verdes y granos tostados.
- Identificar en hojas y granos verdes, metabolitos marcadores de calidad sensorial altamente heredables.

5. METODOLOGÍA

5.1 Estrategia general



5.2 Lugar de muestreo

Las plantas muestreadas de la especie *C. canephora* empleadas en este estudio, se colectaron en la finca experimental “Badenia”, propiedad de la compañía AMSA (Agroindustrias Unidas de México). La finca se encuentra ubicada en la Ciudad de Tapachula, Chiapas, México (latitud 15.02° N, longitud 101.36) a 465 metros sobre el nivel del mar. Esta zona presenta precipitación anual promedio de 4500mm, la temperatura anual registrada es de 26°C y las condiciones de sombra del cultivo son del 60 %.

5.3 Análisis de desarrollo del fruto

5.3.1 Muestreo

El diseño experimental para análisis de desarrollo del fruto consistió en el muestreo de 5 familias de plantas de la especie *C. canephora* de descendencia libre (medios hermanos maternos). Por cada familia se incluyeron 10 plantas.

Se colectaron hojas completamente expandidas y frutos de la posición media de ramas laterales durante 6 etapas del desarrollo del fruto, como se ilustra en la figura 7. Los estadios muestreados fueron designados de la siguiente manera: E1) frutos correspondientes a 10 semanas después de la floración (SDF) que corresponde a la etapa expansión o cabeza de alfiler; E2) fruto verde (20 SDF); E3) fruto amarillo (38 SDF); E4) pintón o amarillo-rojo (44 SDF); E5) rojo (48 SDF) y E6) cereza (50 SDF). Los frutos correspondientes a la etapa de cabeza de alfiler (E1) se seleccionaron de acuerdo a rasgos fisiológicos característicos. De la etapa E2 hasta la etapa E6, con la finalidad de obtener un muestreo homogéneo, se empleó una guía de colores pantone (COLOR BRIDGE TM, coated PANTONE ®) mediante la asignación de un código CMYK específico para cada etapa basándose en lo reportado por (Carvajal Herrera *et al.*, 2011) para frutos de *C. arabica*. La colecta se realizó en un horario establecido de 7 a 11 de la mañana. Para todas las etapas se colectó material vegetal suficiente para obtener por lo menos 500 mg de material en peso seco.

Las muestras de los 10 individuos muestreados por familia fueron mezcladas para obtener una única muestra representativa por familia. Los 5 grupos de muestras formados fueron designados como F1, F2, F3, F4 Y F5, haciendo referencia a las familias analizadas.

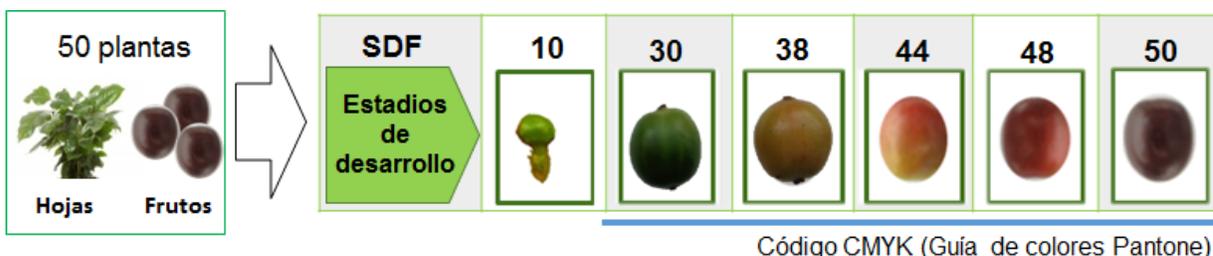


Figura 7. Cronograma de muestreo de etapas del desarrollo del fruto. SDF: semanas después de la floración.

Las muestras se envolvieron en papel absorbente para eliminar humedad dentro de bolsas Ziploc y se preservaron en una hielera con hielo seco y anticongelante durante la colecta y para su transporte.

5.3.2 Procesamiento de las muestras

Se realizó lavado con agua desionizada para eliminar cualquier tipo de partículas o polvo adheridos a la superficie de las muestras. Los frutos (a partir de la etapa de desarrollo E2) fueron disectados en semilla y pericarpio para ser analizadas de manera separada. Ambos tejidos (hojas y frutos) fueron congelados a $-80\text{ }^{\circ}\text{C}$ para su posterior liofilización. Para la extracción de metabolitos, las muestras fueron procesadas por familias (F1-F5) como anteriormente se mencionó.

5.3.3 Extracción de metabolitos

La extracción de metabolitos se realizó a partir de 35 mg de tejido liofilizado. El material se colocó en tubos de 1.5 mL y se le adicionó 450 μL de solvente de extracción (75 % methanol / 24.85 % agua / 0.15 % ácido fórmico). Las muestras se sometieron a sonicación en baño ultrasónico durante 25 min a temperatura ambiente, después de este tiempo, se centrifugaron a 10,000 rpm durante 5 min a 5°C y se recuperó el sobrenadante. La extracción se repitió 2 veces. Los sobrenadantes generados de ambas repeticiones se mezclaron y se filtraron empleando una membrana de nylon de 0.22 μm . Los extractos obtenidos fueron analizados mediante UPLC-MS.

5.3.4 Análisis metabólico

Los extractos tanto de hojas como de frutos fueron analizados empleando un sistema UPLC-ESI-MS/MS system (LCQ Fleet Ion Trap Mass Spectrometer, Thermo Finnigan, San Jose, CA, USA). Los metabolitos fueron separados en una columna Hypersil Gold C18 (dimensiones: 50 x 2.1 mm; tamaño de partícula: 1.9 μm). La temperatura de la columna se mantuvo a 40°C con un flujo constante de 550 $\mu\text{L}/\text{min}$. La fase móvil tanto para el análisis de frutos como para hojas consistió de H₂O con 0.1 % (v/v) ácido fórmico (solvente A) y el solvente B fue metanol con 0.1 % de ácido fórmico.

Los gradientes de solventes empleados para el análisis de los extractos mediante UPLC se especifican en la tabla 9 y tabla 10 para cada uno de los tejidos analizados (frutos y hojas).

Tabla 9.- Gradiente de solventes para frutos

Tiempo (min)	% Solvente A	% Solvente B
0	97	3
1	97	3
5	50	50
5.5	27	73
19.3	0	100
20.3	0	100
21.5	97	3
25	97	3

Tabla 10.- Gradiente de solventes para hojas

Tiempo (min)	% Solvente A	% Solvente B
0	92	8
0.5	92	8
6.5	35	65
20	0	100
21	0	100
21.5	92	8
25	92	8

5.3.5 Condiciones del espectrómetro de masas

Los espectros fueron adquiridos en un rango de m/z 50-1000, las corridas se realizaron en ambos modos de ionización (negativo y positivo) de manera separada. El tiempo de adquisición por scan fue de 500 ms (3 micro-scan).

Los parámetros empleados en la fuente de ionización por electrospray fueron los siguientes para el modo positivo: temperatura del capilar: 290°C; voltaje del capilar: 15 V; voltaje del spray: 4.5 kV; Tube lens: 80V; nitrógeno (sheat gas): 38 unidades arbitrarias (UA); y gas auxiliar: 10 UA.

5.4 Análisis de heredabilidad en hojas y frutos

5.4.1 Diseño experimental y muestreo

Se incluyeron 40 plantas madre de aproximadamente 15 años de edad, así como a su progenie F1 (2 hijas por cada planta madre), en total se analizó una población de 120 plantas. La población F1 fue generada a partir de semillas obtenidas por polinización abierta de las 40 plantas madre seleccionadas, es decir nuestra población de estudio estuvo conformada por familias de medios hermanos maternos (Fig. 8).

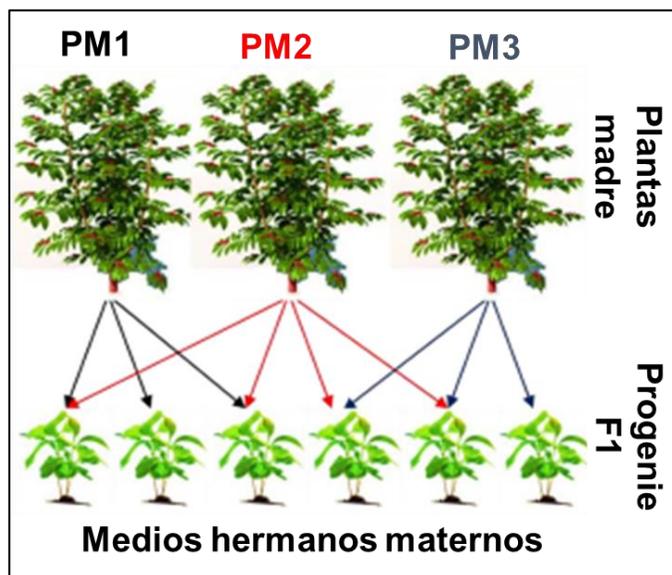


Figura 8. Diseño de medios hermanos maternos para análisis de heredabilidad. PM: planta madre.

El análisis de heredabilidad se llevó a cabo tanto en frutos como en hojas completamente expandidas. El muestreo comprendió 2 colectas: 1) se colectaron únicamente hojas en diciembre de 2012, tanto de plantas F1 (a los 6 meses de edad) así como de plantas madre; y 2) se obtuvieron frutos maduros (cerezas) durante la cosecha de diciembre-enero de 2014 tanto de las plantas madre como de su progenie F1, ya que en esta fecha tuvo lugar la primer cosecha de frutos de la progenie F1.

5.4.2 Procesamiento de las muestras y análisis metabólico

El acondicionamiento y procesamiento de las muestras colectadas (hojas y frutos) se llevó a cabo como a continuación se menciona: aproximadamente 2 gramos de cada muestra fueron congelados en nitrógeno líquido y liofilizados durante tres días. Las muestras liofilizadas se molieron en un molino Mixer Mill MM301 (Retsch®) a 30 Hz durante 15 segundos y se tamizaron hasta un tamaño de partícula < 300 µm.

La extracción de metabolitos así como las condiciones empleadas para el análisis metabólico fueron las ya mencionadas en los apartados 5.3.3 (Extracción de metabolitos), 5.3.4 (Análisis metabólico) y 5.3.5 (Condiciones del espectrómetro de masas).

5.4.3 Estimación de heredabilidad

Como anteriormente se mencionó, en este estudio se incluyeron 40 familias para realizar el análisis de heredabilidad. Para cada familia se contó con las mediciones de niveles de metabolitos de un único parental (planta madre), y en el caso de la progenie, por cada planta madre se obtuvo el promedio de las mediciones de 2 plantas hija.

El cálculo de los niveles de heredabilidad en sentido estricto se llevó a cabo mediante regresión lineal progenie-parentales, en este caso el cálculo se realiza a partir de las mediciones (niveles de metabolitos) de los parentales (la medición de uno sólo o el promedio de ambos padres) y el promedio de su progenie (Falconer & Mackay, 1996).

El coeficiente de regresión (Y_i) calculado a partir de la regresión entre parentales y progenie corresponderá al nivel de heredabilidad con su respectivo error estándar.

En las regresiones parentales-progenie se asume que la correlación ambiental entre padres y progenie no existe.

Dado que el cálculo se realiza a partir de la medición de un único parental, el resultado de la regresión estima $\frac{1}{2}$ de la heredabilidad en sentido estricto ($\frac{1}{2} h^2$).

El modelo estadístico para el cálculo de heredabilidad mediante regresión lineal es:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

Donde:

Y_i = medias de las progenes de las X familias

β_0 = intercepto

β_1 = coeficiente de regresión

X_i = Medición de un único parental a la media de i^{th} family

ϵ_i = Error estándar.

El coeficiente de regresión mide la proporción de covarianza parentales-progenie a la varianza de los padres.

5.5 Análisis de correlación de calidad sensorial del café y perfiles químicos

5.5.1 Muestras

De las 40 plantas muestreadas para análisis de heredabilidad se seleccionaron al azar 15 plantas madre de *C. canephora* y se obtuvieron muestras de aproximadamente 5 kg de frutos completamente maduros (cerezas) por cada planta durante la cosecha 2014/2015, una cantidad suficiente para la estandarización del tostado. Los frutos fueron sometidos a un tratamiento post-cosecha de secado (beneficio seco) para obtener granos verdes libres de cáscara y pergamino, proceso que duró aproximadamente 5 días.

Por otra parte, con la finalidad de tener variabilidad en los perfiles sensoriales de las muestras a evaluar, se incluyeron muestras tanto de la especie *C. canephora* como de *C. arabica* seleccionadas en base a su baja o alta calidad en taza previamente ya evaluada por los proveedores. En este caso, las muestras de grano verde fueron donadas por Agroindustrias Unidas de México (AMSA) y por el laboratorio de Tecnología del café del Instituto Tecnológico de Veracruz.

5.5.2 Evaluación sensorial de la bebida

5.5.4.1 Proceso de tostado del café

Las muestras de café verde que fueron sometidas a proceso de tostado consistieron de lotes que contenían granos libres de defectos físicos y que de acuerdo al análisis granulométrico presentaron un tamaño promedio de grano, esto para obtener un tostado homogéneo de las muestras.

Se realizó un tostado medio con escala de Agtron 60, en base a lo establecido por la SCAA. Se empleó un tostador de tambor rotatorio de capacidad de 600 gramos. Inicialmente el tostado inició con un precalentamiento del horno a 220 °C. La

estandarización del proceso de tostado se llevó a cabo mediante el monitoreo de la temperatura del café al inicio, durante y al término de cada tostado, así como el tiempo final tomado para cada muestra. Una vez estandarizado, se realizó el tostado por triplicado, estos tres tostados al final fueron mezclados para obtener una única muestra. Las muestras se introdujeron en bolsas de plástico (Ziploc®) y se almacenaron a -20 °C hasta su uso.

5.5.4.2 Preparación de la bebida

Los granos tostados de café se molieron empleando un molino eléctrico de discos con nueve niveles de graduación (modelo Kenia, Solocafé). El molido de las muestras fue grueso (nivel 4).

La preparación de la bebida se realizó mediante el método de prensa francesa. Se pesaron 50 gramos de café molido y se colocaron en jarras de vidrio (tipo francesa). En jarras con termostato tipo pistón de 1.5 L se calentó agua a 97°C y se vertió exactamente 1 L en las jarras conteniendo la muestra. Se agitó suavemente para homogeneizar y se dejó reposar durante 5 min. Una vez transcurrido este tiempo se bajó el pistón de la prensa francesa, para la obtención de la bebida que inmediatamente fue sometida al análisis sensorial.

5.5.4.3 Análisis sensorial descriptivo de la bebida de café

El objetivo de esta prueba fue identificar y cuantificar las notas sensoriales de cada una de las muestras. La evaluación se llevó a cabo de acuerdo a las especificaciones establecidas en la norma ISO 6564-1985. El análisis sensorial se llevó a cabo en el Laboratorio de Evaluación sensorial del Instituto Tecnológico de Veracruz.

Se empleó un panel conformado por 11 individuos, quienes fueron previamente entrenados. Los catadores incluidos aprobaron el sistema de preselección, selección y entrenamiento, en base a pruebas de identificación y ordenación de olores y sabores. Las notas sensoriales evaluadas fueron 10 en total: intensidad aromática, calidad aromática, cuerpo, acidez, amargor, astringencia, sabor, agrio, amargor y preferencia global. A cada juez se le presentaron tres muestras de café en tazas de porcelana codificadas con números aleatorizados de tres dígitos.

La calificación de los atributos sensoriales se realizó en una escala de 6 puntos (0 – 5), donde 0 significa ausencia de intensidad de la característica evaluada y 5 la mayor

intensidad percibida para cada uno de los atributos sensoriales evaluados enlistados en el formato de evaluación sensorial (Anexo 11.1 y 11.2).

5.5.3 Extracción de metabolitos

Las muestras de granos tostados fueron molidas en un molino Mixer Mill MM301 (Retsch®) a 30 Hz durante 15 segundos y se tamizaron hasta un tamaño de partícula < 300 µm.

Se pesaron 35 mg de muestra en tubos eppendorf y se les adicionó 450 µL de H₂O a 97 °C. Las muestras se sonicaron en baño ultrasónico durante 25 min a temperatura ambiente, transcurrido este tiempo se centrifugaron a 10,000 rpm durante 5 min a 5°C y se recuperó el sobrenadante. La extracción se repitió una vez más. Los sobrenadantes generados de ambas extracciones fueron mezclados en un sólo tubo y se filtraron empleando una membrana de nylon de 0.22 µm. Los extractos obtenidos fueron posteriormente sometidos al análisis metabolómico.

5.5.4 Análisis metabolómico

Los extractos de granos tostados se analizaron mediante UPLC-ESI-MS/MS. La temperatura de la columna se mantuvo a 40°C con un flujo constante de 550 µL/min. La fase móvil consistió de H₂O con 0.1 % (v/v) ácido fórmico (solvente A) y el solvente B fue metanol con 0.1 % de ácido fórmico.

El gradiente empleado fue: 2 % B, 0 - 1 min; 2 – 66 % B, 1 - 6 min; 66 – 88 % B, 6 - 12 min; 88 - 93 % B, 12 - 19 min; 93 – 97 % B, 19 – 20 min; 97 - 100 % B, 20 - 25 min; 100 % B, 25 – 27 min; 100 % B – 98 % A, 27 – 27.5 min y finalmente, reequilibrio de la columna durante 3.5 min (27.5 - 31 min) con 98 % A.

Los espectros fueron adquiridos en un rango de m/z 50-1000, se realizaron corridas de manera independiente para cada uno de los modos de ionización (positivo y negativo). El tiempo de adquisición por scan fue de 500 ms (3 micro-scan). Los parámetros empleados en la Fuente de ionización por electrospray fueron los siguientes: temperatura del capilar: 330°C; voltage del capilar: 15 V; voltage del spray: 4.5 kV; Tube lens: 80V; nitrógeno (sheat gas): 38 unidades arbitrarias (UA); y gas auxiliar: 10 UA.

6. RESULTADOS

6.1 Desarrollo del fruto

En este trabajo se planteó analizar los perfiles metabólicos de frutos y hojas durante el desarrollo del fruto de *C. canephora* con el objetivo principal de identificar metabolitos biomarcadores que permitan discriminar el nivel de desarrollo o madurez de los frutos.

Con este fin se llevó a cabo un análisis de metabolómica no dirigida mediante UPLC-MS para determinar los niveles de metabolitos a partir de extractos de frutos colectados durante 6 estadios de desarrollo: E1: frutos completos colectados 10 semanas después de la floración (SDF); E2: frutos verdes 30 SDF; E3) frutos amarillos de 38 SDF; E4) frutos pintón de 44 SDF; E5) frutos maduros (rojos) de 48 SDF; y E6) correspondió a frutos cereza de 50 SDF (Fig. 9). A partir de la etapa E2 los frutos se separaron en endospermo y pericarpio, y cada tejido se fue analizado de manera independiente. Como anteriormente se mencionó en la metodología (apartado 5.3.1) se analizaron 5 familias de plantas las cuales fueron designadas como F1-F5.

En cada uno de los 6 estadios (E1 - E6) también se colectaron hojas, esto con el objetivo de dilucidar si los cambios metabólicos que se observan en el fruto durante su desarrollo también pueden verse reflejados en los perfiles metabólicos de otros órganos o tejidos de la planta.



Figura 9. Estadios muestreados para análisis de desarrollo del fruto. SDF: semanas después de la floración.

6.1.1 Evaluación de perfiles metabólicos durante el desarrollo del fruto

Los datos obtenidos de los análisis mediante UPLC-MS fueron procesados y filtrados con el software MZmine 2.21 (Pluskal *et al.*, 2010). En los frutos, se detectaron en total 665 picos válidos con su correspondiente valor de m/z (relación masa-carga), cada uno fue considerado como un metabolito independiente. Mientras que a partir de los perfiles metabólicos de hojas se detectaron 613 metabolitos.

Empleando como variables los niveles de abundancia de estos metabolitos en cada una de las muestras, se realizó un análisis de componentes principales (PCA) para visualizar las diferencias en los perfiles metabólicos de los diferentes estadios de desarrollo del fruto.

La figura 10 A muestra el PCA de los perfiles metabólicos obtenidos a partir de frutos completos. Los dos primeros componentes principales explican el 34 % de la varianza de los datos. Específicamente se puede observar la separación de las muestras en grupos aislados dependientes del estadio de desarrollo, esto se observó para los estadios E1, E2, E5 y E6. Los estadios E3 y E4 (frutos amarillos y amarillo-rojo, respectivamente) se agruparon de manera conjunta, pero fueron claramente discriminadas del resto de las muestras (Fig. 10 A). En los gráficos de PCA de tejido de endospermo y pericarpio (Fig. 10B-C) no fueron incluidas las muestras correspondientes al estadio E1, debido a que en esta fase el endospermo aún no está desarrollado.

El análisis de los perfiles metabólicos de tejido de endospermo (Fig. 10B) ilustra la clara discriminación de un grupo que engloba a muestras de frutos de 30 SDF (estadio E2). El patrón de agrupamiento de los estadios E3-E6 muestra la formación de dos grupos: el primero contiene a muestras de los estadios E3 y E4, y en el segundo se diferenciaron las muestras de los estadios E5 y E6. El análisis de PCA de tejido de pericarpio (Fig. 10C) permitió el correcto agrupamiento de las muestras en 5 grupos correspondientes a los 5 estadios de desarrollo incluidos (E2-E6). El componente principal 1 explicó el 15.4 % de la varianza de los datos mientras que el 10.3 % fue explicado por el componente 2.

Los componentes 1 y 2 del PCA explicaron un porcentaje bajo (20.2 %) de la varianza de los datos, en el caso los perfiles metabólicos obtenidos de hojas (Fig. 10 D). Se puede observar que las muestras correspondientes a los estadios E1 y E2 se agrupan conjuntamente. En el caso de los estadios E3-E6 no fue posible visualizar grupos aislados en base al nivel de desarrollo.

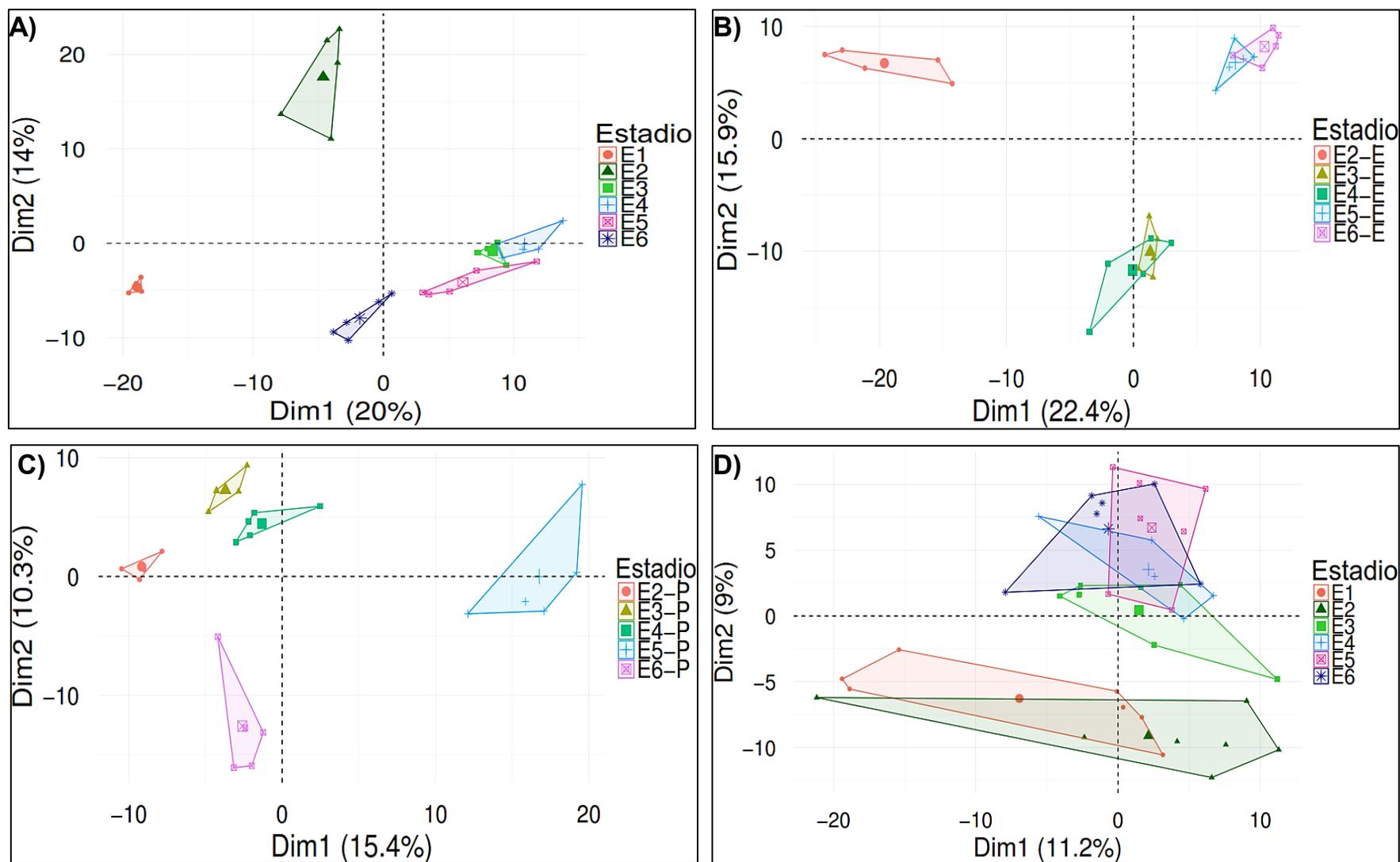


Figura 10. PCA de perfiles metabólicos durante el desarrollo de frutos. A) Frutos completos; B) frutos (endospermo); C) frutos (pericarpio); y D) hojas.

6.1.2 Selección de metabolitos predictivos de estadio de desarrollo

La identificación de un metabolito o de un grupo más restringido de éstos que permiten diferenciar entre estadios de desarrollo es algo que resulta ideal. Para este fin, empleando los niveles de metabolitos determinados en frutos completos y en hojas, se generó un modelo predictivo de bosque aleatorio (Random Forest Model) para la identificación de marcadores de desarrollo del fruto o de nivel de madurez. El algoritmo empleado consistió de tres pasos: entrenamiento, validación, y prueba. El algoritmo se basó en la construcción de 500 árboles de decisión a partir de los cuales se genera el mejor modelo predictivo que permite clasificar las muestras en base a su estadio de desarrollo (categoría) correspondiente. En todos los análisis, del total de muestras el 70 % se emplean para entrenar el modelo, 15 % para la fase de validación y 15 % restante para probar el poder predictivo del modelo generado. La importancia de la variable (metabolito) es determinada por el valor del parámetro Mean Decrease Gini, los metabolitos que presentan los valores más altos de este parámetro representan la contribución más importante en la predicción de la correcta clasificación de las muestras. De esta manera se seleccionaron 20 metabolitos con mayor potencial predictivo que permiten diferenciar entre estadios de desarrollo tanto a partir de perfiles metabólicos de frutos así como perfiles metabólicos de hojas.

Tabla 11.- Matriz de error de modelo predictivo RandomForest generado con perfiles metabólicos de frutos.

Real	Predicho					
	E1	E2	E3	E4	E5	E6
E1	5	0	0	0	0	0
E2	0	5	0	0	0	0
E3	0	0	5	0	0	0
E4	0	0	0	4	1	0
E5	0	0	0	0	5	0
E6	0	0	0	0	1	4

Error promedio de clasificación: 7 %. Los números resaltados en negritas representan los casos en los que las muestras fueron correctamente asignadas por el modelo.

El modelo predictivo generado mediante perfiles metabólicos de frutos permitió clasificar las muestras de acuerdo a su estadio de desarrollo con una tasa de error de 7 % (Tabla 11), lo que representa un modelo sumamente robusto. En el caso de las hojas el modelo predictivo resultó con un error promedio de clasificación de 27 % (Tabla 12).

Tabla 12.- Matriz de error de modelo predictivo de bosques aleatorios generado con perfiles metabólicos de hojas.

Real	Predicho					
	E1	E2	E3	E4	E5	E6
E1	3	2	0	0	0	0
E2	0	5	0	0	0	0
E3	0	0	2	3	0	0
E4	0	0	0	5	1	0
E5	0	0	0	2	3	0
E6	0	0	0	1	0	4

Error promedio de clasificación: 27 %. Los números resaltados en negritas representan los casos en los que las muestras fueron correctamente asignadas por el modelo.

Para determinar la identidad de los 20 iones seleccionados, se llevó a cabo un análisis robusto en el que se conjuntó tres tipos de datos para tener una asignación certera. 1) Información de la relación masa/carga (m/z) obtenida mediante la trampa de iones con la cual contamos en nuestro laboratorio así como información de masa exacta obtenida mediante un equipo de alta resolución de masa (SYNAPT, Langebio). Mediante la lista de masas generada a partir de estos dos equipos se llevó a cabo la búsqueda en SpiderMass (Winkler, 2015) empleando una base de datos que contiene únicamente metabolitos reportados en café y la cual fue generada específicamente para el desarrollo de este trabajo. 2) Por otra parte además de los patrones de fragmentación (MS^3) obtenidos en nuestra trampa de iones, se realizaron corridas adicionales en el equipo SYNAPT con la finalidad de obtener datos de los patrones de fragmentación de cada uno de los metabolitos. Extrayendo esta información fue posible comparar dichos patrones contra la base de datos de Massbank. Esta estrategia permite tener mayor seguridad en la asignación de la identidad de los iones detectados.

En las figuras 11 y 12 se muestran las listas de 20 metabolitos identificados en frutos (Fig. 11 A) y en hojas, así como el ranking de importancia (basado en el Mean Decrease Gini) que indica la contribución que tiene cada uno de los metabolitos para lograr la correcta clasificación de las muestras. Empleando únicamente este grupo de 20 metabolitos, el análisis de PCA presentó un mayor poder discriminante para diferenciar entre las diferentes etapas de desarrollo del fruto (Fig. 11 B y 12 B), comparado con los resultados obtenidos anteriormente (Fig. 10 A y D) en donde se empleó el perfil metabólico completo de las muestras. Este hecho demuestra la capacidad de predicción de este grupo de metabolitos que pueden ser empleados como marcadores específicos de estadio.

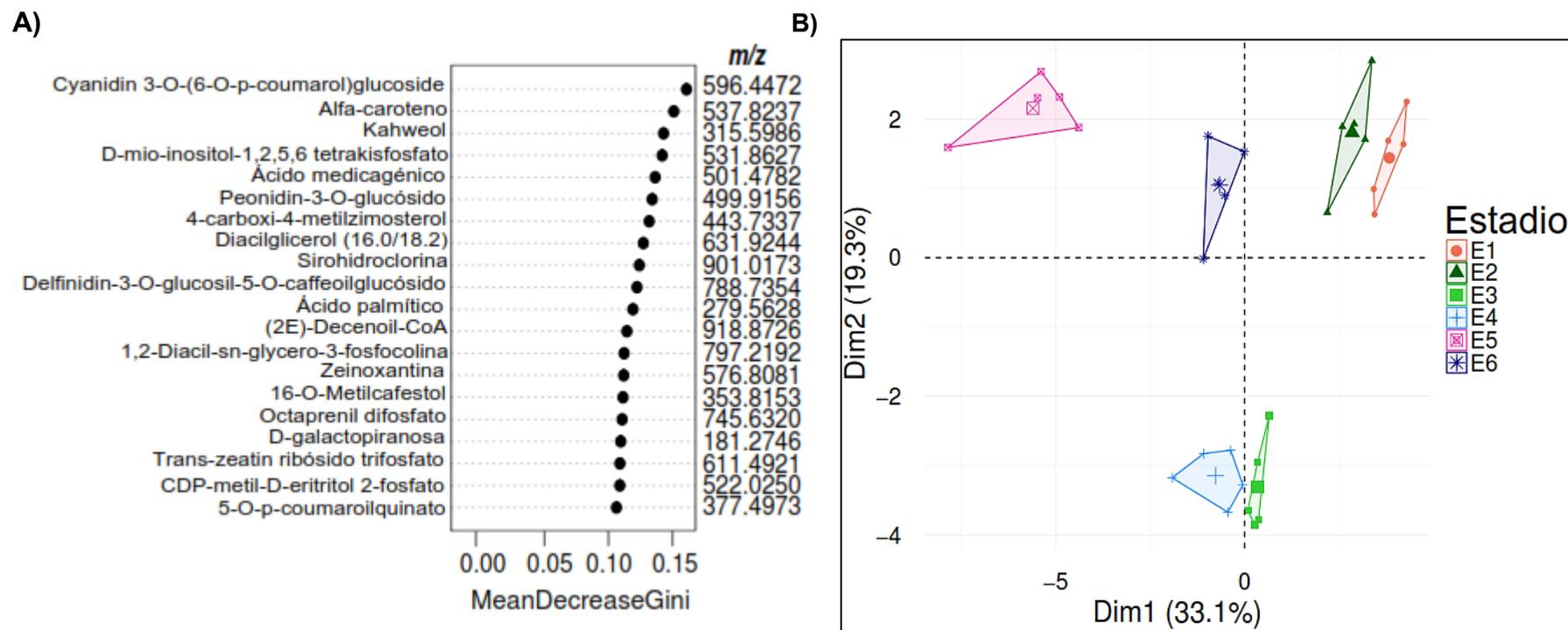


Figura 11. Ranking de importancia de metabolitos detectados en frutos que presentan mayor diferencia durante el desarrollo del fruto. A) Se muestran los 20 metabolitos identificados en frutos mediante el modelo predictivo de bosques aleatorios, empleando el programa Rattle. Los valores de Mean Decrease Gini representan el nivel de importancia de la variables (metabolitos), mientras más alto es este valor mayor es la contribución de la variable. La figura B corresponde al análisis de PCA generado con los 20 metabolitos seleccionados.

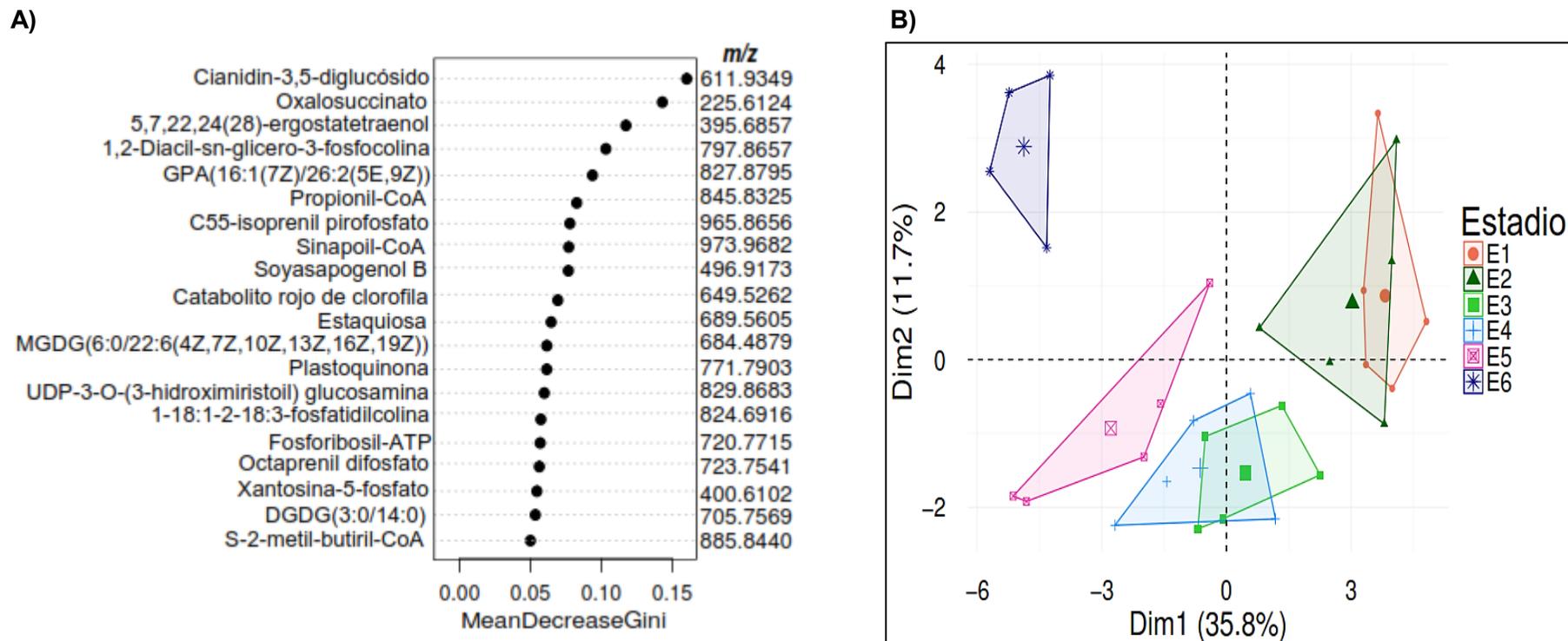


Figura 12. Ranking de importancia de metabolitos detectados en hojas que presentan mayor diferencia durante los estadios de desarrollo del fruto. A) Se muestran los 30 metabolitos identificados en hojas mediante el modelo predictivo de Random Forest empleando el programa Rattle. Los valores de Mean Decrease Gini representan el nivel de importancia de las variables (metabolitos), mientras más alto es este valor mayor es la contribución de la variable. La figura B corresponde al análisis de PCA generado con los 20 metabolitos seleccionados.

Para entender la relevancia biológica de los metabolitos identificados como marcadores, en el proceso fisiológico del desarrollo del fruto de *C. canephora*, se realizó la búsqueda en KEGG para clasificarlos de acuerdo a la vía metabólica en la que participan. De esta manera, los 20 metabolitos identificados en frutos se agruparon en 7 clases diferentes (Fig. 13 A), la clase V correspondiente a metabolismo y biosíntesis de terpenoides y carotenoides, fue la más enriquecida ya que estuvo conformada por 9 de los 20 metabolitos. Por otra parte, en hojas la clase III de metabolismo de ácidos grasos, esfingolípidos y glicerolípidos fue a la que mayor número de metabolitos diferenciales pertenecieron. Adicionalmente, en hojas se identificaron compuestos involucrados en el ciclo de ácidos tricarboxílicos así como en biosíntesis y degradación de aminoácidos (Fig. 13 B).

Se realizó un análisis comparativo del patrón de los niveles de los metabolitos seleccionados en frutos y hojas. En frutos completos (Fig. 14) los niveles de la mayoría de los metabolitos mostraron similitud en los estadios E1 y E6 (10 y 50 SDF, respectivamente) durante la fase de maduración del fruto (30-48 SDF) el perfil metabólico presentó mayor variabilidad.

El patrón de los niveles de metabolitos en hojas (Fig.15), de manera global permite distinguir la transición entre periodos tempranos del desarrollo del fruto (10–30 semanas después de la floración) y el inicio de la maduración del fruto (fruto amarillo de 38 SDF).

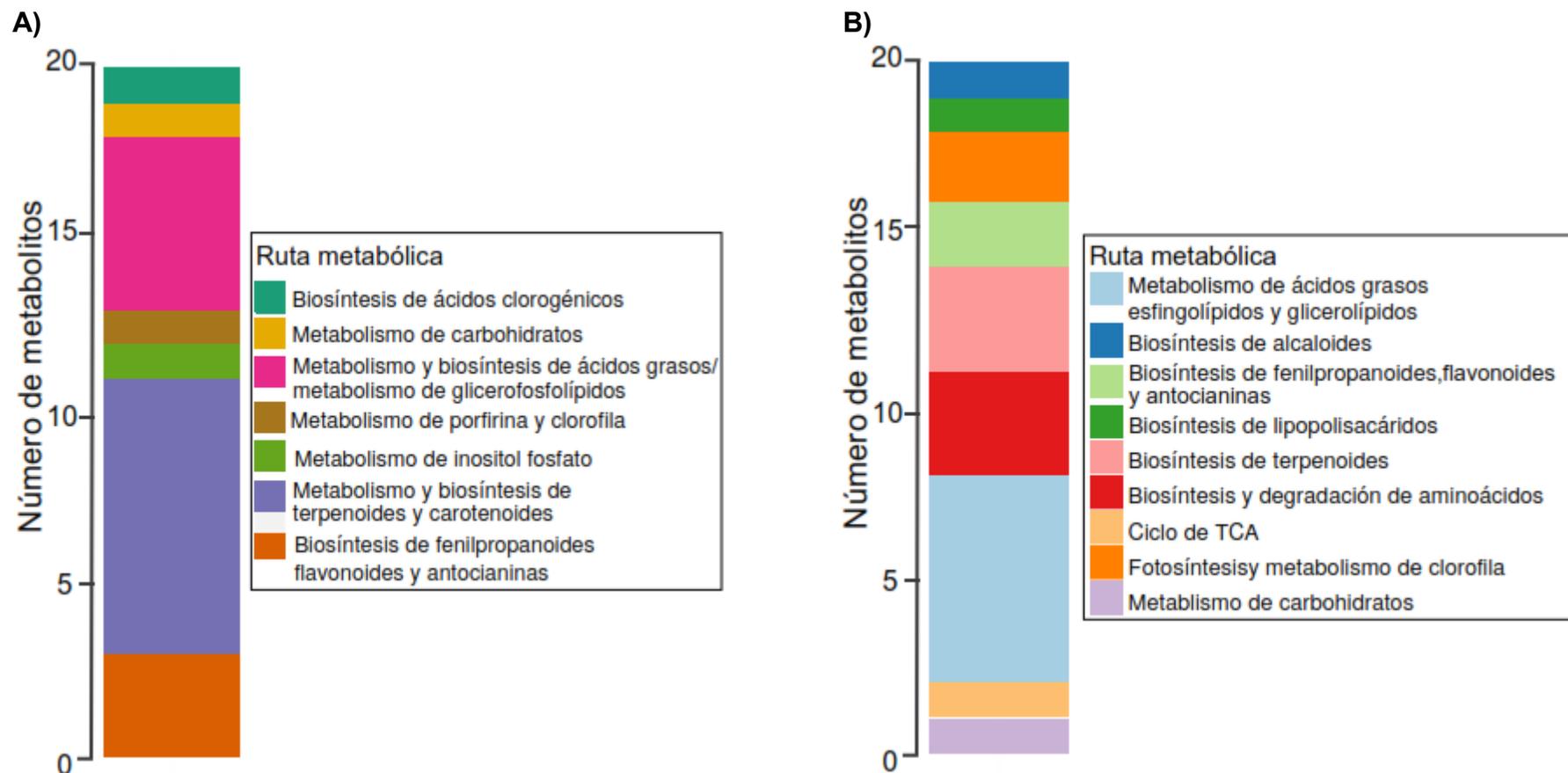


Figura 13. Clasificación de metabolitos seleccionados basada en KEGG. En A) se muestra la clasificación de los metabolitos seleccionados en frutos y B) en hojas.

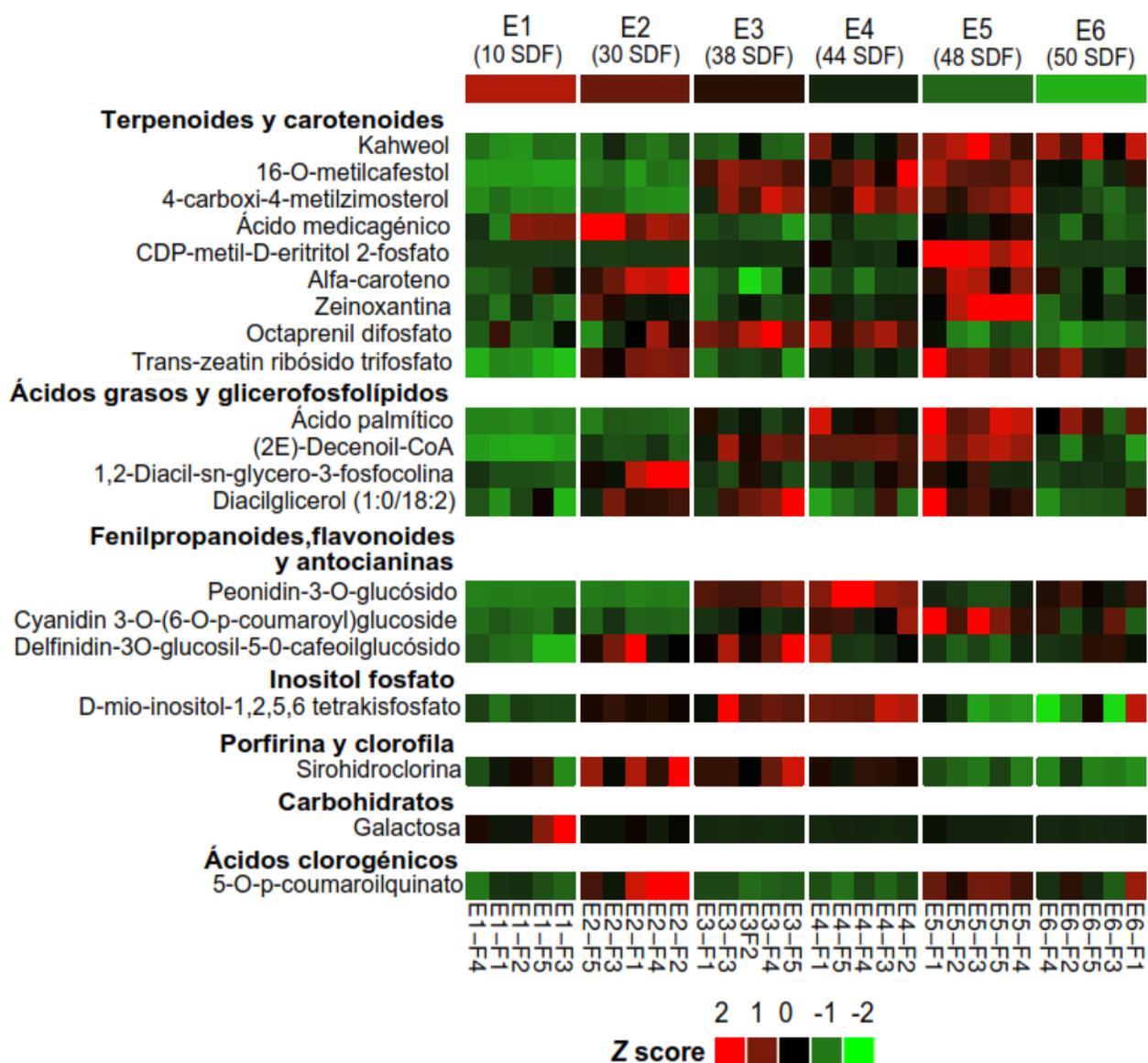


Figura 14. Heatmap de los niveles de metabolitos detectados en frutos durante el desarrollo del fruto. Los valores de intensidad de cada uno de los metabolitos fueron estandarizados a valor Z, las observaciones con valores positivos o negativos de Z representan desviaciones estándar por encima o por debajo de la media. SDF: semanas después de las floración

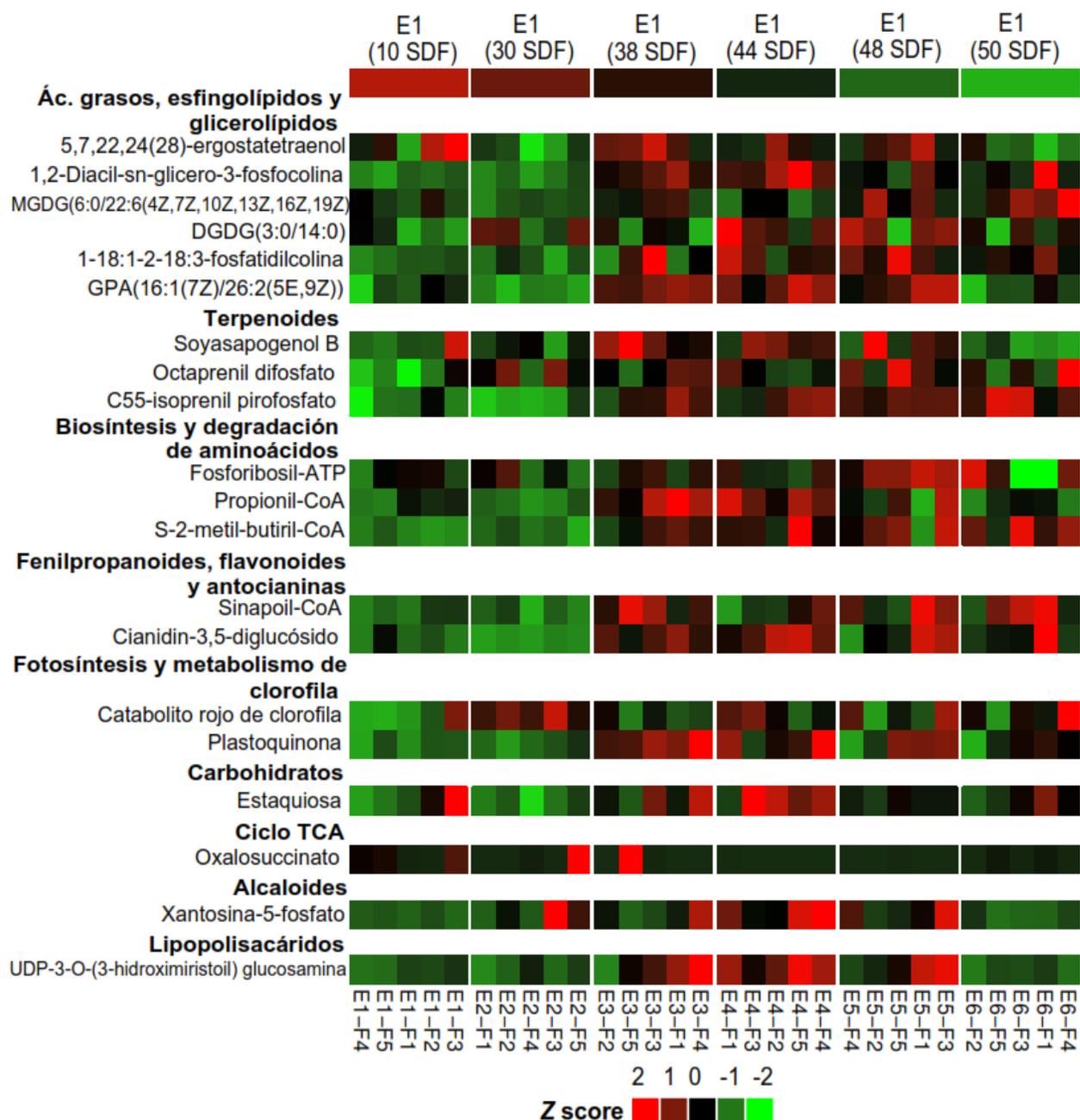


Figura 15. Heatmap del patrón de niveles de metabolitos detectados en hojas durante el desarrollo del fruto. Los valores de intensidad de cada uno de los metabolitos fueron estandarizados a valor Z, las observaciones con valores positivos o negativos de Z representan desviaciones estándar por encima o por debajo de la media. SDF: semanas después de las floración

En frutos se analizaron los tejidos de endospermo y pericarpio por separado, a continuación se describe la evolución de los niveles de los metabolitos en estos tejidos enfocándose en cada una de las clases identificadas en KEGG.

Clase I: Biosíntesis de ácidos clorogénicos:

Se identificó un metabolito de m/z 377.4973 perteneciente a la ruta de biosíntesis de ácidos clorogénicos, cuya identidad fue asignada como 5-O-p-coumaroilquinato. Este metabolito fue detectado desde etapas tempranas del desarrollo del fruto (Fig. 16). Sin embargo a las 30 SDF (estadio E2) se presentó un aumento drástico en sus niveles, principalmente en tejido de endospermo (Fig. 16 B). El 5-O-p-coumaroilquinato presentó disminución significativa en los siguientes estadios, aunque las abundancias observadas en los estadios E5 y E6 superaron a los mostrados en estadios tempranos.

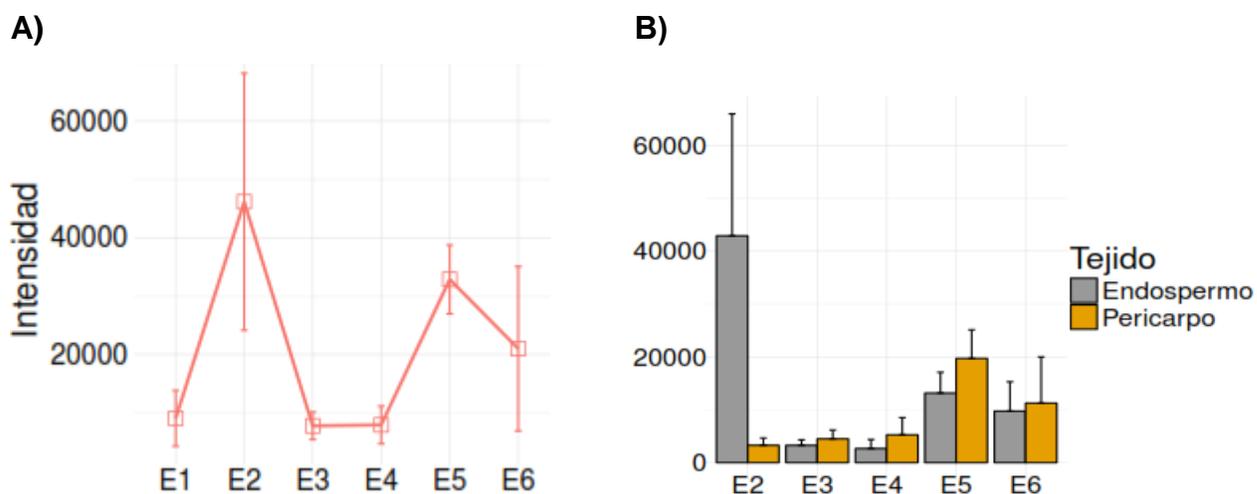


Figura 16. Cambios de los niveles de 5-O-p-coumaroilquinato durante el desarrollo del fruto de café. En las figuras A) y C) se muestran los niveles de los compuestos en frutos completos. Los valores graficados corresponden a la media de las 5 familias de plantas analizadas (F1-F5) \pm desviación estándar.

Clase II: Metabolismo de carbohidratos/metabolismo de galactosa

En frutos completos de frutos en expansión de 10SDF, se observó niveles altos de galactosa (m/z 181.2746) (Fig. 17 A). En la transición de fruto verde a amarillo (E2-E3)

se observó disminución de este compuesto, en estadios posteriores estos niveles se mantuvieron hasta los estadios fruto maduro, mostrando mayor abundancia en pericarpio a partir del estadio E4 que corresponde a frutos amarillo-rojo (Fig. 17 B).

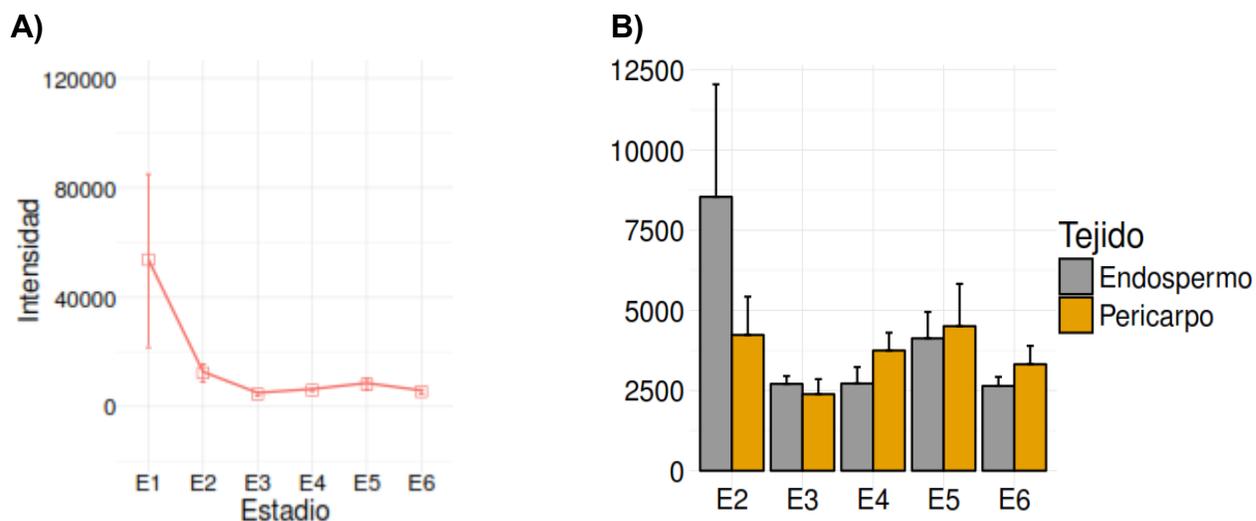


Figura 17. Cambios de los niveles de galactosa durante el desarrollo del fruto de café. A) Frutos completos, B) frutos separados por tejidos. Los valores graficados corresponden a la media de las 5 familias de plantas analizadas (F1-F5) \pm desviación estándar.

Clase III: Metabolismo y biosíntesis de ácidos grasos/ metabolismo de glicerofosfolípidos

El diacilglicerol (16:0/18:2) y (2E)-Decenoil-CoA presentaron un patrón muy similar en sus niveles durante el desarrollo del fruto, ya que éstos aumentaron en estadios tempranos (E1-E3) y en estadios posteriores se mantuvieron, observándose su disminución en el fruto cereza (Estadio E6) (Fig. 18 A, C y E).

El ácido palmítico es el primer ácido graso saturado producido en la biosíntesis de ácidos grasos y es precursor de ácidos grasos de cadena más larga. Este compuesto fue el que se detectó en mayor abundancia en el grupo de los ácidos grasos (Fig. 18 B), se observó aumento en sus niveles de manera dependiente del estadio de desarrollo y alcanzó sus máximos niveles en frutos rojos (Estadio E5), se encontró mayores cantidades de ácido palmítico en el endospermo de los frutos de café.

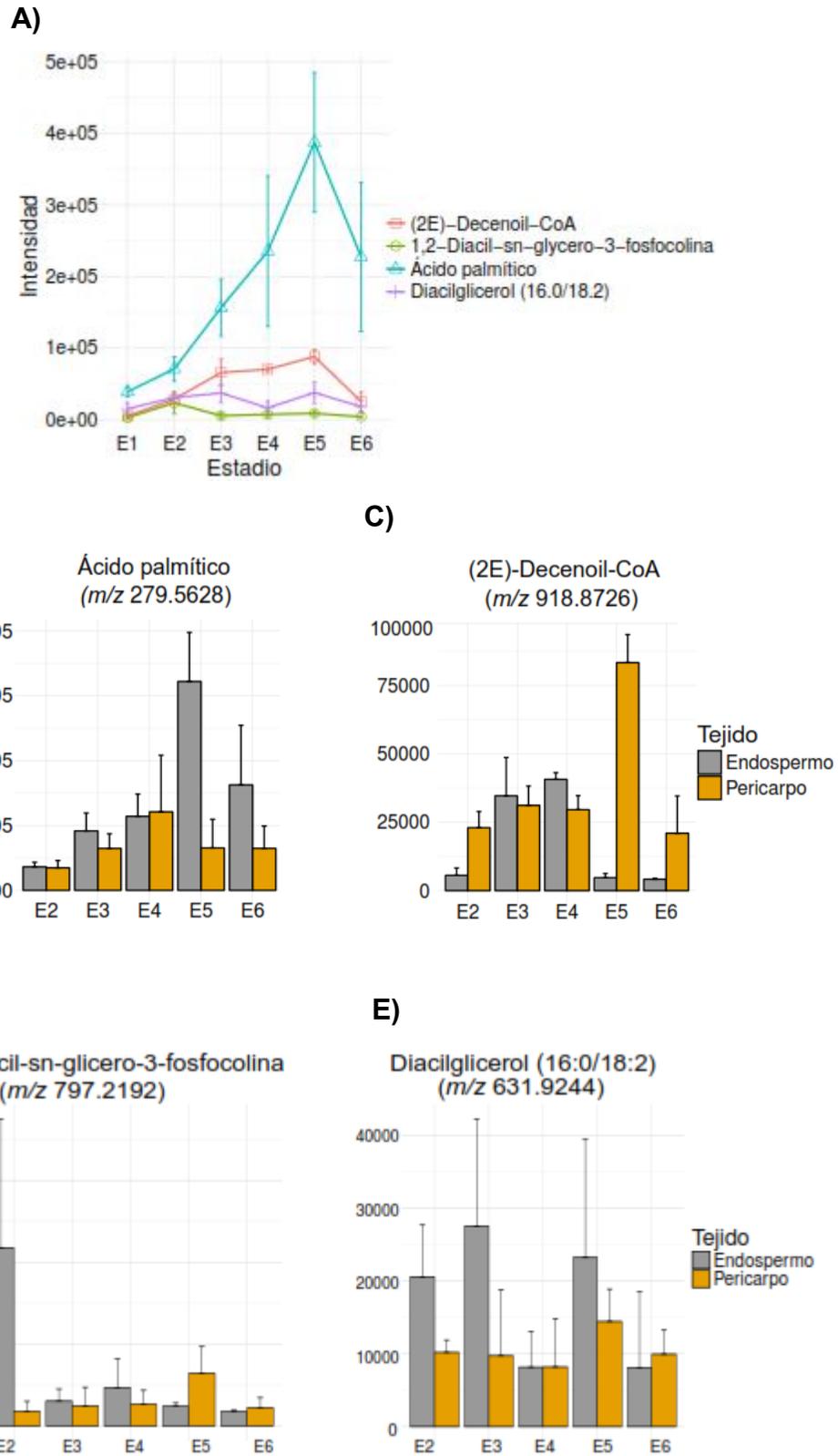


Figura 18. Cambios en los niveles de ácidos grasos durante el desarrollo del fruto de café. A) Niveles de ácidos grasos en frutos completos; Las figuras B - E corresponden a los niveles de los metabolitos en frutos separados por tejidos. Los valores graficados corresponden a la media de las 5 familias de plantas analizadas (F1-F5) \pm desviación estándar.

Clase IV: Metabolismo de porfirina y clorofila

En esta clase se identificó a la sirohidroclorina, la cual es intermediario en la biosíntesis del grupo sirohemo, un grupo prostético estructuralmente parecido al grupo hemo y a la clorofila. El grupo sirohemo forma parte de dos enzimas indispensables en las plantas, sulfito y nitrito reductasas, que participan en la asimilación de sulfatos y nitratos. La sirohidroclorina se observó en mayor abundancia en frutos de 30 semanas después de la floración (Fig. 19 A), durante este estadio y el estadio E4 se encontró mayormente en el endospermo de los frutos (Fig. 19 B). De manera dependiente de la maduración del fruto la sirohidroclorina disminuyó sus niveles los cuales fueron más bajos incluso que el estadio temprano E1.

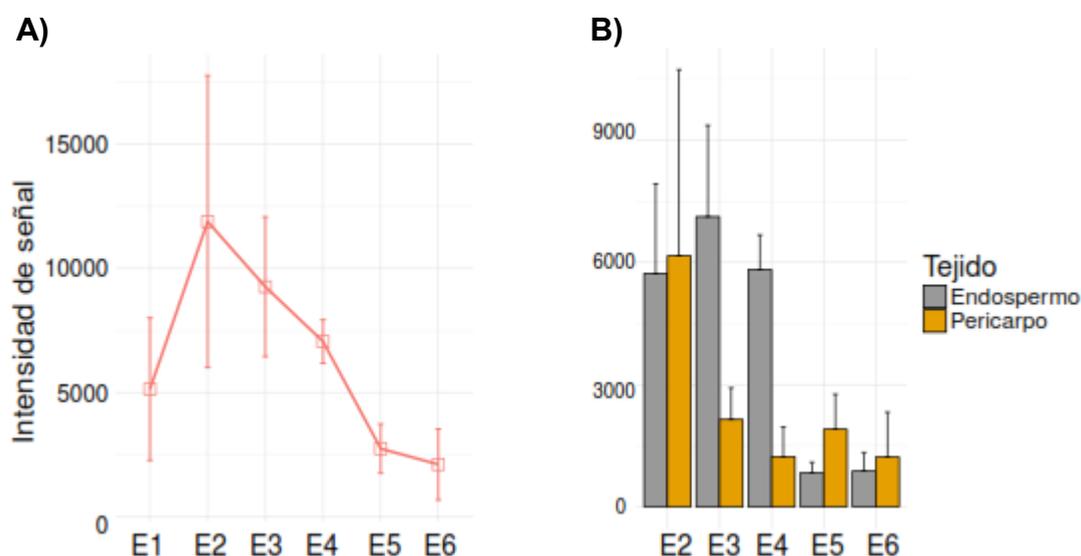


Figura 19. Cambios de los niveles de sirohidroclorina durante el desarrollo del fruto de café. A) Fruto completo y B) fruto separado por tejidos. Los valores graficados corresponden a la media de las 5 familias de plantas analizadas (F1-F5) ± desviación estándar.

Clase V: Metabolismo de inositol fosfato

El D-mio-inositol-1,2,5,6 tetrakisfosfato se encontró como uno de los metabolitos diferenciales durante el desarrollo del fruto.

Está involucrado en la vía de degradación del ácido fítico, el cual es el mio-inositol fosfato más abundante en plantas. El ácido fítico sirve como almacén para formar mio-inositol,

fosfato y nutrientes minerales que son empleados durante la germinación y el crecimiento de la semilla. En la figura 20 A se ilustra el patrón de los niveles del D-mio-inositol-1,2,5,6 tetrakisfosfato, el cual mostró aumento gradual hasta el estadio de fruto amarillo-rojo (Estadio E4), justamente cuando ha terminado el crecimiento y expansión del fruto e inicia propiamente la maduración del fruto, en esta etapa se observó mayor cantidad de este metabolito en el endospermo del fruto (Fig. 20 B). En los estadios posteriores de fruto maduro sus niveles disminuyeron a niveles semejantes al estadio E1.

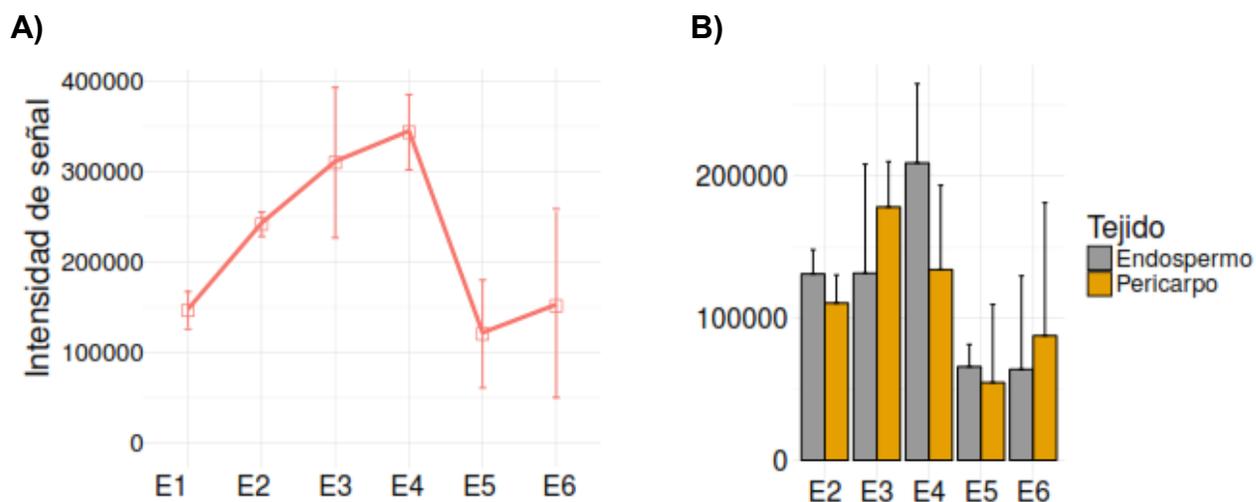


Figura 20. Cambios de los niveles de D-mio-inositol-1,2,5,6 tetrakisfosfato durante el desarrollo del fruto de café. A) Fruto completo y B) fruto separado por tejidos. Los valores graficados corresponden a la media de las 5 familias de plantas analizadas (F1-F5) \pm desviación estándar.

Clase VI: Metabolismo y biosíntesis de terpenoides y carotenoides

En esta clase se asignaron 2 compuestos pertenecientes al grupo de carotenoides: zeinoxantina (m/z 576.8081) y α -caroteno (m/z 537.8237).

Ambos compuestos mostraron un patrón similar en sus niveles de abundancia, y fueron encontrados en mayor cantidad en tejido de pericarpio. En los estadios de fruto verde (E2) y particularmente en fruto rojo (E6) se observaron los picos máximos de abundancia de estos dos carotenoides (Fig. 21 A, B y C).

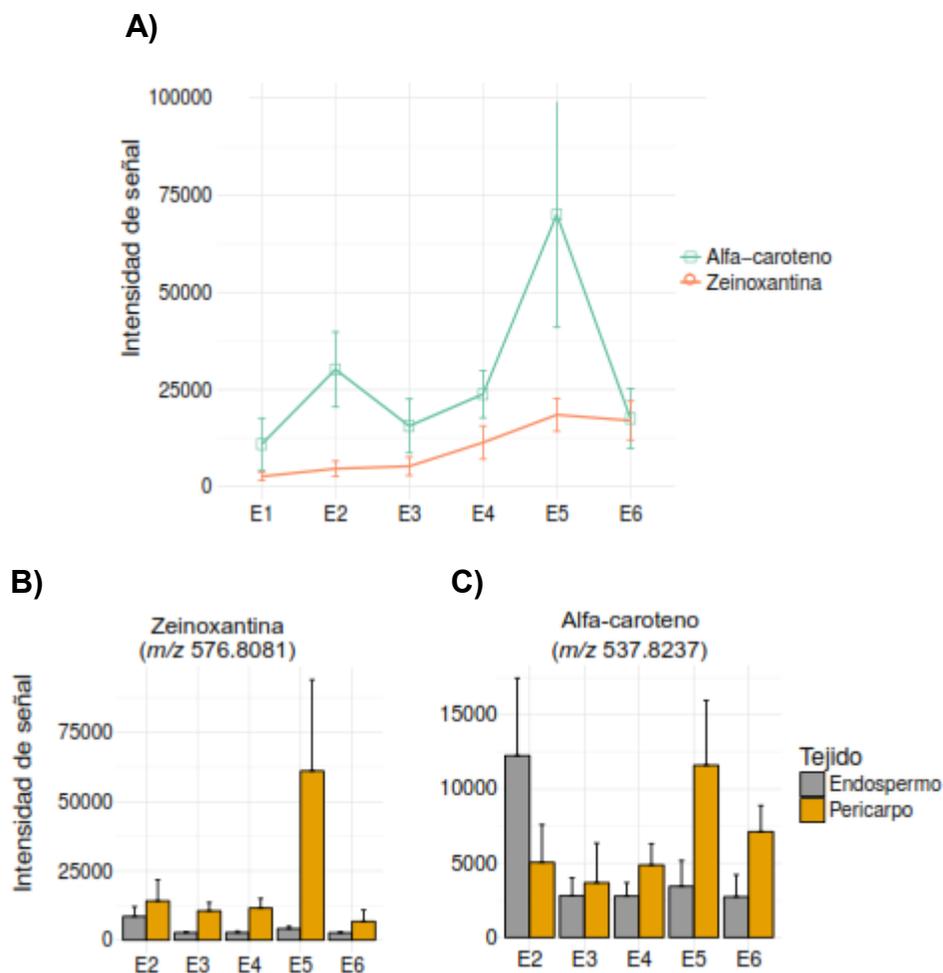


Figura 21. Cambios en los niveles de carotenoides durante el desarrollo del fruto de café. A) Frutos completos, B) y C) niveles de zeinoxantina y alfa-caroteno en pericarpio y endospermo. Los valores graficados corresponden a la media de las 5 familias de plantas analizadas (F1-F5) \pm desviación estándar.

En esta clase también se identificaron 7 compuestos involucrados en la biosíntesis de esqueleto de terpenoides: 16-O-metilcafestol, 4-carboxi-4metilzimosterol, ácido medicagénico, CDP-metil-D-eritritol 2-fosfato, Kahweol, octaprenil difosfato y trans-zeatin ribósido trifosfato (Fig. 22 A-H).

De manera general, en frutos completos la mayoría de estos compuestos aumentaron gradualmente sus niveles en la transición del estadio E1 a E3 (fruto amarillo), excepto el ácido medicagénico que del Estadio E1 al E3 disminuyó sus niveles y se mantuvo constante en los estadios de maduración del fruto (E4-E6). De manera interesante la mayoría de estos compuestos se detectaron con mayor abundancia en tejido de pericarpio en los estadios de frutos maduros (E5 y E6).

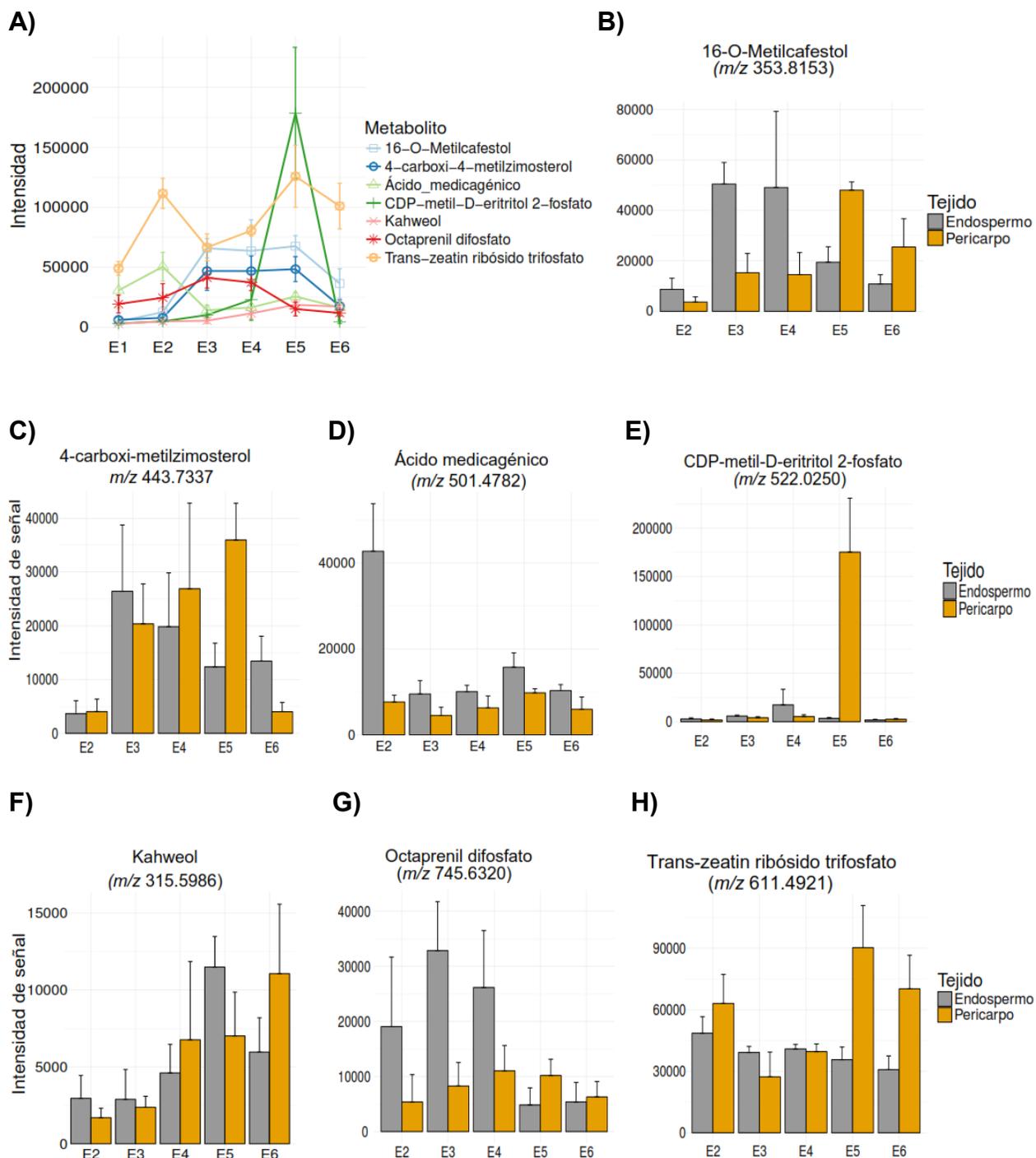


Figura 22. Cambios en los niveles de compuestos involucrados en metabolismo y biosíntesis de terpenoides durante el desarrollo del fruto. A) corresponde a niveles de los compuestos detectados en frutos completos. Las gráficas de barras B - H muestran la abundancia de los compuestos en tejido de pericarpo y endospermo. Los valores graficados corresponden a la media de las 5 familias de plantas analizadas (F1-F5) \pm desviación estándar.

Clase VII: Biosíntesis de fenilpropanoides, flavonoides y antocianinas

En esta clase se clasificaron 3 compuestos involucrados en la biosíntesis de antocianinas (Fig. 23 A). El cianidin-3-O-(6-O-p-coumaroil) glucósido presentó aumento dependiente del nivel de desarrollo del fruto (Fig. 23 B), siendo más abundante en el fruto rojo, ya que en el estadio final (E6) sus niveles nuevamente disminuyeron. El compuesto peonidin-3-O-glucósido se presentó en niveles casi indetectables durante los dos primeros estadios, sin embargo en el estadio E3 correspondiente a frutos amarillos aumentó drásticamente y alcanzó niveles máximos en el estadio de fruto rojo (Fig. 23 C). Por su parte, el delphinidin-3-O-glucosil-5-O-cafeoilglucósido aumentó sus niveles en la transición de fruto en expansión hasta el inicio de la maduración del fruto, en estadios posteriores sus niveles decayeron (Fig. 23 D).

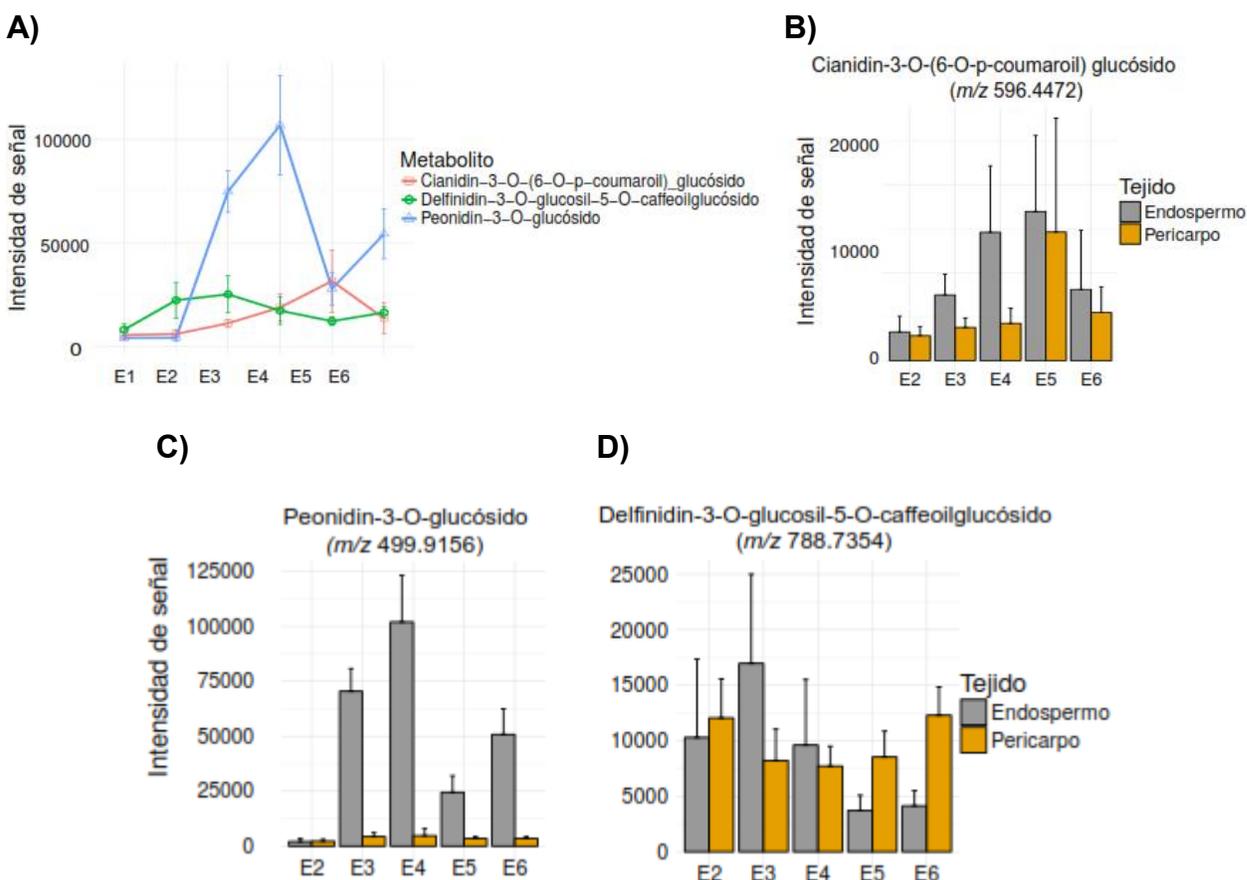


Figura 23. Cambios en los niveles de compuestos involucrados en la biosíntesis de fenilpropanoides, flavonoides y antocianinas durante el desarrollo del fruto. A) corresponde a niveles de los compuestos detectados en frutos completos. Las gráficas de barras B - H muestran la abundancia de los compuestos en tejido de pericarpio y endospermo. Los valores graficados corresponden a la media de las 5 familias de plantas analizadas (F1-F5) \pm desviación estándar.

6.1.3 Cianidin-3,5-diglucósido y cianidin-3-O-(6-O-p-coumaroil) como marcadores predictivos de estadio de desarrollo del fruto de café

De acuerdo al modelo predictivo generado con los datos metabolómicos se logró determinar que los iones m/z 596.4472 en frutos y m/z 611.9349 en hojas fueron las variables más importantes para discriminar entre estadios de desarrollo. Estas dos masas fueron confirmadas mediante los patrones de fragmentación de sus iones y mediante espectrometría de masas de alta resolución y se identificaron como cianidin-3-O-(6-O-p-coumaroil) glucósido y cianidin-3,5-diglucósido, respectivamente. Estos metabolitos mostraron los valores más altos de Mean Decrease Gini en el ranking de contribución de las variables obtenido mediante el modelo predictivo de random forest.

Como puede observarse en la figura 24 A, los niveles de Cianidin-3-O-(6-O-p-coumaroil) glucósido en frutos son diferenciales, ya que claramente aumentan en asociación con el nivel de desarrollo del fruto. En hojas los resultados también son concluyentes, sin embargo en este caso se observan niveles bajos de la antocianina cianidin-3,5-diglucósido en los estadios tempranos E1 y E2, y niveles hasta cuatro veces más altos en frutos de semimaduros a maduros (Fig. 24 B). Por lo tanto, se puede determinar que el cianidin-3,5-diglucósido es un factor discriminante entre del grado de madurez de los frutos de café. Este es un hallazgo interesante ya que este hecho fue posible a partir del análisis de tejidos de hojas.

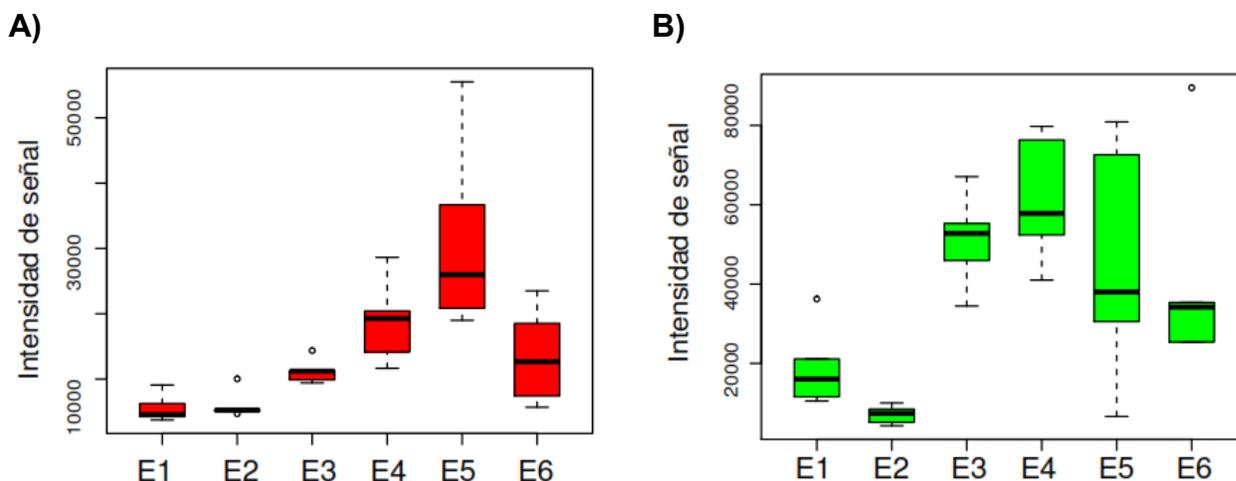


Figura 24. Cambios en los niveles de compuestos marcadores de desarrollo del fruto A) niveles de Cianidin-3-O-(6-O-p-coumaroil) glucósido en frutos y B) niveles de cianidin-3,5-diglucósido.

En conclusión, se propone como marcadores específicos de estadio de desarrollo y madurez del fruto de café a cianidin-3-O-(6-O-p-coumaroil) glucósido y cianidin-3,5-diglucósido, estos compuestos podrían ser implementados en rutinas de análisis de la calidad de productos de café.

6.2 Análisis de Heredabilidad metabólica

El entendimiento acerca de en qué magnitud la variabilidad de los perfiles metabólicos está regulada genéticamente y cómo éstos se relacionan con las características fenotípicas de las plantas es uno de los retos en vías de desarrollar una estrategia dirigida para el mejoramiento genético de cultivos. En cultivos perennes como el café resulta de gran importancia evaluar estos aspectos en la búsqueda de acortar los tiempos de selección. La predicción de manera temprana de las características agronómicas deseables resultaría de gran utilidad para apoyar las estrategias clásicas de mejoramiento genético.

Por ello, una parte importante del presente trabajo fue la estimación de heredabilidad metabólica en frutos y hojas de plantas de *C. canephora*, teniendo como objetivo la selección de aquellos metabolitos que presenten altos niveles de heredabilidad.

La razón de la estimación de heredabilidad de metabolitos en hojas representa la posibilidad de identificar metabolitos altamente heredables en estadios tempranos del desarrollo de la planta. La anticipando a la producción de frutos que se presenta entre 3.5 – 5 años después de la plantación en campo, podría representar la optimización de los tiempos de selección.

Como se mencionó en el apartado de metodología para cumplir con este objetivo, en este estudio se incluyeron 40 familias de plantas de la especie *C. canephora* (la cual se considera una población en recombinación), en total se analizaron 120 plantas (Tabla 13). Cada familia estuvo conformada por 3 integrantes: un parental (planta madre) y dos hijas por cada parental. La progenie (80 plantas) es una población de medios hermanos maternos que tienen en común la planta madre de la cual fueron obtenidas las semillas para su reproducción, por lo tanto se desconoce la identidad del otro parental ya que la progenie fue generada a partir de semillas obtenidas por polinización abierta.

Tabla 13. Familias incluidas en el análisis de heredabilidad

Número	Familia	Planta madre	Hija 1	Hija 2
1	196	PM196	196-1	196-3
2	203	PM203	203-1	203-2
3	205	PM205	205-1	205-2
4	212	PM212	212-1	212-2
5	292	PM292	292-3	292-4
6	295	PM295	295-2	295-3
7	312	PM312	312-3	312-4
8	315	PM315	315-1	315-2
9	316	PM316	316-1	316-2
10	318	PM318	318-2	318-3
11	322	PM322	322-3	322-4
12	323	PM323	323-1	323-4
13	328	PM328	328-3	328-4
14	333	PM333	333-1	333-4
15	336	PM336	336-1	336-2
16	337	PM337	337-3	337-4
17	339	PM339	339-1	339-2
18	340	PM340	340-1	340-3
19	343	PM343	343-1	343-3
20	346	PM346	346-1	346-4
21	349	PM349	349-2	349-3
22	360	PM360	360-3	360-4
23	390	PM390	390-1	390-3
24	408	PM408	408-1	408-4
25	416	PM416	416-1	416-3
26	418	PM418	418-3	418-4
27	419	PM419	419-2	419-4
28	420	PM420	420-1	420-4
29	430	PM430	430-2	430-3
30	432	PM432	432-1	432-2
31	433	PM433	433-2	433-3
32	434	PM434	434-2	434-4
33	442	PM442	442-2	442-3
34	443	PM443	443-3	443-4
35	449	PM449	449-1	449-3
36	450	PM450	450-3	450-4
37	451	PM451	451-1	451-4
38	452	PM452	452-2	452-3
39	453	PM453	453-2	453-3
40	454	PM454	454-2	454-4

6.2.1.1 Variabilidad de fenotipos metabólicos y parentesco entre plantas madre

Las población de plantas madre incluidas en el estudio corresponden a plantas de 15 años de edad, se desconoce el parentesco genéticos entre ellas. Debido a estos antecedentes, el primer análisis que se realizó consistió en establecer el parentesco metabólico y la variabilidad existente entre las 40 plantas madre.

Para ello se llevó a cabo un análisis no dirigido basado en UPLC-MS/MS a partir de extractos metanólicos-acuosos para determinar el perfil metabólico de cada una de las plantas madre. Tomando como variables las abundancias de cada uno de los metabolitos detectados en los extractos de frutos, se evaluó mediante agrupamiento jerárquico el parentesco metabólico entre ellas.

Como se observa en la figura 25 se identificaron tres grandes clados. El primer clado estuvo formado por 10 integrantes relacionados muy cercanamente en cuanto a su perfil metabólico. Las plantas madre PM419, PM318, PM434 y PM450 formaron uno de los grupos de este clado mientras que en una ramificación diferente se agruparon las muestras PM416, PM418, PM337, PM452, PM430 y PM443.

El segundo clado fue el que estuvo constituido por un número mayor de muestras (16), lizadas, se formaron tres subgrupos soportados por más de 75 % de probabilidad. De manera aislada dentro de este clado se localizó el grupo conformado por se agruparon las muestras PM316, PM340 y PM312 (Fig. 25).

En el tercer clado se observaron dos ramas principales, en la primer ramificación de muestra un mayor nivel de semejanza metabólica entre las 10 plantas madre que formaron parte de este grupo. En la otra ramificación más distantes metabólicamente, se agruparon las plantas madre PM203, PM433, PM453 Y PM420 (Fig. 25).

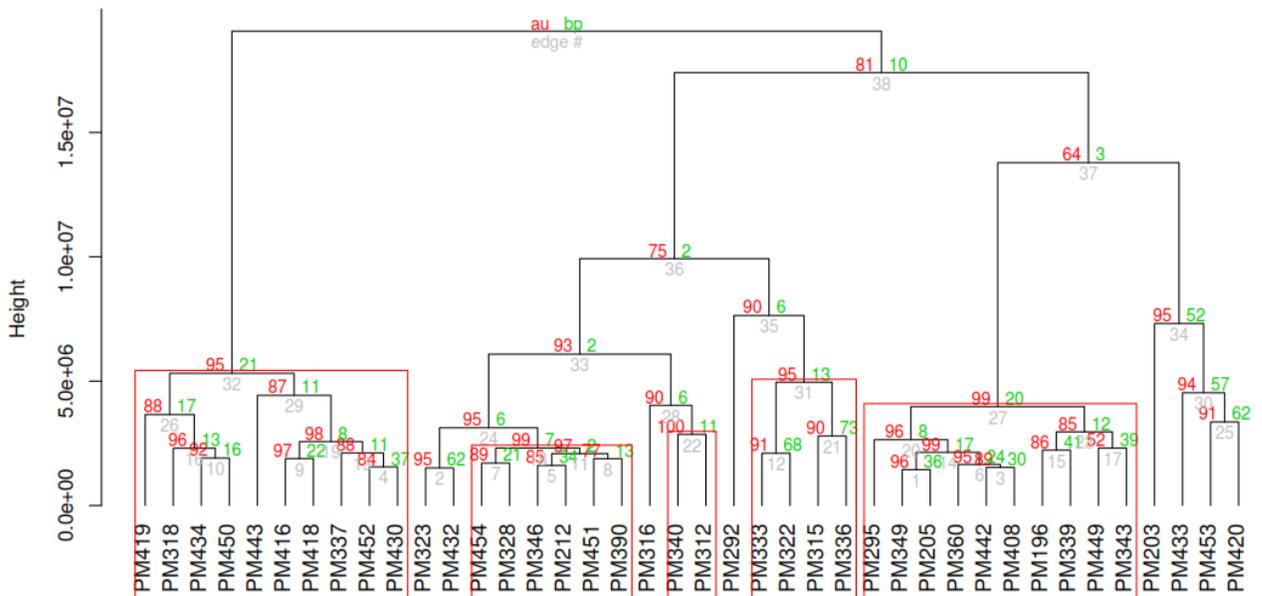


Figura 25. Dendrograma de 40 plantas madre. El agrupamiento jerárquico se realizó mediante aglomeración euclidiana y distancias establecidas por método Ward. Los recuadros rojos indican agrupamientos con probabilidad > 95 %.

De la misma manera, se realizó el análisis no supervisado de agrupamiento jerárquico de las plantas F1 (Fig. 26). Las plantas hija fueron codificados con tres dígitos que corresponden a la planta madre de la cual provienen y el número adicional representa el número de planta hija muestreada en el campo experimental (ver Tabla 13).

En primera instancia, el dendrograma muestra un claro agrupamiento debido al parentesco metabólico presente entre plantas hija que comparten un parental (Fig. 26), en este caso la planta madre de la que fueron reproducidas por semilla.

Se puede observar la formación de un clado pequeño formado por 5 de las familias analizadas mientras que en el otro gran clado se localizaron el resto de las familias. En el primer (clado color rosa) se agruparon las plantas hija de las familias 315, 292, 336, 322 y 333, de manera interesante en esta rama del dendrograma agrupó a 5 familias que habían mostrado estrecho parentesco metabólico en los perfiles metabólicos de las plantas madre (Fig. 26).

En otro subgrupo (rama color verde), se agruparon las familias 419, 434, 450, 443, 416, 418, 337, 452, 430. La relación metabólica entre estas familias, se observó al agrupar las plantas madre, con excepción de la familia 318 que en este caso mostró mayor cercanía metabólica con las familias 442 y 432 (ramificación color azul) (Fig. 26).

En algunos casos, dada la gran similitud que existe entre algunas familias a nivel de perfil metabólico, no es posible distinguir entre medios hermanos y éstos se agrupan con familias con las que están estrechamente relacionadas.

Estos resultados son relevantes y reflejan que aún en plantas de la progenie F1 es posible detectar el parentesco que se demostró entre plantas madre.

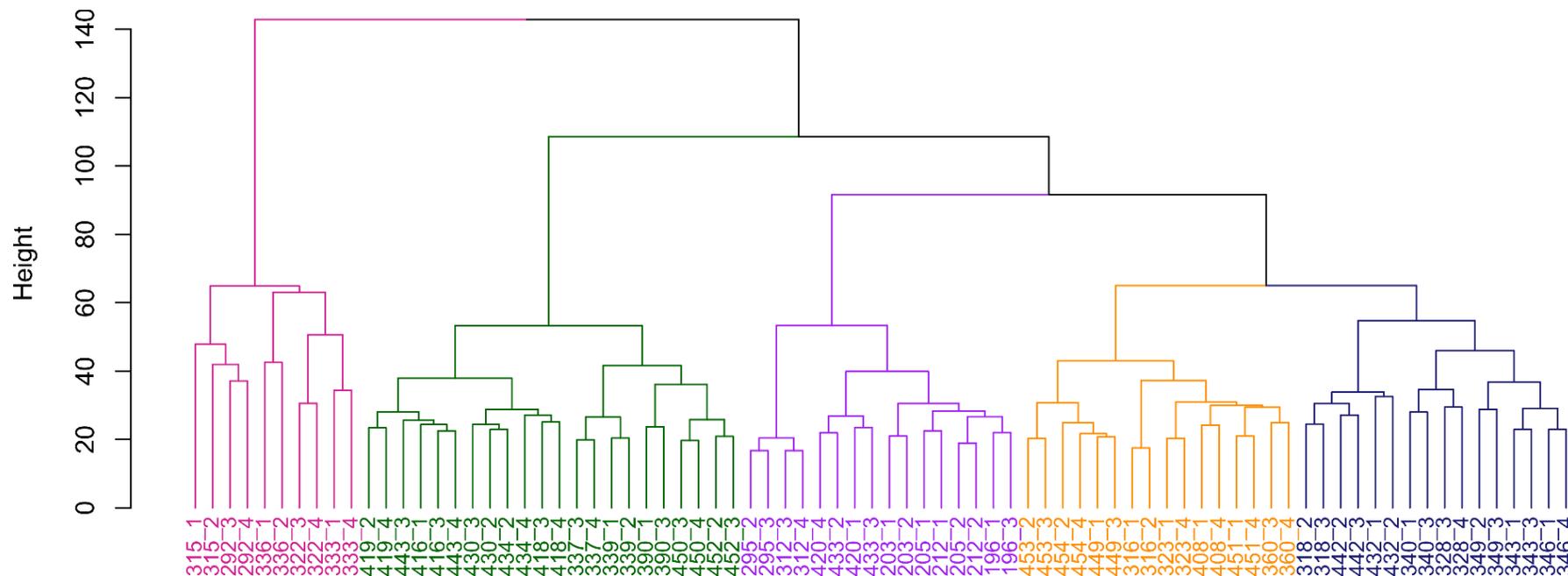


Figura 26. Dendrograma de progenie F1. Se realizó un agrupamiento jerárquico con aglomeración euclidiana y distancias establecidas por método Ward. Se observa la formación de dos clados principales, el más grande (lado derecho) agrupó el 87.5 % de las muestras.

6.2.1.2 Estimación de heredabilidad en frutos y hojas

Mediante UPLC-MS se analizaron los perfiles metabólicos de 40 familias (120 plantas de *C. canephora*). En total, a partir de frutos se detectaron 755 metabolitos que estuvieron presentes en todas las muestras, mientras que en hojas se detectó un número menor, 527 metabolitos.

Para evaluar el porcentaje de la variación en los niveles de los metabolitos que puede ser atribuible al factor genético, se emplearon los niveles de intensidad de cada metabolito para estimar heredabilidad metabólica en sentido estricto. Los niveles de heredabilidad se determinaron mediante el cálculo del coeficiente de regresión de la regresión lineal progenie-parentales aplicando la fórmula que se describió en el apartado de metodología. Como puede observarse en la figura 27 A, el 12.98 % (98 metabolitos) presentaron niveles bajos de heredabilidad (0 - 0.2), mientras que el 27.42 % (207 metabolitos) presentaron heredabilidad moderada. El 59.6 % (450 metabolitos) de los metabolitos mostraron estar principalmente influenciados por el factor genético ya que presentaron niveles de heredabilidad mayores de 0.4 lo que se considera como heredabilidad alta. Este último grupo de metabolitos fue considerado para seguir con el análisis de correlación con los atributos sensoriales de la bebida.

Al 53.77 % de los metabolitos (242 / 450) fue posible asignarle identidad, en el anexo 11.4 se enlistan los 450 metabolitos con sus respectivos niveles de heredabilidad y error estándar.

En hojas el patrón de heredabilidad de los metabolitos fue diferente (Fig. 27 B), 286 de los 527 metabolitos (54.27 %) presentaron niveles de heredabilidad bajos, el 23,72 % (125 metabolitos) se encontraron con niveles de heredabilidad moderada, mientras que para 116 metabolitos (22.01 %) los niveles de heredabilidad fueron mayores a 0.4.

Con la finalidad de establecer si existe correlación entre los 566 metabolitos altamente heredables previamente identificados (450 metabolitos de frutos y 116 de hojas), se realizó posteriormente un análisis de correlación de Spearman entre estos metabolitos y las notas sensoriales de la bebida de café evaluadas.

Con este fin, en el siguiente apartado se describirá primero de qué manera se llevó a cabo el análisis sensorial de la bebida de café y los resultados obtenidos.

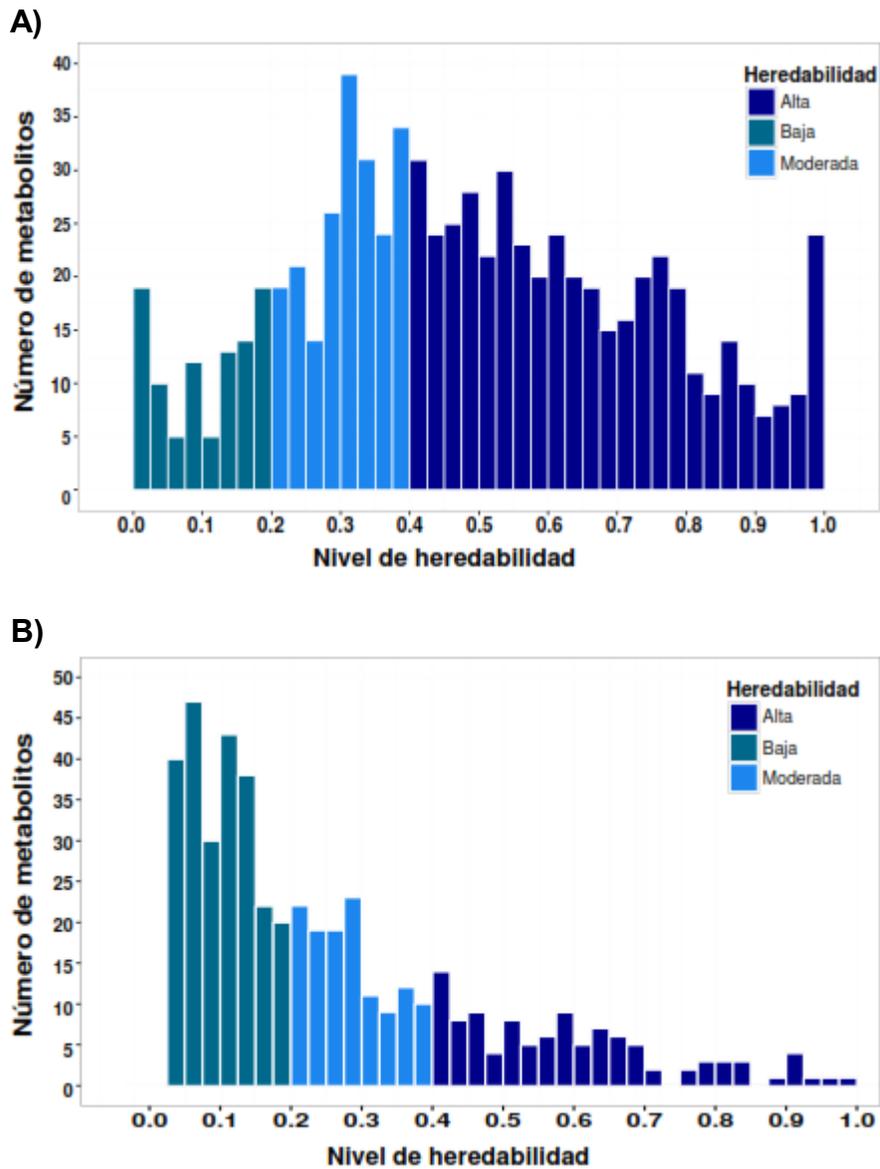


Figura 27. Histograma de heredabilidad de metabolitos estimada en frutos (A) y hojas (B). Heredabilidad Alta: ≥ 0.4 ; moderada: $0.2 - 0.4$; baja: < 0.2 .

6.3 Correlación de perfiles sensoriales de la bebida-metabolitos altamente heredables de frutos y hojas

6.3.1 Evaluación sensorial

Las bebidas preparadas a partir de los granos tostados y molidos de las 21 muestras incluidas en el estudio fueron evaluadas por un panel de 11 catadores entrenados. Se incluyeron 14 muestras correspondientes a *C. canephora* (M480 - M495) y 7 muestras de referencia de la especie *C. arabica* (M498 - M508). Mediante el análisis sensorial descriptivo se calificaron 10 atributos sensoriales de la bebida (Intensidad aromática, calidad aromática, sabor, astringencia, resabio, acidez, amargor, cuerpo, agrio y preferencia) empleando una escala de 0 - 5 puntos. Los resultados de la evaluación de cada una de las muestras se muestran en la tabla 14.

Como puede observarse en la figura 28, las muestras evaluadas presentan diversidad en sus perfiles sensoriales, sin embargo se identifican evidentes diferencias debidas principalmente a la especie evaluada (*C. canephora* o *C. arabica*). Las muestras de *C. arabica* presentaron mayor grado de acidez comparado con las muestras de *C. canephora*. Contrario a lo que se observó en el caso de amargor, ya que en las muestras de la especie *C. canephora* la intensidad de este atributo fue más alta comparado con *C. arabica*. En cuanto al nivel de agrio y cuerpo, las muestras de *C. canephora* presentaron un comportamiento más homogéneo.

Se encontró mayor variabilidad entre todas las muestras para los atributos sensoriales preferencia, intensidad aromática, calidad aromática, sabor, astringencia y resabio (Fig. 28).

Las diferencias observadas entre especie se ven reflejadas visualmente en el análisis de PCA (Fig. 29) en el cual los componentes principales 1 y 2 explican 75.7 % de la variabilidad de los datos. Esto permite diferenciar claramente entre las muestras de *C. arabica* y *C. canephora* indicando las diferencias marcadas que existen entre los perfiles sensoriales de cada una de estas especies.

Tabla 14. Perfil sensorial descriptivo de las bebidas. Los datos representan los promedios de 11 catadores \pm desviación estándar.

Muestra	Intensidad aromática	Calidad aromática	Sabor	Astringencia	Resabio	Acidez	Amargor	Cuerpo	Agrio	Preferencia
M480	3.27 \pm 0.79	3.09 \pm 0.49	2.73 \pm 0.66	2.45 \pm 0.94	3.27 \pm 0.71	0.64 \pm 0.57	3.45 \pm 0.76	2.73 \pm 0.52	0.27 \pm 0.44	2.82 \pm 0.77
M481	2.14 \pm 0.48	1.86 \pm 0.85	2.23 \pm 0.88	2.05 \pm 0.60	2.27 \pm 0.52	0.55 \pm 0.42	3.50 \pm 0.54	2.23 \pm 0.61	0.36 \pm 0.52	2.14 \pm 0.87
M482	3.18 \pm 0.74	3.45 \pm 0.95	3.09 \pm 0.66	2.82 \pm 0.62	3.00 \pm 0.54	0.45 \pm 0.66	2.82 \pm 0.99	2.82 \pm 0.47	0.55 \pm 0.79	3.09 \pm 0.67
M483	2.55 \pm 0.67	2.64 \pm 1.12	2.18 \pm 0.74	2.45 \pm 0.67	2.55 \pm 0.77	0.55 \pm 0.59	3.18 \pm 0.80	2.36 \pm 0.87	0.45 \pm 0.66	2.36 \pm 0.64
M484	2.64 \pm 0.71	2.45 \pm 0.76	2.91 \pm 0.49	2.36 \pm 0.76	2.64 \pm 0.64	0.64 \pm 0.8	3.73 \pm 0.66	2.55 \pm 0.76	0.45 \pm 0.66	2.91 \pm 0.67
M485	2.82 \pm 0.95	2.64 \pm 0.94	2.36 \pm 0.52	3.09 \pm 0.67	3.00 \pm 0.72	0.55 \pm 0.69	3.55 \pm 0.95	2.73 \pm 0.57	0.45 \pm 0.66	2.55 \pm 0.67
M486	3.00 \pm 0.72	3.09 \pm 0.82	3.09 \pm 0.67	1.73 \pm 0.84	2.64 \pm 0.69	0.36 \pm 0.52	3.55 \pm 0.59	2.36 \pm 0.99	0.18 \pm 0.33	3.09 \pm 0.85
M488	2.27 \pm 0.66	2.09 \pm 1.02	2.00 \pm 1.09	2.09 \pm 0.67	2.73 \pm 0.76	0.45 \pm 0.74	3.45 \pm 0.94	2.73 \pm 0.71	1.00 \pm 1.1	2.00 \pm 0.9
M489	2.64 \pm 0.82	2.36 \pm 0.87	1.82 \pm 0.92	2.18 \pm 0.62	2.55 \pm 0.77	1.27 \pm 1.07	2.36 \pm 0.82	2.45 \pm 0.77	1.27 \pm 1.43	1.82 \pm 1.10
M490	3.27 \pm 0.66	2.82 \pm 0.44	2.36 \pm 0.52	2.82 \pm 0.29	3.00 \pm 0.54	0.27 \pm 0.49	3.36 \pm 0.69	2.73 \pm 0.71	0.36 \pm 0.52	2.82 \pm 0.8
M491	2.73 \pm 0.71	3.00 \pm 0.9	3.00 \pm 0.54	2.55 \pm 0.67	2.91 \pm 0.33	1.18 \pm 0.8	3.36 \pm 1.23	2.36 \pm 0.57	0.55 \pm 0.69	2.45 \pm 1.14
M492	2.45 \pm 0.67	1.82 \pm 0.77	1.82 \pm 0.77	2.45 \pm 0.59	2.73 \pm 0.52	1.00 \pm 0.9	3.09 \pm 1.02	2.64 \pm 0.57	0.55 \pm 0.69	1.91 \pm 0.85
M493	2.91 \pm 0.49	2.00 \pm 0.18	1.91 \pm 0.68	2.73 \pm 0.57	2.55 \pm 0.57	0.73 \pm 0.79	2.91 \pm 0.69	2.82 \pm 0.44	0.64 \pm 0.7	1.91 \pm 0.67
M495	2.36 \pm 0.64	1.45 \pm 0.76	1.45 \pm 0.68	1.55 \pm 0.59	1.73 \pm 0.84	0.82 \pm 0.74	2.27 \pm 0.97	2.27 \pm 0.84	0.91 \pm 1.0	1.73 \pm 1.02
M498	2.73 \pm 0.89	1.91 \pm 0.99	1.36 \pm 1.12	2.27 \pm 0.84	2.09 \pm 1.02	3.00 \pm 0.55	1.91 \pm 0.67	2.27 \pm 0.84	2.00 \pm 0.9	0.45 \pm 0.57
M499	3.09 \pm 0.82	1.82 \pm 1.10	1.55 \pm 0.95	2.18 \pm 0.95	1.73 \pm 0.52	2.09 \pm 1.19	2.18 \pm 0.89	2.55 \pm 0.67	1.18 \pm 0.99	1.18 \pm 0.95
M500	3.00 \pm 0.5	2.38 \pm 0.87	2.25 \pm 1.06	2.50 \pm 0.87	2.50 \pm 0.75	3.38 \pm 0.78	2.25 \pm 0.93	2.63 \pm 0.62	1.63 \pm 1.3	1.75 \pm 0.81
M502	3.43 \pm 0.65	2.29 \pm 0.89	1.71 \pm 0.69	1.43 \pm 0.77	2.43 \pm 0.93	3.14 \pm 0.73	2.29 \pm 0.89	1.86 \pm 0.48	1.14 \pm 1.06	1.43 \pm 0.77
M506	2.88 \pm 0.68	1.25 \pm 0.56	1.50 \pm 1.25	2.00 \pm 0.75	2.50 \pm 0.75	3.88 \pm 0.66	1.75 \pm 1.25	2.00 \pm 0.75	3.13 \pm 1.4	0.75 \pm 0.75
M507	2.57 \pm 0.77	1.29 \pm 0.61	1.57 \pm 0.94	2.29 \pm 0.69	1.57 \pm 0.49	4.00 \pm 0.57	1.43 \pm 0.77	1.71 \pm 0.81	1.43 \pm 1.5	0.43 \pm 0.61
M508	2.71 \pm 0.57	2.00 \pm 0.8	2.00 \pm 0.55	1.71 \pm 0.81	1.43 \pm 0.49	3.86 \pm 0.78	1.29 \pm 1.26	2.14 \pm 0.77	1.14 \pm 1.3	1.00 \pm 0.85

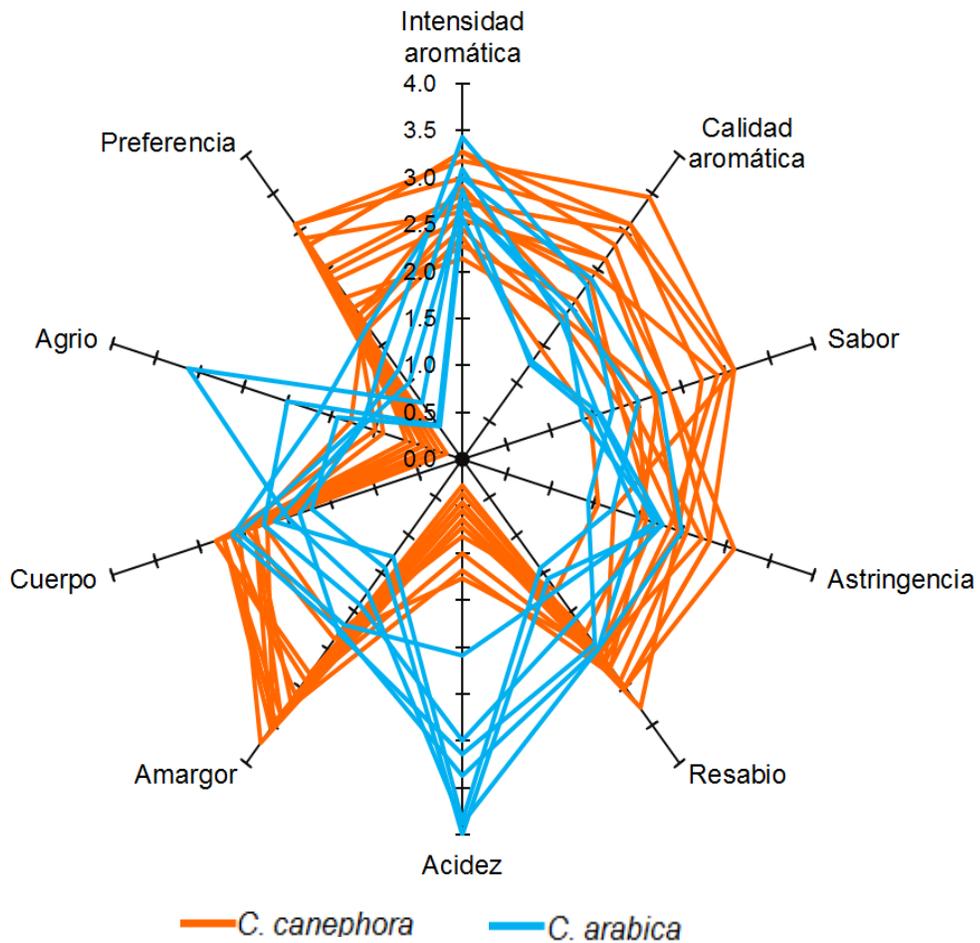


Figura 28. Gráfica de radar de perfiles sensoriales. Cada uno de los valores graficados representa el promedio de 11 jueces catadores. Se observan claras diferencias entre el nivel de acidez de las muestras de la especie *C. arabica* vs *C.canephora*. Los niveles de amargor de las muestras de la especie *C. canephora* fueron más altos comparados con los observados para las muestras de *C. arabica*.

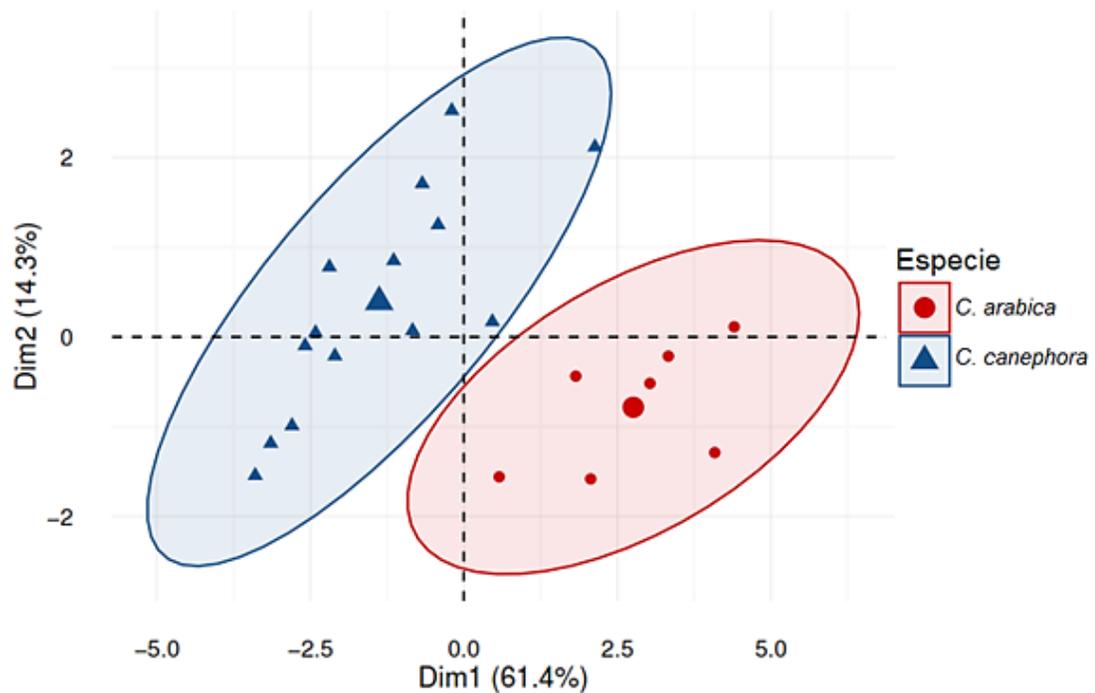


Figura 29. Análisis de componentes principales (PCA) de perfiles sensoriales. Mediante los perfiles sensoriales de las bebidas fue posible discriminar entre especies (*C. canephora* vs *C. arabica*).

Las variables (atributos sensoriales) que más contribuyen para diferenciar entre especies fueron preferencia, amargor y acidez (Fig. 30).

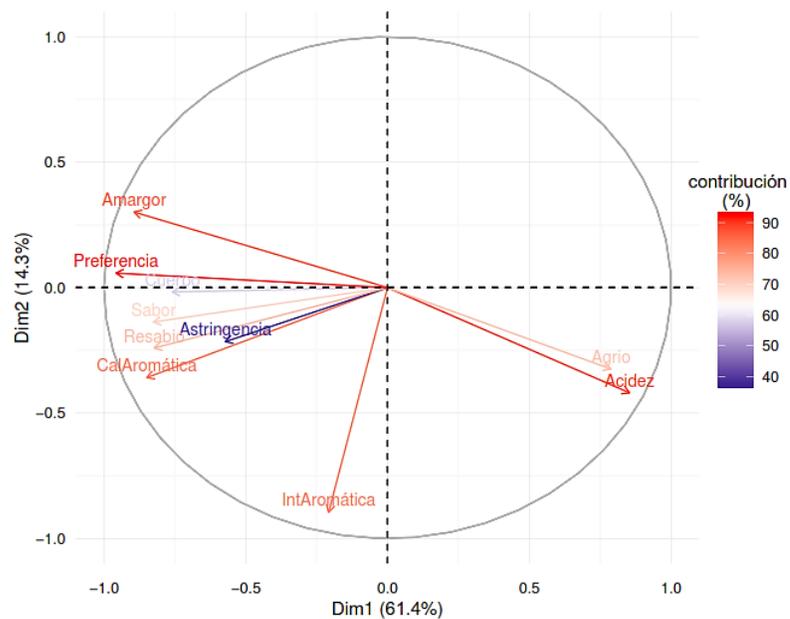


Figura 30. Contribución de variables (caracteres sensoriales) a los componentes 1 y 2. La escala de colores muestra el porcentaje de contribución de cada una de las variables.

Se realizó un análisis de correlación entre atributos sensoriales (Fig. 31), para la mayoría de las comparaciones los coeficientes de correlación de Spearman (CCS) calculados fueron estadísticamente significativos ($p < 0.05$). El atributo intensidad aromática únicamente mostró correlación positiva con calidad aromática (CCS=0.48; $p < 0.05$). Para la calidad aromática de la bebida se encontró correlación positiva con la mayoría de los atributos excepto con acidez (CCS: - 0.55; $p < 0.01$) y agrio (CCS: - 0.62; $p < 0.01$) con los cuales se encontró negativamente correlacionada (Fig. 31). Este mismo patrón de correlación se encontró para los atributos de sabor, astringencia y resabio, mostrando correlación negativa únicamente con acidez y agrio. De manera interesante, el atributo acidez mostró correlación negativa con todos los atributos sensoriales evaluados con excepción del atributo agrio, el coeficiente de correlación de Spearman fue de 0.87 ($p < 0.001$). El atributo preferencia representa la calificación de la percepción global de las características organolépticas de la bebida. Este atributo mostró correlación positiva con calidad aromática, sabor, astringencia, resabio, amargor y cuerpo (Fig. 31), sin embargo se encontró alta correlación negativa estadísticamente significativa con los atributos acidez y agrio (CCS: -0.85 y -0.8; $p < 0.001$, respectivamente).

6.3.2 Análisis metabolómico e identificación de metabolitos marcadores de calidad sensorial en granos tostados de café

Con el objetivo de identificar metabolitos marcadores de calidad sensorial en granos tostados, se analizaron extractos acuosos de granos tostados y molidos (como se detalló en metodología) mediante UPLC-ESI-MS. De esta manera, se detectaron 737 picos con su respectiva m/z y nivel de abundancia en cada una de las 21 muestras de *C. arabica* y *C. canephora* que se incluyeron en este análisis.

Haciendo uso de esta información se realizó un análisis de correlación de Spearman entre los niveles de cada uno de los metabolitos detectados en las muestras *versus* las calificaciones de cada uno de los atributos sensoriales asignadas por el panel de catación. En total se detectaron 206 metabolitos que cumplieron con los filtros establecidos para ser seleccionados como marcadores asociados a los atributos sensoriales. Lo cual consistió en presentar significancia estadística ($p < 0.05$) y CC Spearman > 0.5 en caso de correlación positiva y < -0.5 en caso de correlación negativa.

Se detectaron 469 asociaciones entre los 206 metabolitos y los caracteres sensoriales (Fig. 32).

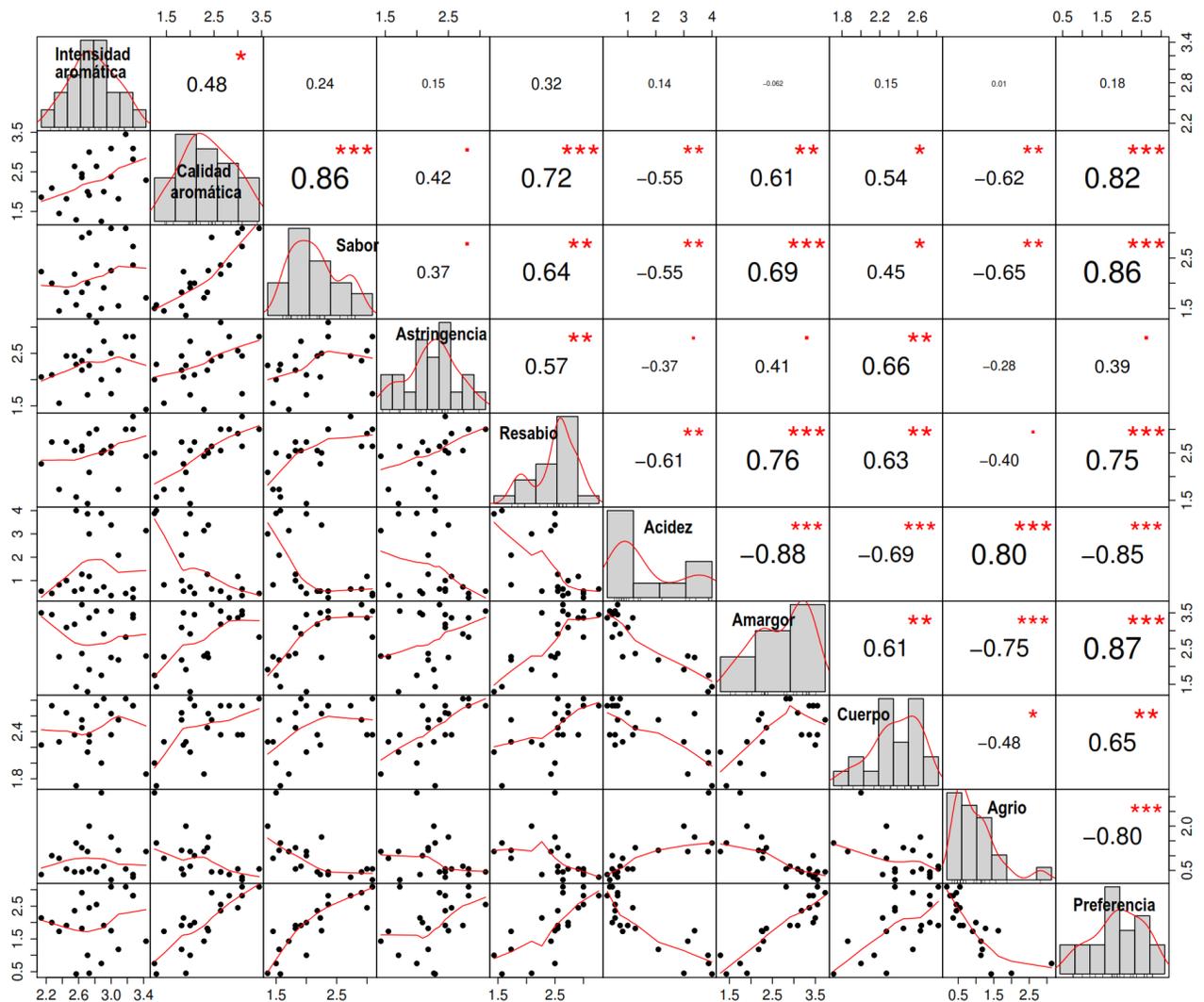


Figura 31. Gráfico de correlación entre caracteres sensoriales. Los niveles de correlación mostrados para cada una de las comparaciones corresponden al valor del coeficiente correlación Spearman entre los caracteres sensoriales. El nivel de significancia estadística está representada con asteriscos: * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

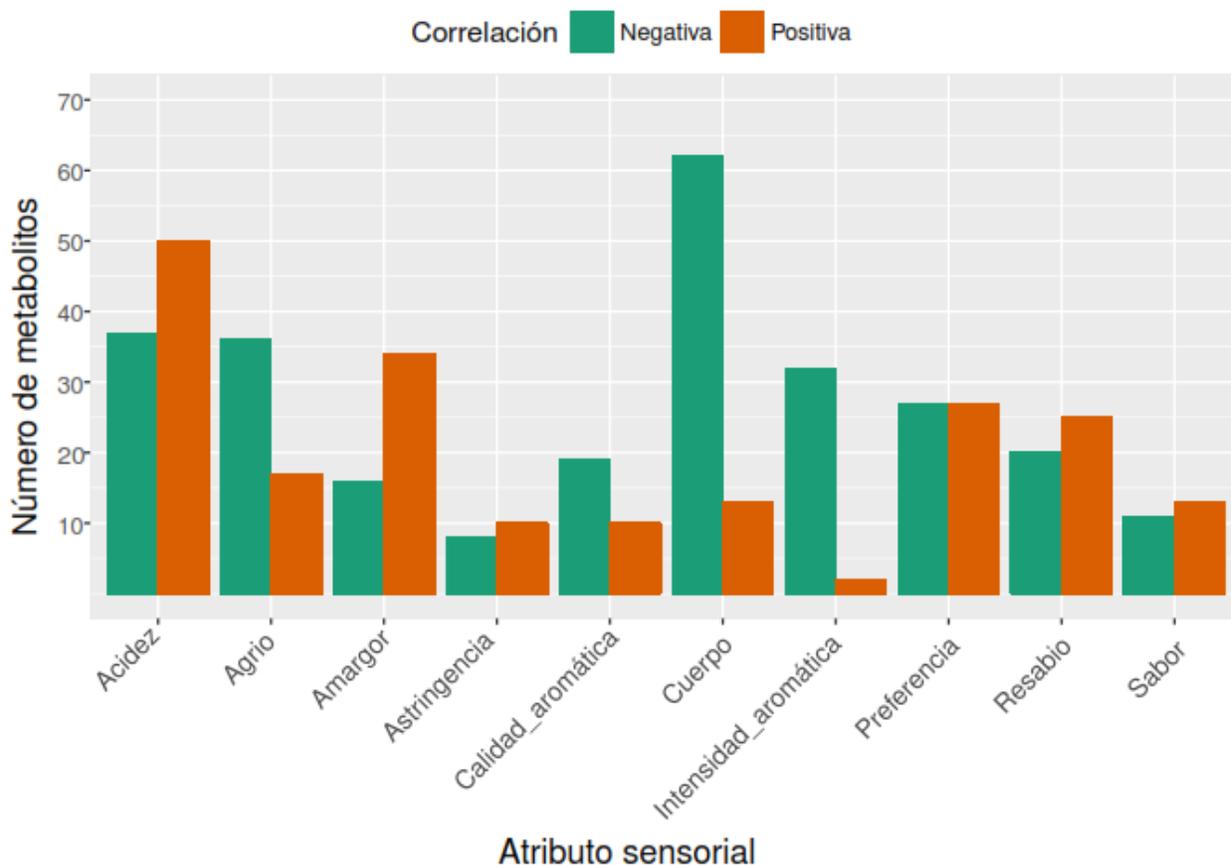


Figura 32. Metabolitos correlacionados con atributos sensoriales de la bebida. Significancia estadística : $p \leq 0.05$; Coeficiente de correlación de Spearman > 0.5 (correlación positiva), y < -0.5 (correlación negativa).

En el anexo 11.5 se enlistan cada una de las masas detectadas con sus respectivos valores de CC Spearman así como su nivel de significancia estadística.

De manera ilustrativa, en la figura 33, se muestra el análisis de correlación entre el ion m/z 727.9535 y 6 de los atributos sensoriales evaluados en la bebida. A esta masa no fue posible asignarle una identidad, sin embargo su patrón de asociación es muy interesante ya que se encontró asociado a múltiples atributos de la bebida. Como puede observarse en la figura el ion m/z 727.9535 mostró únicamente fuerte correlación positiva con el atributo agrio, mientras que mostró correlación negativa para los atributos calidad aromática, amargor, resabio preferencia y sabor. Este hecho podría ser de gran ayuda debido a que únicamente la medición de este metabolito estaría relacionado con la sección de varios atributos en una misma muestra. Varios de los metabolitos que se detectaron presenta un patrón similar al de este ion.

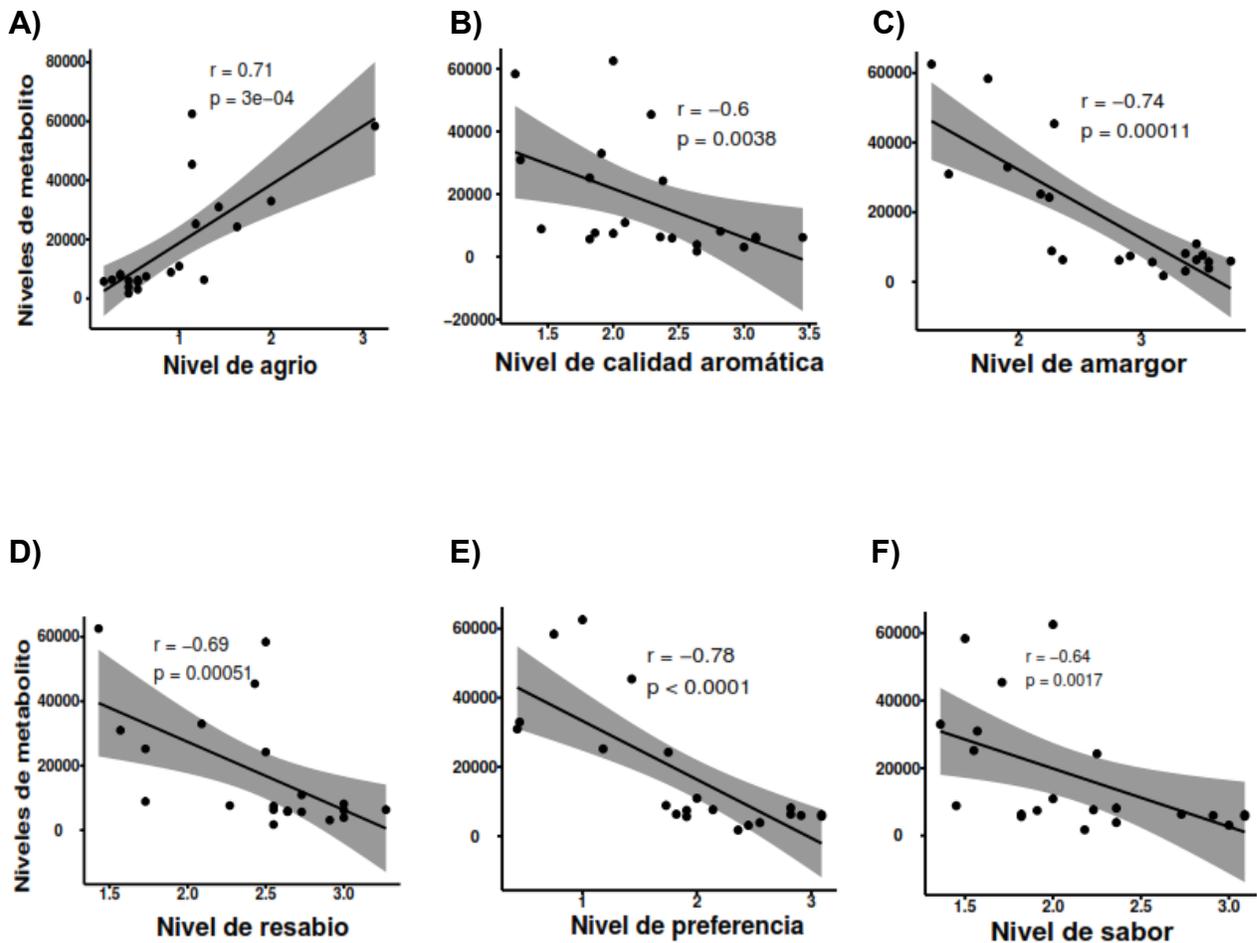


Figura 33. Gráfico de correlación entre m/z 727.9535 y los atributos sensoriales de la bebida. El valor de r corresponde el coeficiente de regresión de Spearman o también llamada rho. En cada una de las gráficas se muestra la correlación pareada entre el ion m/z 727.9535 y A) Agrio, B) calidad aromática, C) amargor, D) resabio, E) preferenciava y E) sabor. Se indica también la significancia estadística de la correlación.

La selección de un grupo de estos metabolitos representa una opción para apoyar el área de la evaluación sensorial que podría estar enfocada a complementar las evaluaciones de los catadores, que hasta ahora es considerada como el estándar de oro para clasificar sensorialmente una la bebida de café.

6.4 Selección de metabolitos marcadores de selección temprana

Hasta este punto se habían determinado los compuestos que son más heredables tanto en frutos (450 metabolitos) como en hojas (116 metabolitos), así también se contaba con los datos de evaluación sensorial de cada una de las muestras. Por lo que el siguiente objetivo fue determinar mediante análisis de correlación si los metabolitos altamente heredables en hojas y frutos se encuentran asociados con atributos sensoriales de la bebida de café.

En total, en frutos se identificaron 233 asociaciones de correlación (Fig. 34) en las que se encontraron involucrados 164 metabolitos con correlaciones ya sea negativas o positivas con al menos un atributo sensorial de la bebida.

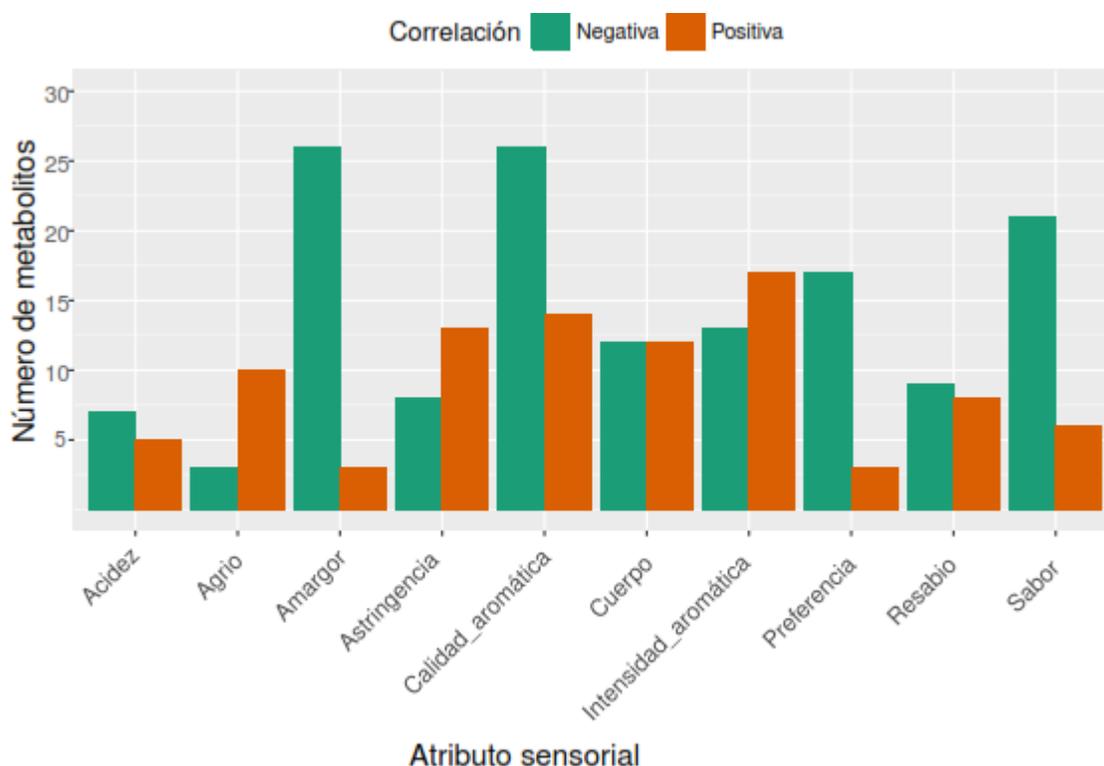


Figura 34. Metabolitos altamente heredables ($H^2 > 0.4$) correlacionados con atributos sensoriales de la bebida. Significancia estadística: $p \leq 0.05$; Coeficiente de correlación de Spearman > 0.5 para: correlación positiva; y valores < -0.5 : correlación negativa.

Para sintetizar la información generada, en la tabla 15 se enlistan dos metabolitos por atributo, uno representando al metabolito que mostró el nivel más alto de correlación positiva y el otro correspondiente al metabolito con el valor más alto de correlación negativa dentro todos los metabolitos explorados.

Tabla 15. Metabolitos detectados en frutos para selección de plantas con características sensoriales deseables.

Atributo	Metabolito asociado	Heredabilidad	CC Spearman	Significancia (valor de p)
Acidez	Indol	0.4638	-0.8167	0.00722
	Cyanidin-3,5-O-diglucósido	0.5085	0.7500	0.01994
Agrio	6-deoxoteasterona	0.4014	-0.8068	0.00860
	Neurosporeno	0.5345	0.7647	0.01638
Amargor	2,3-trans-proanthocyanidin	0.8551	-0.8656	0.00256
	<i>m/z</i> 889.9878853	0.5488	0.7731	0.01455
Astringencia	N-carbamoylputrescina	0.4170	0.8833	0.00159
	<i>m/z</i> 516.7384105	0.5806	-0.8167	0.00722
Calidad aromática	Delfinidin 3-O-glucosil-5-O-caffeoilglucósido	0.9487	-0.8667	0.00250
	D-sedoheptulosa-1,7-bisfosfato	0.4273	0.8000	0.00963
Cuerpo	4-Metil-5-tiazoleetanol fosfato	0.4713	0.8938	0.00116
	Glutathione disulfide	0.4954	-0.8598	0.00296
Intensidad aromática	Molibdopterin-AMP	0.7709	-0.9289	0.00029
	Zeta-caroteno	0.4626	0.9456	0.00012
Preferencia	Isoeugenol	0.5080	-0.8619	0.00281
	All-trans-nonaprenil difosfato	0.9842	0.7280	0.02615
Resabio	Pirofeoforbido-a (PPa)	0.7967	-0.7459	0.02103
	Desulfoglucotropeolin	0.4456	0.8900	0.00130
Sabor	Isoeugenol	0.5080	-0.8908	0.00127
	<i>m/z</i> 337.7305144	0.5326	0.8740	0.00206

El hallazgo de estos metabolitos resulta de gran importancia debido a que cumplen dos características que pueden ayudar a desarrollar una estrategia de selección mucho más robusta que los métodos convencionales de selección y mejoramiento de café. Es decir presentan alta heredabilidad y se encuentran fuertemente ligados con al menos un atributo sensorial de la bebida.

En el caso de hojas, los datos completos de los coeficientes de correlación entre metabolitos altamente heredables vs atributos sensoriales de la bebida se presentan de manera visual en el heatmap de la figura 35.

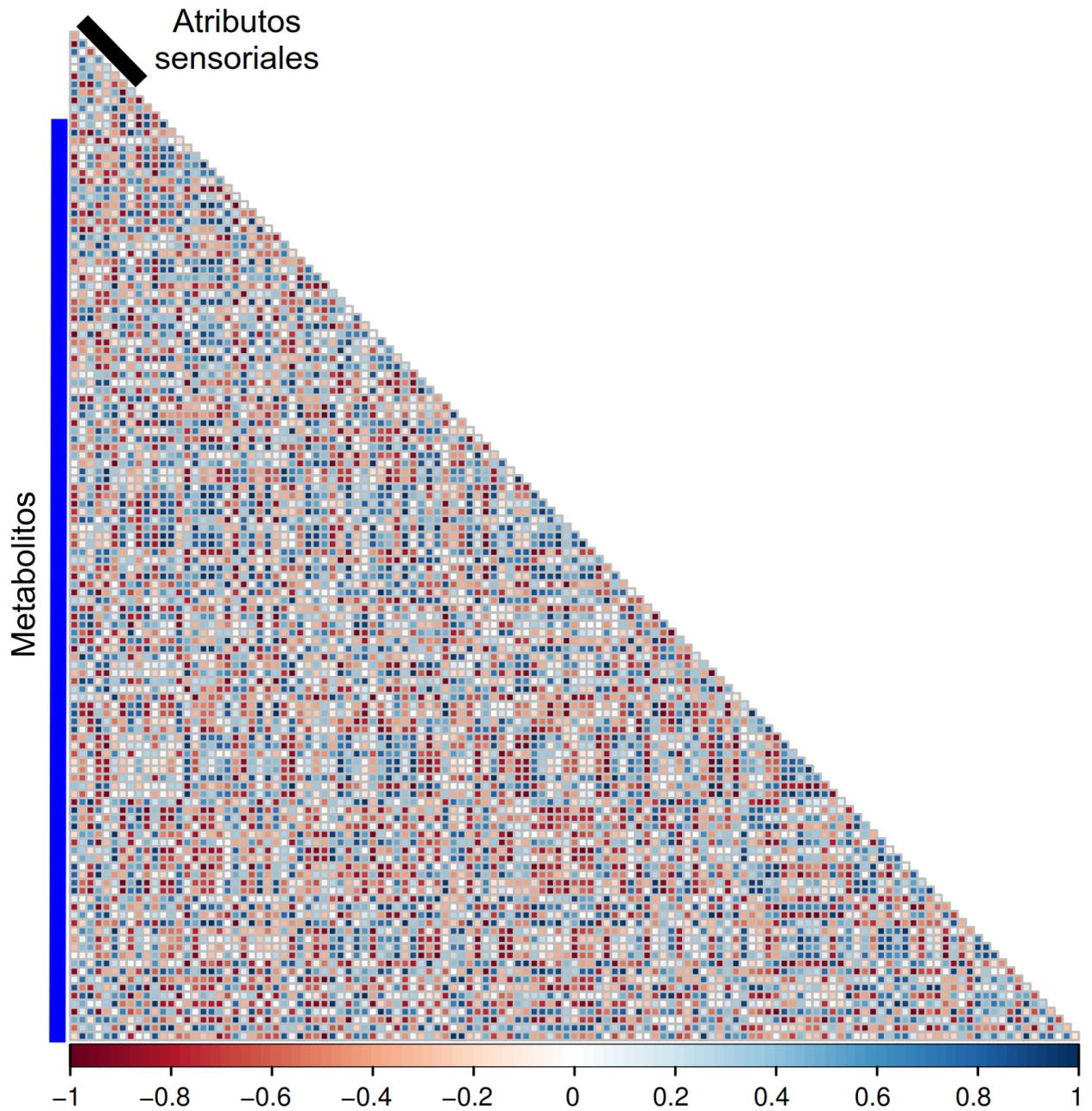


Figura 35. Heat map de correlaciones de metabolitos altamente heredables vs atributos sensoriales de la bebida. Los coeficientes de correlación se calcularon empleando el algoritmo de Spearman en el programa R. Las correlaciones positivas y negativas se muestran en azul y rojo respectivamente, como se muestra en la escala.

En hojas se identificaron 76 asociaciones de correlación con los atributos sensoriales ya sea negativas o positivas que involucraron en total 45 metabolitos. En la tabla 16 se muestran los metabolitos altamente heredables detectados en hojas, que mostraron los niveles más altos de correlación con los atributos sensoriales (Tabla 16).

Tabla 16. Metabolitos detectados en hojas para selección de plantas con características sensoriales deseables.

Atributo	Metabolito asociado	Heredabilidad	CC Spearman	Significancia (valor de p)
Acidez	GA43	1.00	-0.80000001	0.009627923
	Vitamina K1	0.64384592	0.78333336	0.012519868
Agrio	<i>m/z</i> 695.35	0.53506301	-0.77313656	0.014548885
	6-deoxoteasterone	0.64384592	0.77313656	0.014548885
Amargor	Dtdp-L-rhamnose	0.43846674	-0.94120967	0.000153116
	Vitamina K1	0.64384592	-0.85717314	0.003143929
Astringencia	<i>m/z</i> 901.7532	0.68631861	-0.91666669	0.000506619
	10-Nonadecenoic acid	0.61209476	0.66666669	0.049867221
Calidad-aromática	7,8-dihydroneopterin triphosphate	0.90950437	-0.88333333	0.001590501
	zymostenol	0.53959755	0.86666667	0.002495398
Cuerpo	UDP-3-O-(3-hydroxymyristoyl)glucosamine	0.76947057	-0.92786962	0.000309097
	coenzyme A	0.46694948	-0.85976911	0.00295628
Intensidad-aromática	<i>m/z</i> 713.3170	0.9586747	0.80335432	0.009106404
	Caffeine	0.68221764	0.76151294	0.017117127
Preferencia	Vitamina K1	0.64384592	-0.92051017	0.0004311
	propionyl-CoA	0.68719929	0.79498601	0.010444523
Resabio	<i>m/z</i> 695.35	0.53506301	0.73739403	0.023371819
	Pheophytin a5	0.63224843	-0.88995832	0.001304925
Sabor	<i>m/z</i> 695.35	0.53506301	0.79834753	0.009892105
	7,8-dihydroneopterin triphosphate	0.90950437	-0.97482431	8.1338E-06

Para los atributos amargor y cuerpo únicamente se muestran metabolitos con valores de correlación negativa, debido a que no se identificaron metabolitos que estuvieran asociados positivamente. Contrario a lo que ocurrió con el atributo intensidad aromática para el cual no se identificaron metabolitos con correlación negativa.

El hallazgo de estos metabolitos resulta de gran importancia debido a que cumplen dos características que pueden ayudar a desarrollar una estrategia de selección mucho más robusta que los métodos convencionales de selección y mejoramiento de café. Es decir presentan alta heredabilidad y se encuentran fuertemente ligados con al menos un atributo sensorial de la bebida.

7. DISCUSIÓN

Para realizar la discusión detallada de los resultados de este trabajo se dividió en las siguientes secciones: 7.1) Identificación de metabolitos predictivos de estadio de desarrollo, 7.2) Variabilidad de fenotipos metabólicos y Heredabilidad metabólica de familias de *C. canephora* y 7.3) Metabolitos marcadores para selección de plantas con características sensoriales deseables

7.1 Identificación de metabolitos predictivos de estadio de desarrollo

El grado de madurez de los frutos de café es uno de los factores más importantes que determinan la calidad y el sabor de la bebida (Farah, 2012). Sin embargo un hecho que resulta de gran relevancia es la dilucidación de un número restringido de compuestos químicos que podrían ser empleados como discriminantes del nivel de desarrollo y grado de madurez del fruto, es decir su posible uso como biomarcadores.

Debido a lo anterior, en el presente trabajo, se generó un modelo predictivo generado a partir de datos metabolómicos, mediante el cual se logró identificar a 30 metabolitos que en conjunto resultan suficientes para predecir el nivel de desarrollo y madurez de los frutos. Los compuestos que se identificaron como diferenciales comprendieron terpenoides, carotenoides, ácidos grasos, glicerofosfolípidos, fenilpropanoides, antocianinas, carbohidratos, aminoácidos, inositoles y compuestos del metabolismo de clorofila.

Sin embargo, los que de acuerdo al modelo generado presentaron el mayor valor predictivo fueron los compuestos Cianidin-3-O-(6-O-p-coumaroil) de frutos y cianidin-3,5-diglucósido, los cuales se establecieron como de hojas como marcadores predictivos de estadio de desarrollo del fruto de café. Este metabolito presentó un patrón de aumento gradual de sus niveles, alcanzando su máxima abundancia en el fruto rojo. Este comportamiento coincide con con la naturaleza del compuesto identificado, ya que se sabe que en general las antocianinas tienden a acumularse mientras los pigmentos de clorofila desaparecen durante los últimos estadios de maduración del fruto provocando cambios en la coloración de los tejidos de exocarpo y pericarpio (De Castro & Marraccini, 2006).

En el 2013, Setoyama y colaboradores identificaron en granos verdes al triptófano como marcador específico de nivel de madurez, el cual permitió diferenciar entre frutos

maduros e inmaduros en diferentes variedades de *C. arabica*. Este metabolito presenta niveles elevados en frutos inmaduros pero disminuyen de manera dependiente de la maduración del fruto. La diferencia con nuestro estudio es que la discriminación de grado de madurez se llevó a cabo con granos tratados postcosecha, ya que una parte importante de control de calidad de los granos verdes es el nivel de madurez.

Otro hallazgo sumamente interesante en nuestros resultados, fue la identificación de cianidin-3,5-diglucósido como marcador predictivo de estado maduro e inmaduro de frutos a través de un modelo predictivo aplicado a perfiles metabólicos de hojas. Este es el primer estudio en el que se desarrolla un análisis de metabolómica no dirigida para dilucidar si los cambios observados en el fruto durante su desarrollo, también podría verse reflejados en otros tejidos como hojas. Esta cianidina detectada en hojas presentó niveles bajos en los estadios tempranos E1 y E2, y niveles cuatro veces más altos en frutos de estado semimaduros a maduro (Fig. 24 B), por lo que los niveles de este metabolito son un factor discriminante entre el grado de madurez de los frutos de café.

7.2 Variabilidad de fenotipos metabólicos y heredabilidad metabólica de familias de *C. canephora*

El estudio de la herencia genética de caracteres asociados con calidad en taza del café podría optimizar los programas de mejoramiento. Dado que un carácter con alto nivel de heredabilidad representa una ventaja al ser en menor medida afectado por el aspecto ambiental. Este tipo de análisis notoriamente se ha limitado al estudio de compuestos y grupos de estos que han sido asociados con características sensoriales de la bebida, éstos incluyen cafeína, trigonelina, contenido de lípidos y contenido total de ácidos clorogénicos (Mazzafera & Carvalho, 1991; Barre *et al.*, 1998; Montagnon *et al.*, 1998; Ky *et al.*, 1999; Leroy *et al.*, 2011; Akaffou *et al.*, 2012).

En este sentido, nuestro trabajo es pionero en dilucidar la heredabilidad de manera masiva de cientos de metabolitos presentes en hojas y frutos de café. Nuestros resultados muestran que la metodología de UPLC-MS empleada es suficientemente robusta ya que a partir de los perfiles metabólicos de plantas madre de café y su progenie, es posible demostrar el nivel de parentesco que existe entre plantas hija (progenie) que comparten un único parental (Fig. 26). Esto da indicios de que los niveles de algunos metabolitos se

están manteniendo de una generación a otra.

Nuestros resultados mostraron que el 53.77 % de metabolitos detectados en frutos son altamente heredables, mientras que en hojas el 22 % mostraron estar influenciados mayormente por el factor genético. Este grupo de metabolitos identificados podría ser de gran ayuda en futuros proyectos para determinar su asociación con otras características deseables de la planta.

7.3 Metabolitos marcadores para selección de plantas con características sensoriales deseables

El mejoramiento genético de café es tema de suma importancia en el desarrollo de la cafecultura, ya que mediante la aplicación exitosa de estas estrategias se ha permitido la generación de variedades de plantas que poseen características agronómicas y de calidad como productividad, resistencia a enfermedades, adaptabilidad a condiciones climáticas y de suelo, calidad física del grano y características sensoriales de la bebida (Zamarripa Colmenero & Escamilla Prado, 2002; López Morgado *et al.*, 2013). Sin embargo aspectos como la incompatibilidad entre *C. arabica* y *C. robusta*, así como el tiempo que transcurre entre la germinación de la semilla hasta la obtención de la primera cosecha (3.5 - 5 años) crea una desventaja para los métodos tradicionales de mejoramiento. Esto provoca que para la obtención de una variedad mejorada puedan transcurrir hasta 30 años.

Basado en lo anterior, en este trabajo se propuso el desarrollo de una novedosa estrategia asistida por metabolómica, con el objetivo de identificar metabolitos altamente heredables predictivos de características sensoriales de la bebida en estadios tempranos de desarrollo de la planta.

Se logró detectar 164 y 45 metabolitos en frutos y hojas respectivamente que se encuentran asociados negativa o positivamente con al menos un atributo sensorial de la bebida.

En este sentido, recientes estudios en otros cultivos indican que la metabolómica posee potencial para el descubrimiento de metabolitos marcadores predictivos de caracteres fenotípicos (Meyer *et al.*, 2007, Sulpice *et al.*, 2010, Schauer *et al.*, 2006, Schauer *et al.*, 2008, Steinfath *et al.*, 2010).

Una alternativa para asistir la selección de plantas de café con caracteres deseables hasta tiempos recientes ha sido el uso marcadores moleculares o SNPs. Esta tecnología no es tan eficiente cuando los caracteres que son objeto de mejoramiento se ven altamente influenciados por regulación poligénica, nivel de poliploidía de la especie estudiada y efectos epistáticos o por el medio ambiente, que impactan caracteres complejos como sabor, color, tolerancia a estrés, etc.

Por lo que la selección asistida por metabolitos marcadores representa una opción potencial para el mejoramiento de plantas independientemente de la existencia de información genómica, sobre todo en casos en que los marcadores moleculares presentan ciertas limitaciones.

8. CONCLUSIONES

En el presente trabajo se investigó el desarrollo del fruto del café mediante un enfoque de metabolómica no dirigida en la búsqueda de metabolitos marcadores. Así también se desarrolló una estrategia asistida por metabolómica para la selección y mejoramiento de plantas de café con características sensoriales deseables.

- A. Mediante un modelo predictivo generado con perfiles metabolómicos de seis estadios de desarrollo del fruto de café, se identificaron como marcadores específicos de estadio y madurez del fruto a los compuestos cianidin-3-O-(6-O-p-coumaroil) glucósido (en frutos) y cianidin-3,5-diglucósido (en hojas). El empleo de los metabolitos identificados como biomarcadores de estadio de desarrollo de frutos de café, abre la posibilidad de implementar su uso a nivel industrial para la determinación de grado de madurez de granos verdes, un parámetro considerado de gran importancia en el control de calidad de productos derivados del café.
- B. El análisis de heredabilidad de metabolitos en hojas y frutos permitió la identificación de 450 y 116 con heredabilidad mayor a 0.4 en frutos y hojas respectivamente. La variación observada en los niveles de estos cientos de metabolitos en tejidos de *C. canephora* puede ser atribuida predominantemente al factor genético, hecho que se ve reflejado en los niveles altos de heredabilidad que presentaron. Sin embargo es importante recalcar que el fenotipo observado en la plantas refleja la interacción entre el factor ambiental y el factor genético. El uso de estos metabolitos aunado a la caracterización fenotípica, resulta promisorio como apoyo en los programas clásicos de mejoramiento genético del cultivo de café.

Por otra parte, se identificaron 164 metabolitos en frutos y 45 en hojas que podrían ser implementados como marcadores de selección temprana, debido a que además de ser altamente heredables se asociaron con al menos un atributo sensorial de la bebida.

Por lo tanto, la estrategia propuesta en el presente proyecto representa una alternativa para la selección temprana de plantas de café con características sensoriales deseables. Este enfoque permitiría el acortamiento de los tiempos que normalmente toma la selección generacional en cultivos perennes como el café.

Este enfoque podría ser aplicado para otros cultivos en la selección de genotipos

de características agronómicas deseables.

9. PERSPECTIVAS

En futuros trabajos se puede evaluar la respuesta a la selección en progenies de generación avanzada de plantas de *C. Canephora*, con el objetivo de determinar la estabilidad de las características seleccionadas mediante nuestra estrategia.

Sería interesante estudiar si nuestra estrategia desarrollada para seleccionar características sensoriales de la bebida, podría también ser aplicada para mejorar características como rendimiento y resistencia a enfermedades.

Así también, se puede investigar la posible aplicación de la estrategia aquí presentada en plantas de *C. arabica*, para determinar su respuesta, considerando las diferencias que existen entre especies como su genoma tetraploide y su característica de autocompatibilidad.

10. BIBLIOGRAFÍA

- Aerts, R., Berecha, G., Gijbels, P., Hundera, K., Glabeke, S., Vandepitte, K., *et al.* 2013. Genetic variation and risks of introgression in the wild *Coffea arabica* gene pool in south-western Ethiopian montane rainforests. *Evolutionary Applications* **6**: 243–252.
- Aggarwal, R.K., Hendre, P.S., Varshney, R.K., Bhat, P.R., Krishnakumar, V. & Singh, L. 2007. Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. *Theor. Appl. Genet.* **114**: 359–372.
- Akaffou, D.S., Hamon, P., Doulbeau, S., Keli, J., Legnate, H., Campa, C., *et al.* 2012. Inheritance and relationship between key agronomic and quality traits in an interspecific cross between *Coffea pseudozanguebariae* Bridson and *C. canephora* Pierre. *Tree Genetics & Genomes* **8**: 1149–1162.
- Akiyama, M., Murakami, K., Hirano, Y., Ikeda, M., Iwatsuki, K., Wada, A., *et al.* 2008. Characterization of headspace aroma compounds of freshly brewed arabica coffees and studies on a characteristic aroma compound of Ethiopian coffee. *Journal of Food Science* **73**: C335-346.
- Akiyama, M., Murakami, K., Ohtani, N., Iwatsuki, K., Sotoyama, K., Wada, A., *et al.* 2003. Analysis of volatile compounds released during the grinding of roasted coffee beans using solid-phase microextraction. *J. Agric. Food Chem.* **51**: 1961–1969.
- Al-Murish, T.M., Elshafei, A.A., Al-Doss, A.A. & Barakat, M.N. 2013. Genetic diversity of coffee (*Coffea arabica* L.) in Yemen via SRAP, TRAP and SSR markers. *Journal of Food, Agriculture & Environment* **11**: 411–416.
- Alvarado, G., Posada, H.E. & Cortina, H.A. 2005. Castillo: nueva variedad de café con resistencia a la roya.
- Amorim, A.C.L., Hovell, A.M.C., Pinto, A.C., Eberlin, M.N., Arruda, N.P., Pereira, E.J., *et al.* 2009. Green and roasted arabica coffees differentiated by ripeness, process and cup quality via electrospray ionization mass spectrometry fingerprinting. *Journal of the Brazilian Chemical Society* **20**: 313–321.
- Arya, M. & Rao, L.J.M. 2007. An impression of coffee carbohydrates. *Critical Reviews in Food Science and Nutrition* **47**: 51–67.
- Ashihara, H., Sano, H. & Crozier, A. 2008. Caffeine and related purine alkaloids: Biosynthesis, catabolism, function and genetic engineering. *Phytochemistry* **69**: 841–856.
- Ausubel, F.M. 2010. Metabolomics. In: *Current Protocols in Molecular Biology* (F. M. Ausubel *et al.*, eds). John Wiley & Sons, Inc., Hoboken, NJ, USA.

- Baggenstoss, J., Poisson, L., Kaegi, R., Perren, R. & Escher, F. 2008. Coffee Roasting and Aroma Formation: Application of Different Time–Temperature Conditions. *Journal of Agricultural and Food Chemistry* **56**: 5836–5846.
- Barre, P., Akaffou, S., Louarn, J., Charrier, A., Hamon, S. & Noiro, M. 1998. Inheritance of caffeine and heteroside contents in an interspecific cross between a cultivated coffee species *Coffea liberica* var *dewevrei* and a wild species caffeine-free *C. pseudozanguebariae*. *Theoretical and Applied Genetics* **96**: 306–311.
- Bekedam, E.K., Loots, M.J., Schols, H.A., Van Boekel, M.A.J.S. & Smit, G. 2008. Roasting Effects on Formation Mechanisms of Coffee Brew Melanoidins. *Journal of Agricultural and Food Chemistry* **56**: 7138–7145.
- Belitz, H.D., Grosh, W. & Schieberle, P. 2009. *Food Chemistry*, 4th ed. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Bertrand, B., Boulanger, R., Dussert, S., Ribeyre, F., Berthiot, L., Descroix, F., *et al.* 2012. Climatic factors directly impact the volatile organic compound fingerprint in green Arabica coffee bean as well as coffee beverage quality. *Food chemistry* **135**: 2575–2583.
- Bertrand, B., Etienne, H., Cilas, C., Charrier, A. & Baradat, P. 2005. *Coffea arabica* hybrid performance for yield, fertility and bean weight. *Euphytica* **141**: 255–262.
- Bertrand, B., Guyot, B., Anthony, F. & Lashermes, P. 2003. Impact of the *Coffea canephora* gene introgression on beverage quality of *C. arabica*. *Theoretical and applied genetics* **107**: 387–394.
- Bertrand, B., Vaast, P., Alpizar, E., Etienne, H., Davrieux, F. & Charmetant, P. 2006. Comparison of bean biochemical composition and beverage quality of Arabica hybrids involving Sudanese-Ethiopian origins with traditional varieties at various elevations in Central America. *Tree physiology* **26**: 1239–1248.
- Bhumiratana, N., Adhikari, K. & Chambers IV, E. 2011. Evolution of sensory aroma attributes from coffee beans to brewed coffee. *LWT - Food Science and Technology* **44**: 2185–2192.
- Blank, I., Sen, A. & Grosch, W. 1992. Potent odorants of the roasted powder and brew of Arabica coffee. *Z Lebensm Unters Forsch* **195**: 239–245.
- Bosselmann, A.S., Dons, K., Oberthur, T., Olsen, C.S., Ræbild, A. & Usma, H. 2009. The influence of shade trees on coffee quality in small holder coffee agroforestry systems in Southern Colombia. *Agriculture, Ecosystems & Environment* **129**: 253–260.
- Brown, J. & Caligari, P.D.S. 2008. *An Introduction to Plant Breeding*. Blackwell Publishing Ltd, Oxford OX4 2DQ, UK.
- Buffo, R.A. & Cardelli-Freire, C. 2004. Coffee flavour: an overview. *Flavour Frag J* **19**: 99–104.

- Cabos, R.Y.M., Sipes, B.S., Nagai, C., Serracin, M. & Schmitt, D.P. 2010. Evaluation of coffee genotypes for root-knot nematode resistance. *Nematropica* **40**: 191–202.
- Cação, S.M.B., Silva, N.V., Domingues, D.S., Vieira, L.G.E., Diniz, L.E.C., Vinecky, F., *et al.* 2013. Construction and characterization of a BAC library from the *Coffea arabica* genotype Timor Hybrid CIFC 832/2. *Genetica* **141**: 217–226.
- CAFÉS DE MÉXICO. 2016. Plan Integral de Atención al Café PIAC 2016.
- Cantergiani, E., Brevard, H., Krebs, Y., Feria-Morales, A., Amadó, R. & Yeretian, C. 2001. Characterisation of the aroma of green Mexican coffee and identification of mouldy/earthy defect. *Eur Food Res Technol* **212**: 648–657.
- Carvajal Herrera, J.J., Aristizábal Torres, I.D., Oliveros Tascón, C.E. & Mejía Montoya, J.W. 2011. Colorimetría del Fruto de Café (*Coffea arabica* L.) Durante su Desarrollo y Maduración. *Revista Facultad Nacional de Agronomía - Medellín* **64**: 6229–6240.
- Charrier, A. & Eskes, A.B. 2008. Botany and Genetics of Coffee. In: *Coffee: Growing, Processing, Sustainable Production*, pp. 25–56. Wiley-VCH Verlag GmbH & Co. KGaA, Switzerland.
- Clarke, R. & Vitzthum, O.G. 2008. *Coffee: Recent Developments*. John Wiley & Sons.
- Clifford, M.N. 2000. Chlorogenic acids and other cinnamates – nature, occurrence, dietary burden, absorption and metabolism. *J. Sci. Food Agric.* **80**: 1033–1043.
- Costa Freitas, A.M. & Mosca, A.I. 1999. Coffee geographic origin — an aid to coffee differentiation. *Food Research International* **32**: 565–573.
- Coulibaly, I., Revol, B., Noirot, M., Poncet, V., Lorieux, M., Carasco-Lacombe, C., *et al.* 2003. AFLP and SSR polymorphism in a *Coffea* interspecific backcross progeny [(*C. heterocalyx* x *C. canephora*) x *C. canephora*]. *Theor. Appl. Genet.* **107**: 1148–1155.
- Cubry, P., Bellis, F.D., Pot, D., Musoli, P. & Leroy, T. 2012. Global analysis of *Coffea canephora*. *Genet Resour Crop Evol* **60**: 483–501.
- Cubry, P., de Bellis, F., Avia, K., Bouchet, S., Pot, D., Dufour, M., *et al.* 2013. An initial assessment of linkage disequilibrium (LD) in coffee trees: LD patterns in groups of *Coffea canephora* Pierre using microsatellite analysis. *BMC Genomics* **14**: 10.
- Czerny, M. & Grosch, W. 2000. Potent Odorants of Raw Arabica Coffee. Their Changes during Roasting. *Journal of Agricultural and Food Chemistry* **48**: 868–872.
- DaMatta, F.M., Ronchi, C.P., Maestri, M. & Barros, R.S. 2007. Ecophysiology of coffee growth and production. *Brazilian Journal of Plant Physiology* **19**: 485–510.
- Davis, A.P., Chester, M., Maurin, O. & Fay, M.F. 2007. Searching for the relatives of *Coffea* (Rubiaceae, Ixoroideae): the circumscription and phylogeny of *Coffeae* based on plastid sequence data and morphology. *Am. J. Bot.* **94**: 313–329.

- Davis, A.P., Gole, T.W., Baena, S. & Moat, J. 2012. The Impact of Climate Change on Indigenous Arabica Coffee (*Coffea arabica*): Predicting Future Trends and Identifying Priorities. *PLoS ONE* **7**: e47981.
- Davis, A.P., Tosh, J., Ruch, N. & Fay, M.F. 2011. Growing coffee: *Psilanthus* (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea*. *Botanical Journal of the Linnean Society* **167**: 357–377.
- de Azevedo, A.B.A., Kieckbush, T.G., Tashima, A.K., Mohamed, R.S., Mazzafera, P. & Melo, S.A.B.V. de. 2008. Extraction of green coffee oil using supercritical carbon dioxide. *The Journal of Supercritical Fluids* **44**: 186–192.
- De Castro, R.D. & Marraccini, P. 2006. Cytology, biochemistry and molecular changes during coffee fruit development. *Brazilian Journal of Plant Physiology* **18**: 175–199.
- De Kochko, A., Crouzillat, D., Rigoreau, M., Lepelley, M., Bellanger, L., Merot l'Anthoene, V., *et al.* 2015. Dihaploid #*Coffea arabica*# genome sequencing and assembly [W180].
- Denoeud, F., Carretero-Paulet, L., Dereeper, A., Droc, G., Guyot, R., Pietrella, M., *et al.* 2014. The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science* **345**: 1181–1184.
- Dereeper, A., Guyot, R., Tranchant-Dubreuil, C., Anthony, F., Argout, X., de Bellis, F., *et al.* 2013. BAC-end sequences analysis provides first insights into coffee (*Coffea canephora P.*) genome composition and evolution. *Plant Mol. Biol.* **83**: 177–189.
- Dessalegn, Y., Herselman, L. & Labuschagne, M.T. 2008. AFLP analysis among Ethiopian arabica coffee genotypes. *African Journal of Biotechnology* **7**.
- Dirección de Investigación y Evaluación Económica y Sectorial. 2015. *Panorama Agroalimentario: Café 2015*. México.
- Eira, M.T., Silva, E.A., De Castro, R.D., Dussert, S., Walters, C., Bewley, J.D., *et al.* 2006. Coffee seed physiology. *Brazilian Journal of Plant Physiology* **18**: 149–163.
- Eskes, A.B. & Leroy, T. 2008. Coffee Selection and Breeding. In: *Coffee: Growing, Processing, Sustainable Production*, pp. 57–86. Wiley-VCH Verlag GmbH & Co. KGaA, Switzerland.
- Esquivel, P. & Jiménez, V.M. 2012. Functional properties of coffee and coffee by-products. *Food Research International* **46**: 488–495.
- Falconer, D.S. & Mackay, T.F.C. 1996. *Introduction to quantitative genetics*, Fourth Edition. Longman, Essex, England.
- Farah, A. 2012. 2 Coffee Constituents. *Coffee: Emerging health effects and disease prevention* **59**.

- Farah, A. 2009. 15 - Coffee as a speciality and functional beverage A2 - Paquin, Paul. In: *Functional and Speciality Beverage Technology*, pp. 370–395. Woodhead Publishing.
- Farah, A. & Donangelo, C.M. 2006. Phenolic compounds in coffee. *Brazilian Journal of Plant Physiology* **18**: 23–36.
- Farah, A., Monteiro, M.C., Calado, V., Franca, A.S. & Trugo, L.C. 2006. Correlation between cup quality and chemical attributes of Brazilian coffee. *Food Chemistry* **98**: 373–380.
- Fernandez, D., Santos, P., Agostini, C., Bon, M.-C., Petitot, A.-S., C Silva, M., *et al.* 2004. Coffee (*Coffea arabica* L.) genes early expressed during infection by the rust fungus (*Hemileia vastatrix*). *Mol. Plant Pathol.* **5**: 527–536.
- Ferrão, L.F.V., Caixeta, E.T., Souza, F. de F., Zambolim, E.M., Cruz, C.D., Zambolim, L., *et al.* 2012. Comparative study of different molecular markers for classifying and establishing genetic relationships in *Coffea canephora*. *Plant Syst Evol* **299**: 225–238.
- Filho, O.G., Silvarolla, M.B. & Eskes, A.B. 1999. Expression and mode of inheritance of resistance in coffee to leaf miner *Perileucoptera coffeella*. *Euphytica* **105**: 7–15.
- Flament, I. & Bessière-Thomas, Y. 2002. *Coffee flavor chemistry*. Wiley, Chichester ; New York.
- Fujioka, K. & Shibamoto, T. 2008. Chlorogenic acid and caffeine contents in various commercial brewed coffees. *Food Chemistry* **106**: 217–221.
- GAIN Report. 2016. *Tight Coffee Situation In México*. USA.
- Gartner, G.A.L., McCouch, S.R. & Moncada, M.D.P. 2013. A genetic map of an interspecific diploid pseudo testcross population of coffee. *Euphytica* **192**: 305–323.
- Gaspari-Pezzopane, C. de, Bonturi, N., Guerreiro Filho, O., Favarin, J.L. & Maluf, M.P. 2012. Gene expression profile during coffee fruit development and identification of candidate markers for phenological stages. *Pesquisa Agropecuária Brasileira* **47**: 972–982.
- Geleta, M., Herrera, I., Monzón, A. & Bryngelsson, T. 2012. Genetic Diversity of Arabica Coffee (*Coffea arabica* L.) in Nicaragua as Estimated by Simple Sequence Repeat Markers. *The Scientific World Journal* **2012**: e939820.
- Geromel, C., Ferreira, L.P., Guerreiro, S.M.C., Cavalari, A.A., Pot, D., Leroy, T., *et al.* 2006. Biochemical and genomic analysis of sucrose metabolism during coffee (*Coffea arabica*) fruit development. *J. Exp. Bot.* **57**: 3243–3258.
- Gichimu, B.M., Gicheru, E.K., Mamati, G.E. & Nyende, A.B. 2014. Variation and association of cup quality attributes and resistance to coffee berry disease in *Coffea*

- arabica L. Composite cultivar, Ruiru 11. *African Journal of Horticultural Science* **7**.
- Gielessen, R. & Graafland, J.J. 2009. Concepts of price fairness: Empirical research into the Dutch coffee market.
- Gimase, J.M., W., M.T., D., T.K., E., K.G. & C., W.K. 2014. Beverage quality and biochemical attributes of arabusta coffee (*C. arabica* L. x *C. canephora* Pierre) and their parental genotypes. *African Journal of Food Science* **8**: 456–464.
- Ginz, M., Balzer, H.H., Bradbury, A.G. & Maier, H.G. 2000. Formation of aliphatic acids by carbohydrate degradation during roasting of coffee. *European Food Research and Technology* **211**: 404–410.
- Gonzalez-Rios, O., Suarez-Quiroz, M.L., Boulanger, R., Barel, M., Guyot, B., Guiraud, J.-P., *et al.* 2007. Impact of “ecological” post-harvest processing on the volatile fraction of coffee beans: I. Green coffee. *J Food Comp Anal* **20**: 289–296.
- Herrmann, A. & Schauer, N. 2013. Metabolomics-Assisted Plant Breeding. In: *The Handbook of Plant Metabolomics* (W. Weckwerth & G. Kahl, eds), pp. 245–254. Wiley-VCH Verlag GmbH & Co. KGaA.
- Hollywood, K., Brison, D.R. & Goodacre, R. 2006. Metabolomics: Current technologies and future trends. *PROTEOMICS* **6**: 4716–4723.
- Holscher, W. & Steinhart, H. 1995. Aroma compounds in green coffee. In: *Developments in Food Science* (G. Charalambous, ed), pp. 785–803. Elsevier.
- J. Singh, R. 2012. *Genetic Resources, Chromosome Engineering, and Crop Improvement: Medicinal Plants*. CRC Press.
- Joet, T., Laffargue, A., Salmona, J., Doulebeau, S., Descroix, F., Bertrand, B., *et al.* 2009. Metabolic pathways in tropical dicotyledonous albuminous seeds: *Coffea arabica* as a case study. *New Phytol* **182**: 146–162.
- Kochko, A.D., Akaffou, S., Andrade, A.C., Campa, C., Crouzillat, D., Guyot, R., *et al.* 2010a. Advances in *Coffea* Genomics. In: *Advances in Botanical Research*, pp. 23–63. Elsevier.
- Kochko, A.D., Akaffou, S., Andrade, A.C., Campa, C., Crouzillat, D., Guyot, R., *et al.* 2010b. Advances in *Coffea* Genomics. In: *Advances in Botanical Research* (J.-C. K. and M. Delseny, ed), pp. 23–63. Academic Press.
- Koshiro, Y., Jackson, M.C., Katahira, R., Wang, M.-L., Nagai, C. & Ashihara, H. 2007. Biosynthesis of chlorogenic acids in growing and ripening fruits of *Coffea arabica* and *Coffea canephora* plants. *Z. Naturforsch., C, J. Biosci.* **62**: 731–742.
- Koshiro, Y., Jackson, M.C., Nagai, C. & Ashihara, H. 2015. Changes in the content of sugars and organic acids during ripening of *coffea arabica* and *coffea canephora* fruits. *European Chemical Bulletin* **4**: 378–383.

- Koshiro, Y., Zheng, X.-Q., Wang, M.-L., Nagai, C. & Ashihara, H. 2006. Changes in content and biosynthetic activity of caffeine and trigonelline during growth and ripening of *Coffea arabica* and *Coffea canephora* fruits. *Plant Sci* **171**: 242–250.
- Ky, C.-L., Doulebeau, B., Guyot, B., Akaffou, S., Charrier, A., Hamon, S., *et al.* 2000. Inheritance of coffee bean sucrose content in the interspecific cross *Coffea pseudozanguebariae* × *Coffea liberica* “dewevrei.” *Plant Breeding* **119**: 165–168.
- Ky, C.-L., Guyot, B., Louarn, J., Hamon, S. & Noirot, M. 2001. Trigonelline inheritance in the interspecific *Coffea pseudozanguebariae* × *C. liberica* var. *dewevrei* cross. *Theor Appl Genet* **102**: 630–634.
- Ky, C.-L., Louarn, J., Guyot, B., Charrier, A., Hamon, S. & Noirot, M. 1999. Relations between and inheritance of chlorogenic acid contents in an interspecific cross between *Coffea pseudozanguebariae* and *Coffea liberica* var “dewevrei.” *Theor Appl Genet* **98**: 628–637.
- Landey, R.B., Cenci, A., Georget, F., Bertrand, B., Camayo, G., Dechamp, E., *et al.* 2013. High Genetic and Epigenetic Stability in *Coffea arabica* Plants Derived from Embryogenic Suspensions and Secondary Embryogenesis as Revealed by AFLP, MSAP and the Phenotypic Variation Rate. *PLOS ONE* **8**: e56372.
- Lashermes, P., Combes, M.-C., Hueber, Y., Severac, D. & Dereeper, A. 2014. Genome rearrangements derived from homoeologous recombination following allopolyploidy speciation in coffee. *Plant J* **78**: 674–685.
- Lashermes, P., Combes, M.C., Prakash, N.S., Trouslot, P., Lorieux, M. & Charrier, A. 2001. Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meioses. *Genome* **44**: 589–596.
- Lashermes, P., Combes, M.-C., Robert, J., Trouslot, P., D’Hont, A., Anthony, F., *et al.* 1999. Molecular characterisation and origin of the *Coffea arabica* L. genome. *Mol Gen Genet* **261**: 259–266.
- Lepelley, M., Cheminade, G., Tremillon, N., Simkin, A., Caillet, V. & McCarthy, J. 2007. Chlorogenic acid synthesis in coffee: An analysis of CGA content and real-time RT-PCR expression of HCT, HQT, C3H1, and CCoAOMT1 genes during grain development in *C. canephora*. *Plant Science* **172**: 978–996.
- Leroy, T., De Bellis, F., Legnate, H., Kananura, E., Gonzales, G., Pereira, L.F., *et al.* 2011. Improving the quality of African robustas: QTLs for yield- and quality-related traits in *Coffea canephora*. *Tree Genetics & Genomes* **7**: 781–798.
- Leroy, T., De Bellis, F., Legnate, H., Musoli, P., Kalonji, A., Llor Solórzano, R.G., *et al.* 2014. Developing core collections to optimize the management and the exploitation of diversity of the coffee *Coffea canephora*. *Genetica* **142**: 185–199.
- Leroy, T., Marraccini, P., Dufour, M., Montagnon, C., Lashermes, P., Sabau, X., *et al.* 2005. Construction and characterization of a *Coffea canephora* BAC library to study the organization of sucrose biosynthesis genes. *Theoretical and Applied Genetics* **111**:

1032–1041.

- Leroy, T., Ribeyre, F., Bertrand, B., Charmetant, P., Dufour, M., Montagnon, C., *et al.* 2006. Genetics of coffee quality. *Brazilian Journal of Plant Physiology* **18**: 229–242.
- López Morgado, R., Díaz Padilla, G. & Zamarripa Colmenero, A. 2013. *El sistema producto café en México: problemática y tecnología de producción*, Primera. Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Centro de Investigación Regional Golfo, Centro Campo Experimental Cotaxtla, Medellín de Bravo, Veracruz, México.
- López-Gartner, G., Cortina, H., McCouch, S.R. & Moncada, M.D.P. 2009. Analysis of genetic structure in a sample of coffee (*Coffea arabica* L.) using fluorescent SSR markers. *Tree Genetics & Genomes* **5**: 435–446.
- Maldonado, M.M., Lomelí, A.P., Castellanos, J.S., Pérez, J.E.R. & Aguilar, R.M. 2002. Varianza aditiva, heredabilidad y correlaciones en la variedad M1-Fitotecnia de tomate de cáscara (*Physalis ixocarpa* Brot). *Rev Fitotec Mex* **25**: 231–237.
- Maluf, M.P., Silvestrini, M., Ruggiero, L.M. de C., Guerreiro Filho, O. & Colombo, C.A. 2005. Genetic diversity of cultivated *Coffea arabica* inbred lines assessed by RAPD, AFLP and SSR marker systems. *Scientia Agricola* **62**: 366–373.
- Marraccini, P., Rogers, W.J., Allard, C., André, M.L., Caillet, V., Lacoste, N., *et al.* 2001. Molecular and biochemical characterization of endo-beta-mannanases from germinating coffee (*Coffea arabica*) grains. *Planta* **213**: 296–308.
- Masumbuko, L.I., Bryngelsson, T., Mneney, E.E. & Salomon, B. 2003. Genetic diversity in Tanzanian Arabica coffee using random amplified polymorphic DNA (RAPD) markers. *Hereditas* **139**: 56–63.
- Mathieu, F., Malosse, C. & Frérot, B. 1998. Identification of the Volatile Components Released by Fresh Coffee Berries at Different Stages of Ripeness. *J. Agric. Food Chem.* **46**: 1106–1110.
- Mayer, F., Czerny, M. & Grosch, W. 2000. Sensory study of the character impact aroma compounds of a coffee beverage. *Eur Food Res Technol* **211**: 272–276.
- Mayer, F. & Grosch, W. 2001. Aroma simulation on the basis of the odourant composition of roasted coffee headspace. *Flavour and Fragrance Journal* **16**: 180–190.
- Mazzafera, P., Baumann, T.W., Shimizu, M.M. & Silvarolla, M.B. 2009. Decaf and the Steeplechase Towards Decaffito—the Coffee from Caffeine-Free Arabica Plants. *Tropical Plant Biol.* **2**: 63–76.
- Mazzafera, P. & Carvalho, A. 1991. Breeding for low seed caffeine content of coffee (*Coffea* L.) by interspecific hybridization. *Euphytica* **59**: 55–60.
- Meletis, C.D. 2006. Coffee—Functional Food and Medicinal Herb. *Alternative and Complementary Therapies* **12**: 7–13.

- Mérot-L'Anthoëne, V., Mangin, B., Lefebvre-Pautigny, F., Jasson, S., Rigoreau, M., Husson, J., *et al.* 2014a. Comparison of three QTL detection models on biochemical, sensory, and yield characters in *Coffea canephora*. *Tree Genetics & Genomes* **10**: 1541–1553.
- Mérot-L'Anthoëne, V., Mangin, B., Lefebvre-Pautigny, F., Jasson, S., Rigoreau, M., Husson, J., *et al.* 2014b. Comparison of three QTL detection models on biochemical, sensory, and yield characters in *Coffea canephora*. *Tree Genetics & Genomes* **10**: 1541–1553.
- Meyer, R.C., Steinfath, M., Lisek, J., Becher, M., Witucka-Wall, H., Törjék, O., *et al.* 2007. The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* **104**: 4759–4764.
- Minamisawa, M., Yoshida, S. & Takai, N. 2004. Determination of biologically active substances in roasted coffees using a diode-array HPLC system. *Anal Sci* **20**: 325–328.
- Mishra, M.K., Sandhyarani, N., Suresh, N., Kumar, S.S., Soumya, P.R., Yashodha, M.H., *et al.* 2012. Genetic Diversity Among Indian Coffee Cultivars Determined via Molecular Markers. *Journal of Crop Improvement* **26**: 727–750.
- Moncada, M.D.P., Tovar, E., Montoya, J.C., González, A., Spindel, J. & McCouch, S. 2015. A genetic linkage map of coffee (*Coffea arabica* L.) and QTL for yield, plant height, and bean size. *Tree Genetics & Genomes* **12**: 5.
- Mondego, J.M., Vidal, R.O., Carazzolle, M.F., Tokuda, E.K., Parizzi, L.P., Costa, G.G., *et al.* 2011. An EST-based analysis identifies new genes and reveals distinctive gene expression features of *Coffea arabica* and *Coffea canephora*. *BMC Plant Biol.* **11**: 30.
- Montagnon, C., Guyot, B., Cilas, C. & Leroy, T. 1998. Genetic parameters of several biochemical compounds from green coffee, *Coffea canephora*. *Plant Breeding* **117**: 576–578.
- Montagnon, C., Marraccini, P. & Bertrand, B. 2012. Breeding for coffee quality. *Specialty Coffee: Managing Quality. International Plant Nutrition Institute, Southeast Asia Program* 93–122.
- Morales, F.J. & Jiménez-Pérez, S. 2004. Peroxyl radical scavenging activity of melanoidins in aqueous systems. *Eur Food Res Technol* **218**: 515–520.
- Motta, L.B., Soares, T.C.B., Ferrão, M.A.G., Caixeta, E.T., Lorenzoni, R.M., Neto, S., *et al.* 2014. Molecular characterization of arabica and Conilon coffee plants genotypes by SSR and ISSR markers. *Brazilian Archives of Biology and Technology* **57**: 728–735.
- Noir, S., Patheyron, S., Combes, M.-C., Lashermes, P. & Chalhoub, B. 2004. Construction and characterisation of a BAC library for genome analysis of the allotetraploid coffee species (*Coffea arabica* L.). *Theor. Appl. Genet.* **109**: 225–230.

- Noirot, M., Charrier, A., Stoffelen, P. & Anthony, F. 2015. Reproductive isolation, gene flow and speciation in the former *Coffea* subgenus: a review. *Trees* **30**: 597–608.
- Oestreich-Janzen, S. 2010. Chemistry of Coffee in: *Comprehensive Natural Products II: Chemistry and Biology: 10 Volume Set*. In: (L. Mander & H.-W. Liu, eds). Elsevier Science.
- Oliveira, L.S., Franca, A.S., Mendonça, J.C.F. & Barros-Júnior, M.C. 2006. Proximate composition and fatty acids profile of green and roasted defective coffee beans. *LWT - Food Science and Technology* **39**: 235–239.
- Organización internacional del café. 2016. Informe del mercado de Café 2016.
- Paillard, M., Lashermes, P. & Pétiard, V. 1996. Construction of a molecular linkage map in coffee. *Theor. Appl. Genet.* **93**: 41–47.
- P Tshilenge, K.K.N. 2009. Genetic variation in *Coffea canephora* L.(Var. Robusta) accessions from the founder gene pool evaluated with ISSR and RAPD. *African Journal of Biotechnology* **8**: 380–390.
- Pearl, H.M., Nagai, C., Moore, P.H., Steiger, D.L., Osgood, R.V. & Ming, R. 2004. Construction of a genetic map for arabica coffee. *Theor. Appl. Genet.* **108**: 829–835.
- Philippe, L., Benoît, B. & Hervé, E. 2009. Breeding coffee (*Coffea arabica*) for sustainable production. In: *Breeding Plantation Tree Crops: Tropical Species*, pp. 525–543. Springer.
- Pluskal, T., Castillo, S., Villar-Briones, A. & Orešič, M. 2010. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11**: 395.
- Poltronieri, Y., Martinez, H.E.P. & Cecon, P.R. 2011. Effect of zinc and its form of supply on production and quality of coffee beans. *J. Sci. Food Agric.* **91**: 2431–2436.
- Pot, D., Bouchet, S., Marraccini, P., De Bellis, F., Cubry, P., Jourdan, I., *et al.* 2007. Nucleotide diversity of genes involved in sucrose metabolism. Towards the identification of candidates genes controlling sucrose variability in #*Coffea*# sp.
- Priolli, R.H.G., Mazzafera, P., Siqueira, W.J., Möller, M., Zucchi, M.I., Ramos, L.C.S., *et al.* 2008a. Caffeine inheritance in interspecific hybrids of *Coffea arabica* x *Coffea canephora* (Gentianales, Rubiaceae). *Genetics and Molecular Biology* **31**: 498–504.
- Priolli, R.H.G., Ramos, L.C.S., Pot, D., Moller, M., Gallo, P.B. & *et al.* 2008b. Construction of a genetic map based on interspecific F2 population between *Coffea arabica* and *Coffea canephora* and its usefulness for quality related traits : [B213].
- Privat, I., Bardil, A., Gomez, A.B., Severac, D., Dantec, C., Fuentes, I., *et al.* 2011. The “PUCE CAFE” Project: the First 15K Coffee Microarray, a New Tool for Discovering

- Candidate Genes correlated to Agronomic and Quality Traits. *BMC Genomics* **12**: 5.
- Privat, I., Foucrier, S., Prins, A., Epalle, T., Eychenne, M., Kandalaf, L., *et al.* 2008. Differential regulation of grain sucrose accumulation and metabolism in *Coffea arabica* (Arabica) and *Coffea canephora* (Robusta) revealed through gene expression and enzyme activity analysis. *New Phytol* **178**: 781–797.
- Puerta Q., G.I. 1998. Calidad en taza de las variedades de *Coffea arabica* L cultivadas en Colombia.
- Redgwell, R. & Fischer, M. 2006. Coffee carbohydrates. *Brazilian Journal of Plant Physiology* **18**: 165–174.
- Redgwell, R.J., Curti, D., Rogers, J., Nicolas, P. & Fischer, M. 2003. Changes to the galactose/mannose ratio in galactomannans during coffee bean (*Coffea arabica* L.) development: implications for in vivo modification of galactomannan synthesis. *Planta* **217**: 316–326.
- Resende, M., Caixeta, E.T., Ruas Alkimin, E., Vieira Sousa, T., Chamala, S. & G neves, L. 2016. High-Throughput Targeted Genotyping of *Coffea arabica* and *Coffea canephora* Using Next Generation Sequencing. *Plant and Animal Genome*.
- Ribeiro, J.S., Augusto, F., Salva, T.J.G. & Ferreira, M.M.C. 2012. Prediction models for Arabica coffee beverage quality based on aroma analyses and chemometrics. *Talanta* **101**: 253–260.
- Ribeiro, J.S., Augusto, F., Salva, T.J.G., Thomaziello, R.A. & Ferreira, M.M.C. 2009. Prediction of sensory properties of Brazilian Arabica roasted coffees by headspace solid phase microextraction-gas chromatography and partial least squares. *Anal. Chim. Acta* **634**: 172–179.
- Risticvic, S., Carasek, E. & Pawliszyn, J. 2008. Headspace solid-phase microextraction-gas chromatographic-time-of-flight mass spectrometric methodology for geographical origin verification of coffee. *Anal. Chim. Acta* **617**: 72–84.
- SAGARPA. 2016. *Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y alimentación*, <http://www.sagarpa.gob.mx>.
- Salmona, J., Dussert, S., Descroix, F., de Kochko, A., Bertrand, B. & Joët, T. 2008. Deciphering transcriptional networks that govern *Coffea arabica* seed development using combined cDNA array and real-time RT-PCR approaches. *Plant Mol. Biol.* **66**: 105–124.
- Saxena, A. & Cramer, C.S. 2013. Metabolomics: A Potential Tool for Breeding Nutraceutical Vegetables. *Adv Crop Sci Tech* **1**: 2.
- Schauer, N., Semel, Y., Balbo, I., Steinfath, M., Repsilber, D., Selbig, J., *et al.* 2008. Mode of Inheritance of Primary Metabolic Traits in Tomato. *Plant Cell* **20**: 509–523.

- Schauer, N., Semel, Y., Roessner, U., Gur, A., Balbo, I., Carrari, F., *et al.* 2006. Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat Biotech* **24**: 447–454.
- Semmelroch, P. & Grosch, W. 1996. Studies on Character Impact Odorants of Coffee Brews. *J. Agric. Food Chem.* **44**: 537–543.
- Semmelroch, P., Laskawy, G., Blank, I. & Grosch, W. 1995. Determination of potent odourants in roasted coffee by stable isotope dilution assays. *Flavour and Fragrance Journal* **10**: 1–7.
- Silva, M. do C., Várzea, V., Guerra-Guimarães, L., Azinheira, H.G., Fernandez, D., Petitot, A.-S., *et al.* 2006. Coffee resistance to the main diseases: leaf rust and coffee berry disease. *Brazilian Journal of Plant Physiology* **18**: 119–147.
- Silvestrini, M., Junqueira, M.G., Favarin, A.C., Guerreiro-Filho, O., Maluf, M.P., Silvarolla, M.B., *et al.* 2007. Genetic diversity and structure of Ethiopian, Yemen and Brazilian *Coffea arabica*. *Genet Resour Crop Evol* **54**: 1367–1379.
- Smrke, S., Krosiakova, I., Gloess, A.N. & Yeretizian, C. 2015. Differentiation of degrees of ripeness of Catuai and Tipica green coffee by chromatographical and statistical techniques. *Food Chemistry* **174**: 637–642.
- Sobreira, F.M., de Oliveira, A.C.B., Pereira, A.A. & Sakyama, N.S. 2015. Potential of Híbrido de Timor germplasm and its derived progenies for coffee quality improvement. *Australian Journal of Crop Science* **9**: 289.
- Spadone, J.C., Takeoka, G. & Liardon, R. 1990. Analytical investigation of Rio off-flavor in green coffee. *J. Agric. Food Chem.* **38**: 226–233.
- Steiger, L., Nagai, C., Moore, H., Morden, W., Osgood, V. & Ming, R. 2002. AFLP analysis of genetic diversity within and among *Coffea arabica* cultivars. *Theor. Appl. Genet.* **105**: 209–215.
- Steinfath, M., Strehmel, N., Peters, R., Schauer, N., Groth, D., Hummel, J., *et al.* 2010. Discovering plant metabolic biomarkers for phenotype prediction using an untargeted approach. *Plant Biotechnol. J.* **8**: 900–911.
- Sulpice, R., Trenkamp, S., Steinfath, M., Usadel, B., Gibon, Y., Witucka-Wall, H., *et al.* 2010. Network analysis of enzyme activities and metabolite levels and their relationship to biomass in a large panel of Arabidopsis accessions. *Plant Cell* **22**: 2872–2893.
- Sunarharum, W.B., Williams, D.J. & Smyth, H.E. 2014. Complexity of coffee flavor: A compositional and sensory perspective. *Food Research International* **62**: 315–325.
- Teixeira-Cabral, T.A., Sakiyama, N.S., Zambolim, L., Pereira, A.A. & Schuster, I. 2004. Single-locus inheritance and partial linkage map of *Coffea arabica* L. *Cropp Breeding and Applied Biotechnology* **4**: 416–421.

- Teressa, A., Crouzillat, D., Petiard, V. & Brouhan, P. 2010. Genetic diversity of Arabica coffee (*Coffea arabica* L.) collections. *EJAST* **1**: 63–79.
- Toci, A.T. & Farah, A. 2008. Volatile compounds as potential defective coffee beans' markers. *Food Chem* **108**: 1133–1141.
- Toledo, P.R.A.B., Pezza, L., Pezza, H.R. & Toci, A.T. 2016. Relationship Between the Different Aspects Related to Coffee Quality and Their Volatile Compounds. *Comprehensive Reviews in Food Science and Food Safety* **15**: 705–719.
- Tran, H.T., Lee, L.S., Furtado, A., Smyth, H. & Henry, R.J. 2016. Advances in genomics for the improvement of quality in coffee. *J. Sci. Food Agric.* n/a-n/a.
- Tran, T. 2005. Genetic variation in cultivated coffee (*Coffea arabica* L.) accessions in northern New South Wales, Australia. *Masters thesis, Southern Cross University, Lismore, NSW.*
- Trugo, L.C. 2003. COFFEE | Analysis of Coffee Products A2 - Caballero, Benjamin. In: *Encyclopedia of Food Sciences and Nutrition (Second Edition)*, pp. 1498–1506. Academic Press, Oxford.
- Urgert, R., van der Weg, G., Kosmeijer-Schuil, T.G., van de Bovenkamp, P., Hovenier, R. & Katan, M.B. 1995. Levels of the Cholesterol-Elevating Diterpenes Cafestol and Kahweol in Various Coffee Brews. *J. Agric. Food Chem.* **43**: 2167–2172.
- USDA, U.S.D. of A. 2016. *Coffee: World Markets and Trade.*
- Vega, F.E., Ebert, A.W. & Ming, R. 2007. Coffee Germplasm Resources, Genomics and Breeding. In: *Plant Breeding Reviews* (J. Janick, ed), pp. 415–447. John Wiley & Sons, Inc.
- Viani, R. & Illy, A. 2005. *Espresso Coffee: The Science of Quality*, Edición: 2nd ed. Elsevier Ltd, Oxford, Amsterdam ; Boston.
- Vieira, E.S.N., Pinho, É.V. de R.V., Carvalho, M.G.G., Esselink, D.G. & Vosman, B. 2010. Development of microsatellite markers for identifying Brazilian *Coffea arabica* varieties. *Genet Mol Biol* **33**: 507–514.
- Vignoli, J.A., Bassoli, D.G. & Benassi, M.T. 2011. Antioxidant activity, polyphenols, caffeine and melanoidins in soluble coffee: The influence of processing conditions and raw material. *Food Chemistry* **124**: 863–868.
- Visscher, P.M., Hill, W.G. & Wray, N.R. 2008. Heritability in the genomics era — concepts and misconceptions. *Nature Reviews Genetics* **9**: 255–266.
- Winkler, R. 2015. SpiderMass: Semantic database creation and tripartite metabolite identification strategy. *J. Mass Spectrom.* **50**: 538–541.
- Wintgens, J.N. & Zamarripa Colmenero. 2008. Coffee Propagation. In: *Coffee: Growing, Processing, Sustainable Production*, pp. 87–136. Wiley-VCH Verlag GmbH & Co.

KGaA, Switzerland.

Zabotina, O.A. 2013. Metabolite-Based Biomarkers for Plant Genetics and Breeding. In: *Diagnostics in Plant Breeding* (T. Lübberstedt & R. K. Varshney, eds), pp. 281–309. Springer Netherlands.

Zamarripa Colmenero, A. & Escamilla Prado, E. 2002. *Varietades de café en México: origen, características y perspectivas*. Universidad Autónoma Chapingo, Centro Regional Universitario Oriente, Centro Nacional de Investigación para el Desarrollo de Regiones Cafetaleras, Huatusco, Veracruz.

11. ANEXOS

11.1 Formato de evaluación sensorial descriptivo de la bebida de café.

Fecha																														
Juez																														
Instrucciones	<p>Deguste las muestras en el orden presentado e indique el valor del atributo. Enjuáguese bien la boca antes de iniciar la degustación, así como entre muestra y muestra.</p>																													
	Muestra 1					Muestra 2					Muestra 3																			
	0	1	2	3	4	5	0	1	2	3	4	5	0	1	2	3	4	5	0	1	2	3	4	5	0	1	2	3	4	5
Intensidad aromática																														
Calidad aromática																														
Sabor																														
Astringencia																														
Resabio																														
Acidez																														
Amargor																														
Cuerpo																														
Agrio																														
Preferencia																														

11.2 Escala de 6 puntos para evaluación de atributos sensoriales

- 0 No presente (Ausente)
- 1 Reconocible o en el umbral (muy débil)
- 2 Débil (ligero)
- 3 Moderado
- 4 Fuerte
- 5 Muy fuerte

11.3 Presentación de bebidas de café para su degustación



11.4 Metabolitos de frutos con niveles altos de heredabilidad (≥ 0.4).

Metabolito/ <i>m/z</i>	Heredabilidad	Error estándar
N-methylputrescine	0.6519	0.0968
4-Aminobutanoic acid	0.4055	0.1012
Succinic acid	0.6292	0.1157
L-threonine	0.5204	0.1398
N-carbamoylputrescine	0.4170	0.0703
Acetylphosphate	0.5535	0.1219
Indole	0.4638	0.0977
Isoeugenol	0.5080	0.1292
Octanoate	0.6143	0.1136
Acetylphosphate	0.6748	0.0623
Allantoin	0.6085	0.1270
Tartaric acid	0.5288	0.1685
Caffeine	0.8072	0.1157
α -keto adipate	0.4557	0.0955
Serotonin	0.5674	0.1272
phosphoenolpyruvate	0.4060	0.1319
2-oxo-3-carboxy-4,5-cyclopropylhex-5-enoate	0.5908	0.0931
4-Methyl-5-thiazoleethanol phosphate	0.4713	0.1057
Deoxyinosine	0.4260	0.1539
Heptadecane	0.8852	0.1281
Xanthoxin	0.4966	0.1110
5-aminoimidazole ribonucleotide	0.7063	0.1233
1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate	0.4952	0.1243
Oleic acid	0.8063	0.2041
13(S)-hydroperoxyoctadeca-9,11-dienoate	0.7409	0.1415
S-methyl-5'-thioadenosine	0.6339	0.1193
Ent-kaur-16-en-19-ol	0.6063	0.2905
N-acetyl-glucosamine-1-phosphate	0.7557	0.0663
D-ribulose-1,5-bisphosphate	0.4578	0.1271
Arachidate	0.5078	0.1068
2,5-diamino-6-(ribosylamino)-4-(3H)-pyrimidinone 5'-phosphate	0.6269	0.1494
3-deoxy-D-manno-octulosonate 8-P	0.6552	0.2607
3-deoxy-D-manno-octulosonate 8-P	0.7785	0.1265
(2-amino-4-hydroxy-7,8-dihydropteridin-6-yl)methyl diphosphate	0.4338	0.1211
Vinorine	0.4051	0.0905
Dihydroneopterin phosphate	0.5462	0.1071
Xanthosine-5-phosphate	0.5837	0.1381
D-sedoheptulose-1,7-bisphosphate	0.4273	0.1262
Desulfoglucotropeolin	0.4456	0.1148
3-Feruloyl-1,5-lactone	0.7116	0.0770
(6z,8e,10e,14z)-(5s,12r)-5,12-dihydroxyicosa-6,8,10,14-tetraenoate	0.8857	0.0994
Caffeoyl tyrosine	0.7350	0.1150

GA38	0.7291	0.1397
Lignoceric acid	0.5836	0.0941
(2-amino-4-hydroxy-7,8-dihydropteridin-6-yl)methyl diphosphate	0.9586	0.1708
GA43	0.4031	0.1417
Lignoceric acid	0.6041	0.1213
Squalene	0.5357	0.1108
6-deoxoteasterone	0.4014	0.1690
(22?)-hydroxy-campesterol	0.6814	0.1027
Guanosine-5'-diphosphate	0.5326	0.0841
Trans-zeatin riboside monophosphate	1.0820	0.1762
5'-phosphoribosyl-4-(N-succinocarboxamide)-5-aminoimidazole	0.4058	0.1523
Sophoradiol	0.8725	0.2087
Tocopherol	0.8942	0.1082
Trans-zeatin riboside monophosphate	0.6736	0.0923
Flavin mononucleotide	0.5229	0.2572
Delphinidin 3-O-beta-D-glucopyranoside	0.4619	0.0662
Phytol diphosphate	0.4397	0.0795
D-myo-inositol (1,3,4,5)-tetrakisphosphate	0.7654	0.1004
Nostocyclamide	1.0329	0.2797
AA-5-HT	0.7077	0.1117
Phosphoadenosine-5'-phosphosulfate	0.4041	0.0754
CTP	0.5345	0.1421
Nostocyclamide	0.7828	0.2053
Nostocyclamide	0.8223	0.0921
D-myo-inositol (1,2,5,6) tetrakisphosphate	0.7411	0.1578
3,5-O-dicaffeoylquinic acid	0.4002	0.1543
7,8-dihydroneopterin triphosphate	0.4804	0.1115
Molybdenum cofactor	0.8415	0.1084
4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol	0.7999	0.1251
Guanosine 5'-triphosphate	0.6860	0.1535
Phosphoadenosine-5'-phosphosulfate	0.9277	0.2113
Tenuocyclamide C	0.4506	0.1405
Adenosine-5'-phosphate	0.4322	0.1701
Strictosidine	0.4795	0.0779
2'-deoxyguanosine-5'-triphosphate	0.6223	0.1393
Trans-zeatin riboside diphosphate	0.4516	0.1109
Strictosidine	0.4285	0.1433
Pyropheophorbide-a (PPa)	0.4345	0.1197
7,8-dihydroneopterin triphosphate	0.7519	0.1354
3,5-O-dicaffeoylquinic acid	0.4732	0.1659
Phytoene	0.6340	0.1367
Dihydrozeatin-9-N-glucoside-O-glucoside	0.5568	0.1713
Nostocyclamide M	0.5550	0.1289
UDP-4-dehydro-6-deoxy-D-glucose	0.6441	0.1142

UDP-L-rhamnose	0.7788	0.0802
Strictosidine	0.8336	0.1326
Beta-cryptoxanthin	0.5129	0.2560
Pelargonidin-3-O-?-D-glucoside	0.7201	0.1615
Pyropheophorbide-a (PPa)	0.7399	0.0749
1,2-Dilauroyl-sn-Glycero-3-Phosphate	0.4465	0.1498
3-Caffeoyl-4-sinapoylquinic acid	0.7410	0.0851
Molybdenum cofactor	0.4845	0.1427
4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol	0.5628	0.1037
Phosphoribosyl-AMP	0.5244	0.1809
Guanosine 5'-triphosphate	0.6217	0.1938
Protoporphyrin IX	0.4798	0.1325
Zeta-carotene	0.4626	0.0897
CMP-3-deoxy-D-manno-octulosonate	0.4628	0.2239
UDP-D-galactose	0.7406	0.1814
Dihydrozeatin-9-N-glucoside-O-glucoside	0.5392	0.0785
Pyropheophorbide-a (PPa)	0.7967	0.1411
Beta,epsilon-carotene-3-diol	0.9049	0.1348
Neurosporene	0.5345	0.1509
Phosphoribulosylformimino-AICAR-P	0.4695	0.0850
PhosphoribulosylformiminoAICAR-phosphate	0.7467	0.1203
2,3-trans-proanthocyanidin	0.8565	0.1975
Pelargonidin-3-O-rutinoside	0.8042	0.3343
UDP-D-glucuronate	1.0591	0.0894
Phytofluene	0.6332	0.1565
CMP-3-deoxy-D-manno-octulosonate	0.7956	0.1169
Prephytoene diphosphate	0.5382	0.1271
dTDP-4-dehydro-6-deoxy-L-mannose	0.5368	0.1713
Antheraxanthin	0.7584	0.0712
UDP-4-dehydro-6-deoxy-D-glucose	0.6117	0.1045
UDP-L-rhamnose	0.7324	0.1375
ADP-D-glucose	0.6875	0.1724
ADP-D-glucose	0.5880	0.1340
Lactucaxanthin	0.9387	0.1172
1-Palmitoleoyl-2-oleoyl-sn-glycerol	0.9457	0.1311
Cyanidin 3-(p-coumaroyl)-glucoside	0.4355	0.0986
Obyanamide	0.4945	0.1261
2,3-trans-proanthocyanidin	0.9569	0.2185
Pelargonidin 3-(6-p-coumaroyl)glucoside	0.5287	0.1116
2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol	0.5543	0.1238
UDP-D-glucuronate	0.4148	0.1040
ppGpp	0.5423	0.0827
UDP-D-galactose	0.5982	0.2071
UDP-N-acetyl-D-glucosamine	0.5970	0.0982

Presqualene diphosphate	0.6466	0.1035
Isopentenyladenosine-5'-triphosphate	0.5196	0.1870
Isopentenyladenosine-5'-triphosphate	0.5149	0.0679
C11832	0.4182	0.1539
Cyanidin-3,5-O-diglucoside	0.5085	0.1244
Pheophorbide?a	0.5929	0.1025
Phosphoribulosylformimino-AICAR-P	0.5620	0.1902
PhosphoribosylformiminoAICAR-phosphate	0.6111	0.1196
2,3-trans-proanthocyanidin	0.8551	0.0925
Cyanidin 3-(p-coumaroyl)-glucoside	0.5745	0.0926
Pelargonidin-3-O-rutinoside	0.5185	0.1771
UDP-D-galacturonate	0.5609	0.1262
Violaxanthin	0.7623	0.0903
2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol	0.7446	0.0484
All-trans-hexaprenyl diphosphate	0.7015	0.0889
Tetrahydropteroyltri-L-glutamate	0.7269	0.2428
Red chlorophyll catabolite	0.8663	0.0791
Guanosine 5'-diphosphoglucose	0.4871	0.1062
UDP-sulfoquinovose	0.5471	0.1171
Chlorophyllide b	0.6273	0.0602
Pheophorbide?a	0.4185	0.1139
Pelargonidin-3,5-diglucoside	0.5303	0.1665
Glutathione disulfide	0.4954	0.0728
Glutathione disulfide	0.4926	0.1452
Chlorophyllide a	0.7679	0.1020
2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol	0.4098	0.1293
5-methyltetrahydropteroyltri-L-glutamate	0.5966	0.1690
ppGpp	0.8453	0.1593
GDP-L-galactose	0.4725	0.1102
Delphinidin-3,5-diglucoside	0.6607	0.1457
Divinyl chlorophyllide a	0.4023	0.0590
Protoheme IX	0.5654	0.1269
Soyasapogenol B-3-O-?-glucuronide	0.8756	0.1512
5-methyltetrahydropteroyltri-L-glutamate	0.5044	0.1660
Alpha-D-ribose-1-phosphate	0.5028	0.1453
Tetrahydropteroyltri-L-glutamate	0.4362	0.0669
UDP-sulfoquinovose	0.7758	0.0933
Mannan	0.4872	0.0803
Chlorophyllide b	0.5513	0.0560
Coproporphyrinogen III	0.4924	0.1028
Coproporphyrinogen III	0.4755	0.0497
Nicotinate adenine dinucleotide	0.5103	0.1108
NADH	0.7543	0.2142
Microginin 690	0.4075	0.1784

All-trans-heptaprenyl diphosphate	0.5194	0.1381
Phytate	0.5675	0.0970
Delphinidin 3,5-diglucoside	0.8635	0.0995
Alpha-D-ribose-1-phosphate	0.7487	0.1730
NADH	0.6367	0.1213
Stachyose	0.5039	0.1793
Dephospho-CoA	0.6141	0.1277
Phosphoribosyl-ATP	0.5311	0.0731
Molybdopterin-AMP	0.6111	0.0973
All-trans-octaprenyl diphosphate	0.4208	0.1607
Molybdopterin-AMP	0.7709	0.1828
Dephospho-CoA	0.9125	0.0445
Ubiquinol-8	0.6332	0.1763
Microginin 690	0.5401	0.0703
1-sn-Phosphatidylcholine, 2,3-dipalmitoyl	0.7488	0.3365
Phosphoribosyl-ATP	0.4745	0.1178
NADPH	0.9202	0.0804
NADP+	0.5699	0.1035
NADPH	0.6961	0.0661
Molybdopterin-AMP	0.7425	0.1625
Ubiquinol-8	0.6745	0.1608
1-sn-Phosphatidylcholine, 2,3-dipalmitoyl	0.7841	0.0752
Bisdemalonylsalvianin	0.9662	0.1577
Phosphoribosyl-ATP	0.6859	0.3014
All-trans-octaprenyl diphosphate	1.0310	0.1043
Molybdopterin-AMP	0.7134	0.0928
1-18:3-2-trans-16:1-phosphatidylglycerol	0.7507	0.1905
1-18:3-2-16:0-phosphatidylglycerol	0.7011	0.0979
Palmitoyl-linoleoyl	0.8081	0.1539
1-16:0-2-18:1-phosphatidylcholine	0.7930	0.2278
1-18:3-2-18:3-phosphatidylcholine	0.9584	0.1071
Delphinidin 3,5,3'-triglucoside	0.6190	0.0760
Delphinidin 3-O-glucosyl-5-O-caffeoylglucoside	0.9487	0.2687
Coenzyme A	0.7178	0.0978
All-trans-nonaprenyl diphosphate	0.9842	0.0417
Atractyloside	0.4168	0.0909
Nostocyclopeptide A3	0.4498	0.1012
Coenzyme A	0.6451	0.0656
Carboxyatractyloside	0.7585	0.1757
Delphinidin 3-O-glucosyl-5-O-caffeoylglucoside	1.3290	0.1966
Maltopentaose	0.7625	0.1037
Delphinidin 3,5,3'-triglucoside	0.6344	0.1431
UDP-3-O-(3-hydroxymyristoyl)glucosamine	0.7285	0.0902
Isobutyryl-CoA	0.8425	0.1508

Crotonyl-CoA	0.7292	0.1220
Symplostatin 1	0.7755	0.0731
Monodemalonylsalvianin	0.6917	0.1538
Acetoacetyl-CoA	0.6667	0.1960
1,2-Dilinoleoyl-3-palmitoyl-rac-glycerol	0.6189	0.1122
Methylacrylyl-CoA	0.6512	0.0698
Succinyl-CoA	0.6094	0.0629
Benzoyl-CoA	0.6692	0.0956
Chlorophyll	0.7371	0.1269
Benzoyl-CoA	0.7681	0.0473
Succinyl-CoA	1.0294	0.1626
B-Chlorophyll	1.0093	0.2089
Bacteriochlorophyll?a	0.6828	0.1124
Glutaryl-CoA	0.5840	0.1057
All-trans-undecaprenyl diphosphate	0.4191	0.1497
Bacteriopheophytin	1.0320	0.0334
Trans-?2, cis-?4-decadienoyl-CoA	0.8815	0.0745
Decanoyl-CoA	0.7032	0.0456
All-trans-undecaprenyl diphosphate	0.7828	0.2236
Salvianin	0.8605	0.0658
2-trans,5-cis-tetradecadienoyl-CoA	0.6125	0.0503
3,5-tetradecadienoyl-CoA	0.8331	0.1191
Myristoyl-CoA	0.4554	0.0752
Feruloyl-CoA	0.7847	0.0639
<i>m/z</i> 337.7305	0.5326	0.1087
<i>m/z</i> 356.2268	0.4408	0.1327
<i>m/z</i> 365.8207	0.4194	0.1326
<i>m/z</i> 376.4213	0.4576	0.1246
<i>m/z</i> 381.9092	0.9265	0.0635
<i>m/z</i> 434.5937	0.4917	0.1640
<i>m/z</i> 457.9324	0.5395	0.1463
<i>m/z</i> 486.3118	0.6119	0.0891
<i>m/z</i> 499.8166	0.5101	0.1185
<i>m/z</i> 500.0636	0.5590	0.1530
<i>m/z</i> 500.5095	0.6520	0.0957
<i>m/z</i> 513.7461	0.5374	0.0862
<i>m/z</i> 513.9543	0.4020	0.1705
<i>m/z</i> 514.1586	0.7347	0.2254
<i>m/z</i> 516.7384	0.5806	0.1592
<i>m/z</i> 517.2225	0.5624	0.1152
<i>m/z</i> 518.1848	0.7658	0.2068
<i>m/z</i> 532.8962	0.9864	0.1010
<i>m/z</i> 534.7274	0.5273	0.0982
<i>m/z</i> 536.0019	0.6556	0.0979

<i>m/z</i> 536.2469	0.5554	0.1131
<i>m/z</i> 538.0196	0.7868	0.1472
<i>m/z</i> 545.7044	0.8079	0.1935
<i>m/z</i> 552.1077	0.7347	0.0835
<i>m/z</i> 557.3547	0.7569	0.0583
<i>m/z</i> 557.8234	0.6576	0.1865
<i>m/z</i> 561.8470	0.8227	0.0957
<i>m/z</i> 571.6767	0.9830	0.1609
<i>m/z</i> 573.4349	0.5355	0.1303
<i>m/z</i> 574.1632	0.9321	0.1443
<i>m/z</i> 586.8697	0.4662	0.1078
<i>m/z</i> 587.9178	0.5639	0.1139
<i>m/z</i> 595.6598	0.9864	0.1398
<i>m/z</i> 597.0433	0.5513	0.0955
<i>m/z</i> 597.1639	0.4033	0.0993
<i>m/z</i> 597.4666	0.6744	0.0955
<i>m/z</i> 597.6050	0.8386	0.1653
<i>m/z</i> 597.6393	0.6147	0.0976
<i>m/z</i> 598.0612	0.4552	0.0920
<i>m/z</i> 598.5407	1.1561	0.3004
<i>m/z</i> 599.9172	0.6367	0.0565
<i>m/z</i> 599.9840	0.6123	0.1022
<i>m/z</i> 609.6398	0.9338	0.0783
<i>m/z</i> 609.9024	0.7648	0.0642
<i>m/z</i> 613.5385	1.0004	0.0909
<i>m/z</i> 613.9050	0.4643	0.1264
<i>m/z</i> 619.2760	0.4782	0.1267
<i>m/z</i> 620.2668	0.5523	0.0735
<i>m/z</i> 621.4600	0.5839	0.0783
<i>m/z</i> 622.5163	0.5033	0.1838
<i>m/z</i> 628.9870	0.5866	0.1333
<i>m/z</i> 631.6777	0.9269	0.0808
<i>m/z</i> 632.0465	0.7677	0.2291
<i>m/z</i> 635.6373	0.5751	0.1167
<i>m/z</i> 638.8974	0.4402	0.0890
<i>m/z</i> 639.8521	0.7176	0.1542
<i>m/z</i> 640.9883	0.4363	0.1196
<i>m/z</i> 641.3805	0.5779	0.1682
<i>m/z</i> 641.6339	0.4841	0.1064
<i>m/z</i> 642.8470	0.4330	0.0993
<i>m/z</i> 645.0176	0.5463	0.1668
<i>m/z</i> 646.8831	0.6649	0.1357
<i>m/z</i> 648.7114	0.6646	0.0678
<i>m/z</i> 649.3521	0.5440	0.2591

<i>m/z</i> 652.5124	0.7716	0.2289
<i>m/z</i> 653.9153	0.8731	0.1531
<i>m/z</i> 654.9806	0.5816	0.1743
<i>m/z</i> 655.6807	0.5892	0.0926
<i>m/z</i> 655.9988	0.7768	0.1285
<i>m/z</i> 657.9407	0.6591	0.1586
<i>m/z</i> 658.8858	0.4851	0.0853
<i>m/z</i> 658.9735	0.5425	0.1299
<i>m/z</i> 659.9839	0.4635	0.1094
<i>m/z</i> 663.7809	0.7098	0.1097
<i>m/z</i> 667.7796	0.4866	0.1074
<i>m/z</i> 668.0492	0.9500	0.1044
<i>m/z</i> 668.9216	0.5051	0.2187
<i>m/z</i> 669.4101	0.5671	0.0877
<i>m/z</i> 670.8874	0.4243	0.0904
<i>m/z</i> 671.7527	0.4836	0.1111
<i>m/z</i> 672.1082	0.4310	0.1371
<i>m/z</i> 676.0284	0.4609	0.1209
<i>m/z</i> 678.3576	0.6051	0.0886
<i>m/z</i> 679.9644	1.1140	0.2340
<i>m/z</i> 680.9191	0.6912	0.2093
<i>m/z</i> 681.7438	0.6822	0.1678
<i>m/z</i> 684.0899	0.6448	0.1003
<i>m/z</i> 684.4961	0.4325	0.0833
<i>m/z</i> 691.9720	0.6881	0.1066
<i>m/z</i> 694.7038	0.4051	0.1416
<i>m/z</i> 695.7722	0.4627	0.0921
<i>m/z</i> 695.8965	0.5168	0.1096
<i>m/z</i> 696.3254	0.5660	0.1250
<i>m/z</i> 696.7964	0.6959	0.1512
<i>m/z</i> 697.8008	0.6256	0.1230
<i>m/z</i> 699.4197	0.5304	0.1452
<i>m/z</i> 699.8030	0.4532	0.0840
<i>m/z</i> 700.3195	0.4301	0.1010
<i>m/z</i> 704.1657	0.5856	0.1396
<i>m/z</i> 705.6388	0.6299	0.0677
<i>m/z</i> 707.7433	0.6300	0.0947
<i>m/z</i> 710.9163	0.6311	0.0943
<i>m/z</i> 711.7637	0.6111	0.1326
<i>m/z</i> 712.0248	0.6519	0.0822
<i>m/z</i> 714.8598	0.6229	0.1050
<i>m/z</i> 715.5803	0.7855	0.1008
<i>m/z</i> 715.8470	0.7343	0.0668
<i>m/z</i> 717.1903	0.8438	0.0444

<i>m/z</i> 717.1952	0.7732	0.0785
<i>m/z</i> 718.6067	0.8567	0.0774
<i>m/z</i> 718.7314	0.6272	0.1070
<i>m/z</i> 719.3605	1.0060	0.0272
<i>m/z</i> 726.6106	0.4625	0.1233
<i>m/z</i> 727.0199	0.9746	0.0727
<i>m/z</i> 728.1401	0.7072	0.1091
<i>m/z</i> 729.6312	0.8361	0.1040
<i>m/z</i> 730.7957	1.0397	0.0704
<i>m/z</i> 730.9537	0.7068	0.1296
<i>m/z</i> 731.4041	0.5871	0.1207
<i>m/z</i> 732.6810	0.7175	0.1012
<i>m/z</i> 733.1851	0.8627	0.0719
<i>m/z</i> 734.4908	0.8173	0.0982
<i>m/z</i> 735.6302	0.9221	0.0884
<i>m/z</i> 736.1153	0.8945	0.0651
<i>m/z</i> 736.8592	0.6188	0.2370
<i>m/z</i> 737.5399	0.4138	0.1169
<i>m/z</i> 738.8142	0.5927	0.1254
<i>m/z</i> 739.8847	0.6587	0.1379
<i>m/z</i> 741.8405	0.6915	0.1264
<i>m/z</i> 744.7437	0.4110	0.0928
<i>m/z</i> 748.7589	0.6155	0.1765
<i>m/z</i> 750.7444	0.7536	0.1514
<i>m/z</i> 752.7969	0.4237	0.1143
<i>m/z</i> 753.7778	0.8072	0.0462
<i>m/z</i> 753.7968	0.4259	0.1332
<i>m/z</i> 755.0653	0.4838	0.0626
<i>m/z</i> 757.8108	1.0403	0.1422
<i>m/z</i> 758.9998	1.0361	0.0565
<i>m/z</i> 759.8082	0.4475	0.1124
<i>m/z</i> 760.0094	0.8530	0.1513
<i>m/z</i> 761.5839	0.5692	0.1144
<i>m/z</i> 761.6315	0.8763	0.1735
<i>m/z</i> 762.7094	0.4090	0.1595
<i>m/z</i> 764.8122	0.7810	0.1580
<i>m/z</i> 771.6912	0.8845	0.1085
<i>m/z</i> 773.0535	0.9073	0.0834
<i>m/z</i> 773.1402	0.5316	0.1542
<i>m/z</i> 775.7207	0.6339	0.1709
<i>m/z</i> 776.6643	0.8677	0.1140
<i>m/z</i> 777.2834	0.9010	0.1021
<i>m/z</i> 788.8339	0.4495	0.1390
<i>m/z</i> 793.4099	0.4082	0.1164

<i>m/z</i> 794.6728	1.0847	0.0943
<i>m/z</i> 797.5153	0.8602	0.0556
<i>m/z</i> 797.6951	0.4400	0.1900
<i>m/z</i> 801.6881	0.4227	0.1221
<i>m/z</i> 802.4507	0.4763	0.0769
<i>m/z</i> 805.5855	0.6207	0.1378
<i>m/z</i> 805.8779	0.4884	0.0815
<i>m/z</i> 811.2030	0.6070	0.1124
<i>m/z</i> 811.8690	0.4980	0.0950
<i>m/z</i> 817.0730	0.6988	0.0697
<i>m/z</i> 817.3395	0.7149	0.0922
<i>m/z</i> 818.8724	0.9747	0.1360
<i>m/z</i> 833.8109	0.8612	0.1035
<i>m/z</i> 835.1677	0.7176	0.1203
<i>m/z</i> 835.2768	0.8545	0.0456
<i>m/z</i> 839.0698	0.9023	0.0896
<i>m/z</i> 844.9400	0.8202	0.1010
<i>m/z</i> 856.5722	0.5144	0.1443
<i>m/z</i> 860.5079	1.2000	0.2290
<i>m/z</i> 861.7196	0.5024	0.1045
<i>m/z</i> 869.0134	0.8751	0.1103
<i>m/z</i> 872.2035	0.7920	0.0977
<i>m/z</i> 873.9837	0.5266	0.0859
<i>m/z</i> 875.5625	0.6743	0.1639
<i>m/z</i> 879.5471	0.5509	0.1742
<i>m/z</i> 881.4225	0.5305	0.0687
<i>m/z</i> 882.7670	0.9707	0.0496
<i>m/z</i> 883.1307	0.6887	0.1085
<i>m/z</i> 885.1049	0.8449	0.0959
<i>m/z</i> 885.4505	0.7643	0.0457
<i>m/z</i> 889.9879	0.5488	0.0899
<i>m/z</i> 895.0077	0.4775	0.0448
<i>m/z</i> 900.9528	0.7806	0.1156
<i>m/z</i> 901.3689	0.4097	0.0884
<i>m/z</i> 903.7270	0.5932	0.1043
<i>m/z</i> 913.5948	1.2629	0.2084
<i>m/z</i> 923.2354	0.4820	0.3316
<i>m/z</i> 926.5988	0.5478	0.1147
<i>m/z</i> 926.7478	1.1206	0.0656
<i>m/z</i> 937.3961	0.4883	0.1331
<i>m/z</i> 939.3691	0.4806	0.1804
<i>m/z</i> 940.7196	0.4652	0.1753
<i>m/z</i> 941.7020	0.4821	0.1222
<i>m/z</i> 943.4075	0.7547	0.0484

<i>m/z</i> 950.0953	1.0804	0.1816
<i>m/z</i> 953.2170	0.6946	0.1258
<i>m/z</i> 954.3670	0.5187	0.1396
<i>m/z</i> 963.4135	0.9600	0.0270
<i>m/z</i> 967.3747	0.8241	0.0892
<i>m/z</i> 968.7753	0.4366	0.1436
<i>m/z</i> 969.2803	0.5276	0.0395
<i>m/z</i> 981.2534	0.7526	0.1082
<i>m/z</i> 997.6674	0.7921	0.0576
<i>m/z</i> 999.4229	0.4117	0.1322
<i>m/z</i> 999.4478	0.4548	0.0794
<i>m/z</i> 999.4611	0.8872	0.0477

11.5 Metabolitos de granos tostados correlacionados con atributos sensoriales de la bebida de café

(m/z)	Atributo sensorial	CC Spearman	Valor de p	Metabolito (m/z)	Atributo sensorial	CC Spearman	Valor de p
104.3321	Acidez	0.5309	1.33E-02	495.4826	Resabio	0.5339	1.27E-02
124.3265	Acidez	-0.5732	6.60E-03	495.4826	Sabor	0.5332	1.28E-02
138.3208	Acidez	0.7378	1.35E-04	496.9602	Acidez	0.6005	4.00E-03
138.3208	Amargor	-0.5824	5.60E-03	496.9602	Cuerpo	-0.5420	1.12E-02
138.3208	Calidad aromática	-0.5936	4.55E-03	498.7040	Acidez	0.6675	9.46E-04
138.3208	Cuerpo	-0.7574	7.00E-05	498.7040	Agrio	0.5928	4.62E-03
138.3208	Preferencia	-0.6779	7.32E-04	498.7040	Amargor	-0.6597	1.14E-03
138.3208	Resabio	-0.6519	1.36E-03	498.7040	Preferencia	-0.5219	1.52E-02
138.3208	Sabor	-0.5241	1.47E-02	498.8342	Acidez	0.5856	5.29E-03
138.4974	Resabio	0.5509	9.66E-03	498.8342	Cuerpo	-0.5668	7.38E-03
163.5760	Cuerpo	-0.5472	1.03E-02	500.6816	Acidez	0.5556	8.92E-03
191.7415	Astringencia	-0.7463	1.02E-04	500.6816	Agrio	0.6430	1.67E-03
191.7415	Cuerpo	-0.6157	2.96E-03	500.6816	Amargor	-0.6071	3.52E-03
191.7415	Resabio	-0.6532	1.32E-03	500.6816	Astringencia	-0.6845	6.21E-04
195.2429	Amargor	0.5388	1.17E-02	500.6816	Calidad aromática	-0.5852	5.33E-03
220.6596	Cuerpo	0.5348	1.25E-02	500.6816	Cuerpo	-0.6373	1.89E-03
220.6596	Resabio	0.5711	6.85E-03	500.6816	Preferencia	-0.6363	1.93E-03
268.6044	Intensidad aromática	0.5299	1.35E-02	500.6816	Resabio	-0.7399	1.26E-04
311.3270	Cuerpo	-0.5354	1.24E-02	500.6816	Sabor	-0.5657	7.53E-03
321.3260	Intensidad aromática	-0.6391	1.81E-03	516.7468	Astringencia	0.6610	1.10E-03
336.6077	Calidad aromática	-0.5780	6.06E-03	516.7468	Resabio	0.5574	8.66E-03
360.6934	Astringencia	0.6135	3.10E-03	518.5260	Intensidad aromática	-0.5156	1.67E-02
360.6934	Resabio	0.5143	1.71E-02	520.6462	Intensidad aromática	-0.5390	1.17E-02
368.9743	Calidad aromática	0.5507	9.68E-03	520.8069	Acidez	0.6493	1.45E-03
369.5516	Intensidad aromática	-0.5202	1.56E-02	522.9324	Cuerpo	-0.6353	1.97E-03
371.6194	Intensidad aromática	-0.6170	2.89E-03	533.1086	Acidez	-0.5205	1.56E-02
373.2366	Intensidad aromática	-0.5618	8.04E-03	533.1086	Agrio	-0.7081	3.28E-04
395.3556	Acidez	0.6890	5.52E-04	533.1086	Amargor	0.6415	1.72E-03

395.3556	Agrio	0.5570	8.72E-03	533.1086	Astringencia	0.5550	9.02E-03
395.3556	Amargor	-0.5928	4.63E-03	533.1086	Calidad aromática	0.5761	6.27E-03
395.3556	Preferencia	-0.5538	9.20E-03	533.1086	Preferencia	0.7423	1.17E-04
437.2851	Cuerpo	-0.5446	1.07E-02	533.1086	Resabio	0.6415	1.72E-03
468.5869	Preferencia	-0.5057	1.94E-02	533.1086	Sabor	0.7152	2.68E-04
469.9502	Acidez	0.5270	1.41E-02	535.8444	Agrio	-0.5309	1.33E-02
492.2000	Agrio	-0.5042	1.98E-02	535.8444	Calidad aromática	0.6411	1.74E-03
495.4826	Astringencia	0.5777	6.09E-03	535.8444	Preferencia	0.6123	3.17E-03
495.4826	Preferencia	0.5291	1.37E-02	535.8444	Sabor	0.6105	3.29E-03
537.7658	Acidez	0.6832	6.42E-04	601.8384	Acidez	0.7072	3.37E-04
537.7658	Cuerpo	-0.6556	1.25E-03	601.8384	Cuerpo	-0.6438	1.64E-03
537.7658	Preferencia	-0.5616	8.07E-03	601.8384	Preferencia	-0.5934	4.57E-03
537.7658	Resabio	-0.5215	1.53E-02	601.8384	Resabio	-0.5306	1.33E-02
538.0293	Amargor	0.5811	5.74E-03	602.6829	Intensidad aromática	0.5078	1.88E-02
539.6973	Acidez	-0.5986	4.15E-03	604.8364	Cuerpo	-0.5798	5.87E-03
539.6973	Agrio	-0.6671	9.56E-04	618.8237	Cuerpo	-0.5642	7.72E-03
539.6973	Amargor	0.5505	9.71E-03	630.7962	Intensidad aromática	-0.5000	2.10E-02
539.6973	Preferencia	0.5200	1.57E-02	632.5865	Agrio	-0.6860	5.96E-04
542.7310	Cuerpo	-0.5831	5.53E-03	653.6051	Astringencia	0.5120	1.76E-02
542.7310	Resabio	-0.5150	1.69E-02	659.7348	Acidez	-0.6057	3.61E-03
544.8249	Cuerpo	-0.5962	4.34E-03	663.6694	Acidez	-0.7521	8.41E-05
552.6862	Acidez	-0.6018	3.90E-03	663.6694	Amargor	0.5018	2.05E-02
552.6862	Amargor	0.5928	4.63E-03	665.7307	Acidez	-0.8009	1.29E-05
560.8530	Cuerpo	-0.6680	9.35E-04	665.7307	Agrio	-0.6143	3.05E-03
564.2915	Cuerpo	-0.5864	5.21E-03	665.7307	Amargor	0.6480	1.49E-03
565.6959	Acidez	-0.6552	1.27E-03	665.7307	Calidad aromática	0.7003	4.08E-04
565.6959	Agrio	-0.5785	6.01E-03	665.7307	Cuerpo	0.5315	1.32E-02
565.6959	Preferencia	0.5401	1.15E-02	665.7307	Preferencia	0.7137	2.80E-04
571.4393	Acidez	0.6871	5.80E-04	665.7307	Resabio	0.5769	6.18E-03
571.4393	Cuerpo	-0.6823	6.55E-04	665.7307	Sabor	0.6743	8.03E-04

571.4393	Preferencia	-0.5297	1.35E-02	667.0015	Acidez	-0.6708	8.74E-04
573.9350	Acidez	-0.5563	8.82E-03	667.0015	Agrio	-0.5726	6.67E-03
574.8521	Cuerpo	-0.6210	2.66E-03	667.0015	Amargor	0.5070	1.90E-02
576.0403	Acidez	0.5530	9.31E-03	668.2198	Agrio	-0.6241	2.50E-03
576.0403	Cuerpo	-0.5700	6.98E-03	668.2198	Amargor	0.5869	5.16E-03
581.1178	Acidez	-0.5986	4.15E-03	668.2198	Calidad aromática	0.5761	6.27E-03
581.1178	Calidad aromática	0.6021	3.88E-03	668.2198	Preferencia	0.6032	3.80E-03
581.1178	Cuerpo	0.5537	9.21E-03	668.2198	Sabor	0.6060	3.60E-03
581.1178	Resabio	0.5450	1.06E-02	671.4762	Acidez	0.6805	6.86E-04
588.8381	Cuerpo	-0.5296	1.36E-02	671.4762	Agrio	0.5160	1.67E-02
588.8381	Intensidad aromática	-0.5254	1.45E-02	674.8121	Calidad aromática	-0.5728	6.65E-03
595.4353	Acidez	0.5830	5.55E-03	674.8121	Intensidad aromática	-0.5865	5.20E-03
595.4353	Cuerpo	-0.6588	1.16E-03	674.8121	Resabio	-0.5150	1.69E-02
595.4353	Preferencia	-0.5921	4.68E-03	677.7247	Astringencia	0.5133	1.73E-02
600.0708	Cuerpo	-0.5485	1.00E-02	679.6623	Amargor	-0.5336	1.27E-02
600.0708	Resabio	-0.6278	2.31E-03	688.7849	Intensidad aromática	-0.5423	1.11E-02
690.8389	Acidez	-0.6376	1.88E-03	722.7305	Preferencia	-0.5219	1.52E-02
690.8389	Agrio	-0.6091	3.38E-03	722.7305	Resabio	-0.5104	1.81E-02
690.8389	Cuerpo	0.5413	1.13E-02	722.7305	Sabor	-0.5488	9.99E-03
690.8389	Preferencia	0.5635	7.81E-03	723.7042	Resabio	0.5385	1.18E-02
690.8471	Cuerpo	-0.5655	7.55E-03	724.0173	Cuerpo	-0.5126	1.75E-02
693.5954	Acidez	-0.5413	1.13E-02	724.0173	Resabio	-0.5926	4.64E-03
693.5954	Agrio	-0.5231	1.50E-02	727.9535	Acidez	0.6291	2.25E-03
693.5954	Amargor	0.6259	2.40E-03	727.9535	Agrio	0.7108	3.05E-04
693.5954	Preferencia	0.5330	1.29E-02	727.9535	Amargor	-0.7449	1.07E-04
693.5954	Resabio	0.5046	1.97E-02	727.9535	Astringencia	-0.5569	8.73E-03
696.6353	Cuerpo	0.5067	1.91E-02	727.9535	Calidad aromática	-0.6034	3.78E-03
696.6353	Resabio	0.5111	1.79E-02	727.9535	Preferencia	-0.7774	3.38E-05
706.7726	Cuerpo	-0.5759	6.29E-03	727.9535	Resabio	-0.6917	5.14E-04
709.4590	Acidez	-0.7918	1.90E-05	727.9535	Sabor	-0.6417	1.71E-03

709.4590	Agrio	-0.6586	1.17E-03	729.2108	Acidez	0.6044	3.71E-03
709.4590	Amargor	0.6240	2.50E-03	729.2108	Agrio	0.6580	1.19E-03
709.4590	Cuerpo	0.5811	5.73E-03	729.2108	Preferencia	-0.6298	2.21E-03
709.4590	Preferencia	0.6298	2.21E-03	729.2108	Resabio	-0.5704	6.93E-03
709.4590	Resabio	0.5522	9.45E-03	731.1906	Acidez	0.5576	8.63E-03
709.6914	Acidez	-0.5882	5.04E-03	731.1906	Calidad aromática	-0.5982	4.18E-03
713.2619	Acidez	-0.7352	1.46E-04	731.1906	Preferencia	-0.5486	1.00E-02
713.2619	Agrio	-0.5935	4.57E-03	731.1906	Resabio	-0.5639	7.76E-03
713.2619	Amargor	0.5856	5.28E-03	732.7101	Intensidad aromática	-0.5897	4.90E-03
713.2619	Astringencia	0.5224	1.51E-02	733.0316	Agrio	-0.5915	4.74E-03
713.2619	Calidad aromática	0.5416	1.12E-02	733.0316	Amargor	0.5720	6.74E-03
713.2619	Cuerpo	0.7405	1.24E-04	733.5720	Preferencia	0.5011	2.07E-02
713.2619	Preferencia	0.6974	4.41E-04	733.5720	Resabio	0.5156	1.67E-02
713.2619	Resabio	0.7529	8.18E-05	734.7701	Calidad aromática	-0.5104	1.81E-02
713.2619	Sabor	0.5995	4.08E-03	735.3842	Intensidad aromática	-0.6385	1.84E-03
715.4369	Acidez	0.5446	1.07E-02	736.7362	Cuerpo	-0.5629	7.90E-03
721.4833	Acidez	-0.5882	5.04E-03	737.1966	Calidad aromática	0.6216	2.63E-03
721.4833	Agrio	-0.7121	2.94E-04	737.1966	Preferencia	0.5739	6.52E-03
721.4833	Amargor	0.6532	1.32E-03	737.1966	Resabio	0.5997	4.06E-03
721.4833	Preferencia	0.6123	3.17E-03	737.1966	Sabor	0.5650	7.61E-03
721.4833	Resabio	0.5196	1.58E-02	740.3981	Acidez	-0.5810	5.74E-03
722.7305	Agrio	0.5779	6.08E-03	740.3981	Amargor	0.5323	1.30E-02
722.7305	Amargor	-0.5811	5.74E-03	741.0432	Agrio	-0.5707	6.90E-03
742.3064	Calidad aromática	-0.5020	2.04E-02	805.6283	Preferencia	0.5096	1.83E-02
743.4958	Amargor	0.5310	1.32E-02	808.7892	Cuerpo	-0.5204	1.56E-02
743.4958	Intensidad aromática	-0.5319	1.31E-02	816.6401	Acidez	0.5016	2.05E-02
743.6496	Acidez	0.5667	7.39E-03	818.6564	Acidez	0.5335	1.27E-02
743.6496	Cuerpo	-0.5602	8.26E-03	818.6564	Cuerpo	-0.6112	3.25E-03
743.6496	Preferencia	-0.5232	1.49E-02	820.6494	Intensidad aromática	-0.5722	6.72E-03
743.6496	Resabio	-0.5313	1.32E-02	821.6050	Acidez	-0.5628	7.91E-03

750.7424	Cuerpo	-0.5426	1.10E-02	821.6050	Agrio	-0.6378	1.87E-03
756.6541	Intensidad aromática	-0.5026	2.02E-02	821.6050	Amargor	0.6376	1.87E-03
758.8740	Acidez	0.5290	1.37E-02	821.6050	Resabio	0.5189	1.59E-02
760.8277	Acidez	0.6122	3.18E-03	822.7915	Amargor	-0.5401	1.15E-02
760.8277	Cuerpo	-0.5851	5.34E-03	822.7997	Cuerpo	-0.6177	2.85E-03
762.7181	Intensidad aromática	-0.6697	8.97E-04	822.8374	Acidez	-0.5172	1.63E-02
778.6983	Astringencia	0.6493	1.45E-03	825.0698	Resabio	0.5522	9.45E-03
778.8175	Cuerpo	-0.5387	1.17E-02	826.3390	Cuerpo	-0.5126	1.75E-02
780.6371	Astringencia	-0.6539	1.30E-03	833.5510	Acidez	0.5895	4.92E-03
780.8106	Acidez	0.5439	1.08E-02	840.8394	Intensidad aromática	-0.5013	2.06E-02
782.8089	Acidez	0.8139	7.16E-06	842.5597	Astringencia	-0.5146	1.70E-02
782.8089	Agrio	0.5049	1.96E-02	842.5597	Calidad aromática	-0.5806	5.78E-03
782.8089	Amargor	-0.5427	1.10E-02	842.5597	Cuerpo	-0.6301	2.20E-03
782.8089	Cuerpo	-0.6634	1.04E-03	842.5597	Preferencia	-0.6006	3.99E-03
782.8089	Preferencia	-0.6162	2.94E-03	842.5597	Resabio	-0.6558	1.25E-03
782.8089	Resabio	-0.5163	1.66E-02	842.5597	Sabor	-0.6294	2.24E-03
784.7896	Acidez	0.5088	1.85E-02	843.6908	Acidez	0.6350	1.98E-03
786.7998	Acidez	0.5576	8.63E-03	846.6486	Preferencia	-0.5336	1.27E-02
786.7998	Cuerpo	-0.5753	6.37E-03	846.6486	Sabor	-0.5098	1.82E-02
792.6658	Calidad aromática	-0.5247	1.46E-02	850.5937	Acidez	-0.5244	1.47E-02
792.6708	Intensidad aromática	-0.6034	3.78E-03	850.5937	Agrio	-0.6280	2.30E-03
794.7285	Cuerpo	-0.5034	2.00E-02	850.5937	Amargor	0.6968	4.49E-04
799.5759	Intensidad aromática	-0.5358	1.23E-02	850.5937	Calidad aromática	0.5124	1.76E-02
802.7472	Agrio	-0.5108	1.80E-02	850.5937	Preferencia	0.5187	1.60E-02
803.4994	Acidez	-0.5830	5.55E-03	850.5937	Sabor	0.5462	1.04E-02
803.4994	Agrio	-0.6111	3.25E-03	850.7066	Intensidad aromática	-0.6073	3.50E-03
803.4994	Amargor	0.5746	6.44E-03	852.7605	Cuerpo	-0.5060	1.93E-02
803.4994	Preferencia	0.5778	6.08E-03	854.7416	Acidez	-0.7742	3.80E-05
805.6283	Agrio	-0.5296	1.35E-02	854.7416	Agrio	-0.6469	1.53E-03
805.6283	Amargor	0.5687	7.14E-03	854.7416	Amargor	0.7020	3.90E-04

854.7416	Cuerpo	0.5720	6.74E-03	884.9105	Calidad aromática	0.5306	1.33E-02
854.7416	Preferencia	0.6701	8.88E-04	884.9105	Cuerpo	0.5341	1.26E-02
854.7416	Sabor	0.6190	2.77E-03	884.9105	Preferencia	0.6662	9.76E-04
855.8055	Intensidad aromática	-0.5780	6.06E-03	884.9105	Resabio	0.6643	1.02E-03
856.1590	Acidez	0.5628	7.91E-03	884.9105	Sabor	0.5169	1.64E-02
856.1590	Cuerpo	-0.5217	1.53E-02	886.1028	Cuerpo	-0.5028	2.02E-02
856.4901	Acidez	-0.6370	1.90E-03	886.9233	Acidez	-0.8491	1.13E-06
856.4901	Agrio	-0.5485	1.00E-02	886.9233	Agrio	-0.7257	1.96E-04
856.4901	Amargor	0.5304	1.34E-02	886.9233	Amargor	0.6727	8.33E-04
857.0291	Acidez	-0.6278	2.31E-03	886.9233	Cuerpo	0.6830	6.44E-04
857.0291	Cuerpo	0.5250	1.45E-02	886.9233	Preferencia	0.7605	6.29E-05
864.6799	Intensidad aromática	-0.5390	1.17E-02	886.9233	Resabio	0.7086	3.24E-04
866.5994	Agrio	-0.5251	1.45E-02	886.9233	Sabor	0.5650	7.61E-03
866.7868	Cuerpo	-0.5498	9.82E-03	896.5138	Agrio	-0.5863	5.22E-03
868.6639	Astringencia	-0.5016	2.05E-02	896.9311	Acidez	0.5283	1.38E-02
868.6639	Cuerpo	-0.6203	2.70E-03	907.8284	Acidez	-0.5836	5.48E-03
868.6867	Acidez	-0.6656	9.91E-04	907.8284	Agrio	-0.5244	1.47E-02
868.6867	Agrio	-0.6020	3.89E-03	907.8284	Amargor	0.5538	9.20E-03
868.6867	Amargor	0.7189	2.41E-04	914.9457	Acidez	-0.6532	1.32E-03
868.6867	Intensidad aromática	-0.5624	7.96E-03	914.9457	Cuerpo	0.5119	1.77E-02
868.6867	Preferencia	0.5661	7.47E-03	924.7319	Cuerpo	-0.5668	7.38E-03
871.0105	Acidez	0.5771	6.16E-03	925.6957	Acidez	-0.6409	1.75E-03
871.6169	Acidez	-0.5120	1.76E-02	925.6957	Agrio	-0.6866	5.87E-04
871.6169	Cuerpo	0.5315	1.32E-02	925.6957	Amargor	0.6045	3.70E-03
876.5870	Astringencia	-0.5172	1.63E-02	925.6957	Preferencia	0.5551	9.00E-03
876.5870	Intensidad aromática	-0.5540	9.17E-03	927.5186	Acidez	-0.7131	2.85E-04
876.9652	Cuerpo	-0.5276	1.40E-02	927.5186	Agrio	-0.5889	4.97E-03
880.9984	Acidez	-0.6519	1.36E-03	927.5186	Amargor	0.6961	4.57E-04
880.9984	Agrio	-0.7590	6.64E-05	927.5186	Preferencia	0.6207	2.68E-03
880.9984	Amargor	0.7319	1.63E-04	927.5186	Resabio	0.5645	7.67E-03

880.9984	Preferencia	0.6974	4.41E-04	934.0657	Acidez	0.5290	1.37E-02
880.9984	Resabio	0.6024	3.86E-03	953.7383	Acidez	0.5101	1.82E-02
880.9984	Sabor	0.5715	6.80E-03	953.7383	Agrio	0.5225	1.51E-02
882.6911	Cuerpo	-0.6092	3.37E-03	954.7219	Cuerpo	-0.6092	3.37E-03
884.9105	Acidez	-0.6324	2.10E-03	955.6416	Acidez	0.5686	7.15E-03
884.9105	Agrio	-0.6358	1.95E-03	955.6416	Cuerpo	-0.5740	6.51E-03
884.9105	Amargor	0.5882	5.03E-03	963.4788	Acidez	0.6213	2.64E-03
963.4788	Agrio	0.6130	3.13E-03	1002.8096	Preferencia	-0.6942	4.81E-04
963.4788	Amargor	-0.7169	2.55E-04	1002.8096	Sabor	-0.6990	4.23E-04
963.4788	Calidad aromática	-0.5163	1.66E-02	1008.7261	Intensidad aromática	-0.5982	4.18E-03
963.4788	Preferencia	-0.6656	9.91E-04	1011.3676	Amargor	0.5674	7.30E-03
963.4788	Resabio	-0.6910	5.23E-04	1011.6594	Agrio	0.5023	2.03E-02
963.4788	Sabor	-0.5260	1.43E-02	1011.6594	Calidad aromática	-0.6287	2.27E-03
965.1912	Acidez	0.6955	4.64E-04	1011.6594	Preferencia	-0.5206	1.55E-02
965.1912	Agrio	0.6169	2.89E-03	1012.5358	Acidez	-0.7807	2.96E-05
965.1912	Amargor	-0.5622	7.98E-03	1012.5358	Agrio	-0.6222	2.60E-03
965.1912	Calidad aromática	-0.5650	7.61E-03	1012.5358	Amargor	0.6402	1.77E-03
965.1912	Cuerpo	-0.5185	1.60E-02	1012.5358	Preferencia	0.6201	2.71E-03
965.1912	Preferencia	-0.7059	3.50E-04	1012.5358	Resabio	0.5000	2.10E-02
965.1912	Resabio	-0.5639	7.76E-03	1019.4653	Cuerpo	-0.5557	8.91E-03
965.1912	Sabor	-0.5520	9.47E-03	1037.7176	Intensidad aromática	-0.5696	7.03E-03
966.3255	Amargor	-0.5713	6.82E-03	1038.2898	Acidez	0.6871	5.80E-04
971.5907	Cuerpo	-0.5374	1.20E-02	1038.2898	Amargor	-0.5447	1.07E-02
971.5907	Resabio	-0.5404	1.14E-02	1038.2898	Calidad aromática	-0.5403	1.15E-02
973.8260	Astringencia	0.5433	1.09E-02	1038.2898	Cuerpo	-0.7222	2.18E-04
986.6929	Cuerpo	-0.5844	5.40E-03	1038.2898	Preferencia	-0.6298	2.21E-03
992.6854	Sabor	0.5676	7.28E-03	1038.2898	Resabio	-0.5782	6.04E-03
992.7195	Acidez	-0.5875	5.10E-03	1038.8151	Acidez	0.5075	1.89E-02
992.7195	Agrio	-0.5700	6.98E-03	1038.8151	Cuerpo	-0.5309	1.33E-02
992.7195	Amargor	0.7267	1.91E-04	1039.6114	Acidez	0.5322	1.30E-02

992.7195	Preferencia	0.5057	1.94E-02	1039.6114	Calidad aromática	-0.5371	1.21E-02
992.7195	Resabio	0.5561	8.86E-03	1039.6114	Cuerpo	-0.5178	1.62E-02
994.4843	Intensidad aromática	-0.5332	1.28E-02	1040.8240	Acidez	0.5185	1.60E-02
995.5423	Acidez	-0.5413	1.13E-02	1040.8240	Cuerpo	-0.5446	1.07E-02
995.5423	Agrio	-0.5694	7.06E-03	1042.6635	Acidez	0.5537	9.22E-03
995.5423	Amargor	0.6474	1.51E-03	1065.6091	Astringencia	0.5777	6.09E-03
995.5423	Preferencia	0.5629	7.89E-03	1068.5749	Acidez	0.6233	2.54E-03
995.5423	Resabio	0.5919	4.70E-03	1068.5749	Cuerpo	-0.5975	4.24E-03
997.5751	Cuerpo	-0.5126	1.75E-02	1068.5829	Acidez	0.6402	1.77E-03
1000.7175	Amargor	0.5447	1.07E-02	1068.5829	Agrio	0.5935	4.57E-03
1002.8096	Acidez	0.5543	9.12E-03	1068.5829	Preferencia	-0.6747	7.94E-04
1002.8096	Agrio	0.5785	6.01E-03	1068.5829	Sabor	-0.5351	1.24E-02
1002.8096	Amargor	-0.5902	4.86E-03	1072.5519	Agrio	0.5694	7.06E-03
1002.8096	Calidad aromática	-0.5787	5.99E-03	1080.5541	Cuerpo	-0.6144	3.04E-03
1101.7143	Calidad aromática	-0.5527	9.37E-03	1269.8941	Astringencia	-0.5862	5.23E-03
1101.7143	Intensidad aromática	-0.8231	4.59E-06	1269.8941	Calidad aromática	-0.5501	9.78E-03
1109.3398	Intensidad aromática	-0.6691	9.12E-04	1288.7485	Agrio	-0.6293	2.24E-03
1118.2765	Acidez	0.5511	9.62E-03	1301.7150	Intensidad aromática	-0.5923	4.67E-03
1200.8903	Acidez	0.5693	7.07E-03	1303.7795	Agrio	0.5922	4.68E-03
1200.8903	Preferencia	-0.5343	1.26E-02	1306.0586	Acidez	0.6194	2.75E-03
1213.6684	Amargor	-0.5271	1.41E-02	1306.0586	Preferencia	-0.5083	1.86E-02
1213.6684	Calidad aromática	-0.5195	1.58E-02	1344.5560	Intensidad aromática	-0.5078	1.88E-02
1213.6684	Preferencia	-0.6025	3.84E-03	1515.9527	Acidez	0.7827	2.74E-05
1213.6684	Sabor	-0.5637	7.78E-03	1515.9527	Agrio	0.5036	1.99E-02
1233.6400	Cuerpo	-0.5204	1.56E-02	1515.9527	Amargor	-0.5109	1.79E-02
1241.6182	Intensidad aromática	-0.5403	1.15E-02	1515.9527	Cuerpo	-0.6726	8.37E-04
				1515.9527	Preferencia	-0.6129	3.13E-03

11.6 Lista de publicaciones y participaciones en congresos

Publicaciones:

- **Gamboa-Becerra, R.**, Montero-Vargas, J.M., Martínez-Jarquín, S., Gálvez-Ponce, E., Moreno-Pedraza, A. & Winkler, R. 2016. Rapid Classification of Coffee Products by Data Mining Models from Direct Electrospray and Plasma-Based Mass Spectrometry Analyses. *Food Anal. Methods* 1–10.
- Barraza, A., Cabrera-Ponce, J.L., **Gamboa-Becerra, R.**, Luna-Martínez, F., Winkler, R. & Álvarez-Venegas, R. 2015. The Phaseolus vulgaris PvTRX1h gene regulates plant hormone biosynthesis in embryogenic callus from common bean. *Frontiers in Plant Science* 6.
- **Gamboa-Becerra, R.**, Ramírez-Chávez, E., Molina-Torres, J. & Winkler, R. 2015. MSI.R scripts reveal volatile and semi-volatile features in low-temperature plasma mass spectrometry imaging (LTP-MSI) of chilli (*Capsicum annum*). *Analytical and Bioanalytical Chemistry*, doi: 10.1007/s00216-015-8744-9.

Congresos, estancias y cursos:

9-11 Nov 2016 – Cancún, Quintana Roo

▪ **Latin Food Congress, IAFP's 5th Latin American Symposium in Food Safety; 7th Food Science, Biotechnology and Safety Meeting**

Trabajo presentado: A novel metabolomics approach to predict the sensory profile of coffee beverage.

7-11 Dic 2015 –Querétaro, México

▪ **XVI National Congress of Plant Biochemistry and Molecular Biology & IX Symposium Mexico/USA**

Trabajo presentado: Early selection of coffee (*C. canephora*) plants with desirable cup quality traits by a novel metabolomics approach

7-11 Dic 2015 –Querétaro, México

▪ **XVI National Congress of Plant Biochemistry and Molecular Biology & IX Symposium Mexico/USA**

Trabajo presentado/coautor: Effect of overexpression of a developmental Regulator in the metabolite accumulation in *Arabidopsis Thaliana*

1-5 nov 2015 – Antalya, Turquía

▪ **2nd International Plant Breeding Congress & EUCARPIA-Oil and Protein Crops Section Conference**

Trabajo presentado: Early selection of coffee (*C. canephora*) with desirable cup quality traits by using heritable metabolic markers

3-5 jul 2015 – Ciudad de México

▪ **Convención internacional del café, México 2015**

Ponencia “Herramientas biotecnológicas para el sector agro-industrial del café”. Presentada en el panel de investigación y desarrollo.

2-8 nov 2014 – Guadalajara, Jalisco

▪ **XXX Congreso Nacional de Bioquímica**

Trabajo presentado: Un-targeted metabolic profiling for the Discovery of highly heritable metabolites linked to coffee (*C. canephora*) cup quality

4 mar-4 jun 2014 – Veracruz, Veracruz

▪ **Estancia: Laboratorio de tecnología del Café de la Unidad de Investigación y Desarrollo en Alimentos, Instituto Tecnológico de Veracruz**

Objetivo de la estancia: Entrenamiento en análisis sensorial descriptivo de la bebida de café mediante panel de catación. Análisis de muestras de proyecto de doctorado

Oct 2013 – Irapuato, Guanajuato

▪ **Scientific writing workshop, BioScience writers-CINVESTAV.**

Rapid Classification of Coffee Products by Data Mining Models from Direct Electrospray and Plasma-Based Mass Spectrometry Analyses

Roberto Gamboa-Becerra¹ · Josaphat Miguel Montero-Vargas¹ ·
Sandra Martínez-Jarquín¹ · Eligio Gálvez-Ponce² · Abigail Moreno-Pedraza¹ ·
Robert Winkler¹ 

Received: 12 July 2016 / Accepted: 13 October 2016
© Springer Science+Business Media New York 2016

Abstract In coffee manufacture, analytical methods with high-throughput and cost efficacy are required for process development and quality control. Thus, we investigated the applicability of direct mass spectrometry methods to distinguish coffee products according to species, geographic origin and processing. We tested the performance of the established method direct-injection electrospray mass spectrometry (DIESI-MS) and the emerging method low-temperature plasma ionization mass spectrometry (LTP-MS). Both methods are capable of classifying coffee products, but DIESI-MS and LTP-MS yield complementary information about the chemical composition of the samples. DIESI-MS shows a

broad molecular weight range of compounds. In contrast, LTP-MS detects mainly low molecular weight compounds, which correspond to quality-related ingredients, such as caffeine and purines. LTP-MS displays a high potential for rapid quality control measurements and online monitoring, because no sample processing is required. Data mining methods support the discovery of ‘important’ compounds, which are responsible for the discrimination between sample groups, and reveal associated chemical processes.

Keywords Low-temperature plasma (LTP) · Direct-injection electrospray (DIESI) · Mass spectrometry · Data mining · Coffee

Electronic supplementary material The online version of this article (doi:10.1007/s12161-016-0696-y) contains supplementary material, which is available to authorized users.

✉ Robert Winkler
robert.winkler@cinvestav.mx

Roberto Gamboa-Becerra
roberto.gamboa@cinvestav.mx

Josaphat Miguel Montero-Vargas
josaphat.montero@cinvestav.mx

Sandra Martínez-Jarquín
sandra.martinez@cinvestav.mx

Eligio Gálvez-Ponce
egalvez@ecomtrading.com

Abigail Moreno-Pedraza
abigail.moreno@cinvestav.mx

¹ Department of Biotechnology and Biochemistry, Cinvestav Unidad Irapuato, Km. 9.6 Libramiento Norte Carr. Irapuato-León, 36821 Irapuato, GTO, Mexico

² Agroindustrias Unidas de México S.A. de C.V. (AMSA), Bosque de Alisos 45A, Colonia Bosques de las Lomas, México, DF, Mexico

Introduction

Coffee is a highly appreciated beverage and, according to data of the Food and Agriculture Organization of the United Nations (FAO), is one of the most valuable agricultural products worldwide (FAOSTAT 2013). Currently, the species *Arabica* (*Coffea arabica*) and *Robusta* (*Coffea canephora*) are traded on the global market.

Coffee plantations require a tropical climate and a high degree of manual labour (Oestreich-Janzen 2010). Therefore, coffee is almost exclusively grown in developing countries, where about 20–25 million families in more than 50 countries are directly dependent on the production and exportation of coffee (Lewin et al. 2004).

Our research institute is located in Mexico, which is a typical representative for coffee-producing countries and their socio-economic problems. In Mexico, about 500,000 producers are cultivating farms with less than 10 ha (Jha et al. 2011). Considering the total production area of approx.

750,000 ha and the existence of more than 3000 producers with areas of 10–100 ha (Jha et al. 2011), the average farm size of smallholders can be estimated to be 1–1.5 ha. Usually, green coffee is exported to the consumer countries, where the grains are roasted and also where the production of instant (soluble) coffee takes place. As a consequence, in the global coffee value chain, the growers only obtain a small fraction of the total retail value. Therefore, even considering ‘fair trade’ prices, the family incomes from the sales of green coffee are insufficient for most of producers.

Currently, approximately 15–20 % of the final sales price goes to the local coffee growers and exporters. The benefits along the value chain can be enormous: the price of coffee may increase up to 11,000 % from the farm to the point of sale (Jha et al. 2011). Since the market is strongly controlled by transnational companies, smallholders have little influence and are vulnerable to other remote and local events, such as climate change (Ponte 2002; Eakin et al. 2009).

Local manufacture of specialty and certified coffees—Mexico e.g. is pioneer in the exportation of organic coffees—contributes to a more independent and sustainable production with respect to economy as well as social and cultural circumstances.

Therefore, in a current project related to the manufacture of soluble coffee in Mexico, we investigated the feasibility of different analytical strategies for process development and quality control. The ‘cup quality’ of the final product is dependent on the primary material and the unit operations of coffee processing. Appropriate chemical analyses support the evaluation of intermediary and final coffee products (Buffo and Cardelli-Freire 2004; Farah et al. 2005).

Direct-injection electrospray mass spectrometry (DIESI-MS) is an efficient method for the high-throughput analysis of biological samples. Therefore, DIESI-MS has been used for the metabolic phenotyping of plants, such as the characterisation of the physiological state of maize plants, and for classifying several thousand coffee trees (García-Flores et al. 2012; Montero-Vargas et al. 2013; García-Flores et al. 2015).

Contrary to other widely used methods in the analysis of plant materials, such as near infra-red spectroscopy (NIRS), the datasets obtained by DIESI-MS allow not only classification but also the quantification of specific molecules. Therefore, DIESI-MS permits e.g. the screening of a plant population for desired selection markers, such as high/low caffeine content in the case of coffee trees (Montero-Vargas et al. 2013). Compared to analytical standard techniques, such as high-pressure liquid chromatography (HPLC) or gas chromatography (GC), with their respective detectors, e.g. mass spectrometer (MS), diode array detector (DAD) or flame ionisation detector (FID), the sample preparation is simple and the cost per sample low. Consequently, DIESI-MS has been adopted for the evaluation of coffee products, e.g. for the detection of illegal admixtures of cheaper *Robusta* coffee

(*C. canephora*) and high-quality *Arabica* coffee (*C. arabica*) and for process-related and cup quality studies (Mendonça et al. 2008; Amorim et al. 2009; Garrett et al. 2012).

Low-temperature plasma ionisation mass spectrometry (LTP-MS) represents a novel strategy for the direct measurement of organic molecules from surfaces without any sample pre-treatment (Harper et al. 2008). Recently, we presented an improved LTP probe design, which enables the direct detection of organic molecules with different degrees of volatility (Martínez Jarquín and Winkler 2016; Martínez-Jarquín et al. 2016). Coupled to a quadrupole MS analyser, molecules related to coffee aroma and quality, such as vanillin, guaiacol, and caffeine, could be analysed by LTP-MS (Martínez-Jarquín and Winkler 2013).

Hence, we evaluated the potential of LTP-MS for the direct analysis of coffee products in comparison to DIESI-MS.

Materials and Methods

Coffee Samples

One hundred coffee samples were included: 95 roasted bean samples representing the two commercially important coffee species (28 of *C. canephora* and 67 of *C. arabica*) and 5 soluble coffee samples. The samples were of different geographical origins and post-harvest processing (dry and wet) (Table S1). We subjected roasted coffee beans and the lyophilised coffee extracts of those batches to both direct MS methods.

DIESI-MS Analysis of Methanolic Extracts

For the DIESI-MS analysis, we obtained methanolic extracts of 95 roasted coffee and 5 lyophilised samples, as described before (Montero-Vargas et al. 2013). Samples were analysed in triplicate. Mass spectra were acquired in ESI-positive mode, using a Micromass ZQ 2000 quadrupole instrument (Waters). For the measurements, the capillary voltage was set to 3 kV, the cone voltage to 60 kV, and the extractor voltage was 3 V. The source temperature and desolvation temperature were set to 80 and 150 °C, respectively. Spectra in continuous mode were acquired in an *m/z* range of 15–2000, with a run duration of 1 min, a scan time of 10 s and an inter scan time of 0.1 s.

Direct LTP-MS Analysis

For the direct analysis with LTP-MS, 500 mg of every sample was located randomly in a 96-well plate. Extraction of metabolites was not necessary, but 500 µL of methanol was added to every sample to moisten a few seconds before analysis. The measurements were performed in triplicate. The parameters

employed for the MS were the same as mentioned above, with the exception of cone voltage and nitrogen flow which were adjusted to 0 V and 0 L/h, respectively. The sample was directly exposed to the plasma beam, and spectra were collected as mentioned above.

Data Analysis

Spectra were processed using a TOPPAS pipeline (OpenMS TOPP tool) (Sturm et al. 2008) as reported previously (Montero-Vargas et al. 2013) and applying an m/z binning with 0.8- m/z intervals for DIESI-MS and LTP-MS, considering the mass tolerance of the mass detector (0.4 m/z). We compared the chemical profiles using an unsupervised hierarchical clustering analysis (HCA) approach and generating metabolic heatmaps with the ion intensities to evaluate the discriminating power of fingerprint data. In order to determine the contribution of every ion, we used Rattle (Graphical User Interface, R package) (Williams 2011) for the generation of Random Forest models. The Random Forest algorithm consisted of training, validation and testing steps employing m/z features (DIESI and LTP-MS data). In all comparisons, 70 % of samples were used to train the model, 15 % were used for the validation step and the testing was done using 15 % of total samples. Five hundred decision trees were built to finally obtain a predictive model which was able to classify samples dependent on species, geographical origin or coffee processing. The variable (ion) importance was determined by a measure of the mean decrease in the Gini index; ions with higher values play the most important role in predicting correct classification of coffee samples. For graphical box plot generation, we employed RStudio packages.

Putative metabolite identification was conducted employing SpiderMass (Winkler 2015a). A coffee metabolite database was automatically generated with a list of coffee compounds previously reported in literature (Farah et al. 2005; Mendonça et al. 2008; Amorim et al. 2009; Alonso-Salces et al. 2009; Garrett et al. 2012; Weschenfelder et al. 2015; Moreno et al. 2015) and the CoffeaCyc database (<http://archive.gramene.org/pathway/coffeacyc.html>). Subsequently, matching with the DIESI and LTP-MS peak lists was performed. The complete workflow can be reproduced using the MASSyPup64 platform (<http://lababi.bioprocess.org/index.php/lababi-software/78-massypup>) (Winkler 2015b).

Results and Discussion

Comparison of DIESI-MS and LTP-MS Fingerprinting

We evaluated the DIESI-MS fingerprints of 95 roasted coffee samples (IDs M1–M95) and 5 lyophilised samples (IDs M96–

M100). In addition, direct analysis with LTP-MS was conducted for samples M91–M95 and their respective lyophilised batches (M96–M100). All samples were analysed in triplicate. Figure 1a–d shows characteristic spectra of lyophilised and roasted coffee extracts analysed by DIESI and LTP-MS. The spectra in positive mode of DIESI-MS for roasted coffee samples (Fig. 1a) display a wide range of ions up to a mass-to-charge ratio of about 1000, but the strongest signals were detected below m/z 250. The most intense ion detected was m/z 94.254, which likely corresponds to methyl pyridinium $[M+H]^+$, a trigonelline degradation product generated during coffee roasting (Amorim et al. 2009). The second most abundant ion was m/z 138.136, which is associated with protonated trigonelline (Garrett et al. 2012). In addition, an ion with very low mass-to-charge ratio (m/z 39.616) was detected. For lyophilised samples, DIESI-MS spectra in positive mode show a smaller m/z range around m/z 600 in comparison with roasted coffee. Similar to roasted coffee, we found m/z 138.136 and m/z 94.254 as the two most intense signals. In addition to roasted coffee spectra, we observed ions such as m/z 53.509, m/z 78.327, m/z 381.136 and m/z 543.166.

In the LTP-MS spectra of roasted samples, we observed masses around m/z 200 which are less abundant compared with data obtained by DIESI-MS. The most intense ions detected were m/z 60.157, m/z 80.022, m/z 95.918, m/z 109.895, m/z 135.907 and m/z 194.919. The last one corresponds to protonated caffeine.

Contrary to roasted beans, in lyophilised samples, we observed additional higher masses between m/z 200 and 800, such as m/z 429.002, m/z 445.002, m/z 683.990 and m/z 757.954.

These distinct mass ranges from the DIESI-MS and LTP-MS data each cover a wide mass range detecting volatile and semi-volatile compounds, which are important for the aroma of coffee. Molecules which confer taste, body and colour are found in the higher mass range. Therefore, LTP-MS and DIESI-MS are complementary for the measurement of quality-related chemical profiles.

Using ion intensities of the DIESI-MS data, we generated a metabolic heatmap with the 40 strongest signals. In general, the HCA displays a grouping mainly with respect to the species (*C. canephora* and *C. arabica*) and lyophilisation process (Fig. S1). Samples M96–M100 belong to lyophilised coffee and were perfectly discriminated in a separate branch from the rest of the roasted coffee samples. Roasted coffee samples of *C. canephora* (M1–M24 and M91–M94) clustered together in a large group (with exception of samples M10 and M16) while in a third branch of the dendrogram only *C. arabica* samples were grouped (Fig. S1). Hence, an HCA approach with DIESI-MS metabolic fingerprinting data is suitable for differentiating at level of species and coffee processing (roasted or lyophilised coffee).

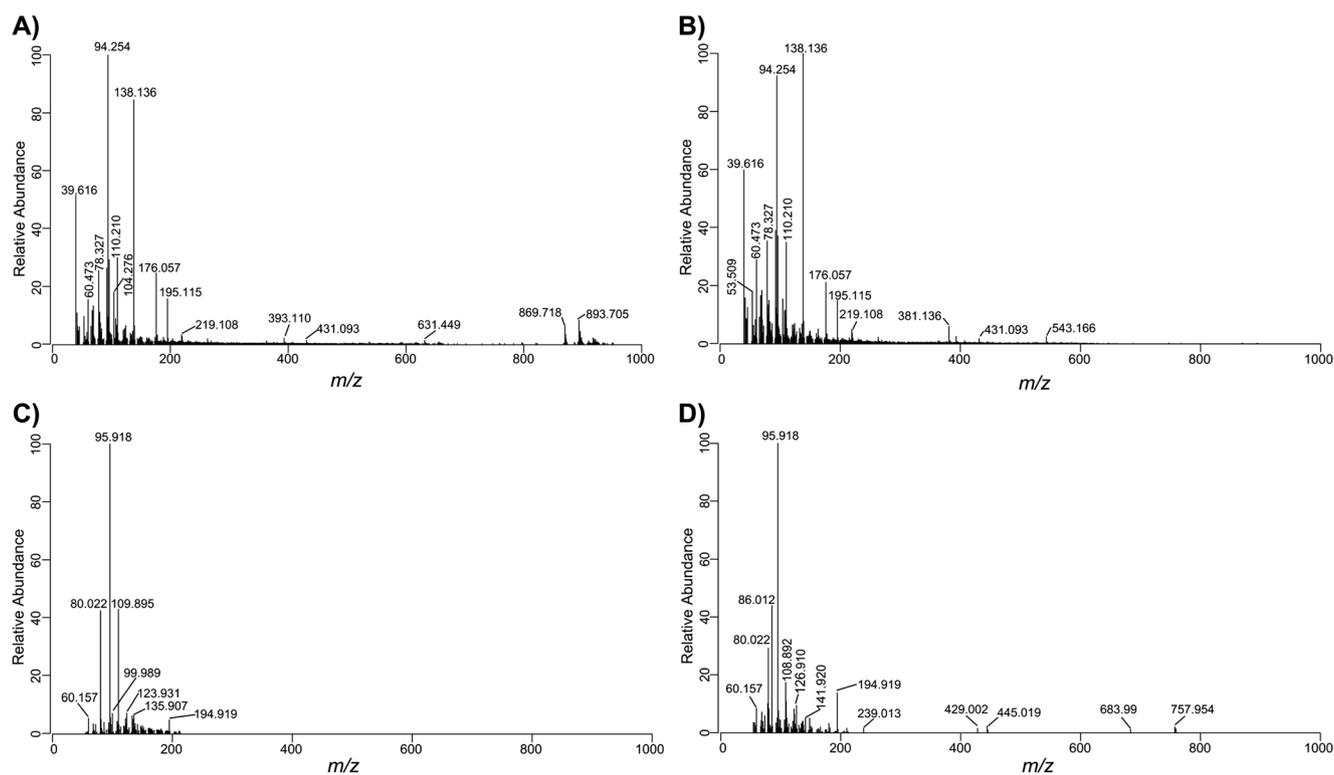


Fig. 1 Characteristic master spectra of roasted and lyophilised coffee extracts analysed by DIESI (a, b) and LTP-MS (c, d). Figures show the spectra of ions detected in positive mode

Creation of Random Forest Model for Differentiating Between Species and Coffee Processing

In order to determine the most important ions that distinguish between species and coffee processing by both DIESI and LTP-MS data, we generated a Random Forest model with 500 tree decisions, which delivers the 30 most important ions for every comparison (Figs. 2a–c and 4a, b). Ions with bigger values of a mean decrease in Gini are better at discriminating between coffee classes. In addition, to determine the putative identity of these ions, first we summed up the spectra of all samples analysed either by DIESI-MS or by LTP-MS to generate a master spectrum and extract a list of average masses. Then, we conducted a matching with a custom coffee database, generated in SpiderMass (Winkler 2015a). Tables 1 and 2 list putative compounds for DIESI-MS and LTP-MS, respectively.

Classification of Roasted *C. arabica* and *C. canephora* Beans by Species

Several methods have been proposed to identify metabolite markers distinguishing between coffee species and varieties (Martín et al. 1998; Rogers et al. 1999; Pizarro et al. 2007; Wei et al. 2012; Nascimento et al. 2015). Here, we analysed the methanolic DIESI-MS fingerprint of *C. arabica* and *C. canephora* coffee samples to establish metabolite markers

specific of the species. Figure 2a shows the 30 most important ions defining if a coffee-roasted sample corresponds to *C. arabica* or *C. canephora* species. The predictive Random Forest model was able to identify both species with an error rate of 1.01 % applying the model to a testing sample set (which was not included in the model building). Between the most important identified ions, we found three compounds which belong to fatty acids: m/z 101.217 (tiglic acid), m/z 169.121 (mesaconic acid) and m/z 279.232 (linoleic acid). The two last ions were detected at higher levels in *C. canephora* compared with *C. arabica* samples (Table S2). Linoleic acid was previously reported as one of the most important fatty acids to discriminate between coffee varieties (Martín et al. 2001; Romano et al. 2014). Compounds corresponding to carboxylic acids such as m/z 77.301 (glycolic acid), m/z 113.176 (oxalic acid) and m/z 127.140 (malonic acid) were detected as well. m/z 521.216 (atractyloside analogue II) and m/z 767.304 (atractyloside analogue I) were detected with stronger signal intensities in *C. arabica* in comparison to *C. canephora* (Table S2). Previously, these two diterpene glycoside compounds have been proposed as important compounds for quantification of *Arabica-Robusta* blends (Garrett et al. 2012).

The coffee roasting process leads to trigonelline degradation. Some of its degradation products were detected, including m/z 92.255 (1-pyrroline) and m/z 131.165 (1,4-dimethylpyridinium). Other compounds with high abundance

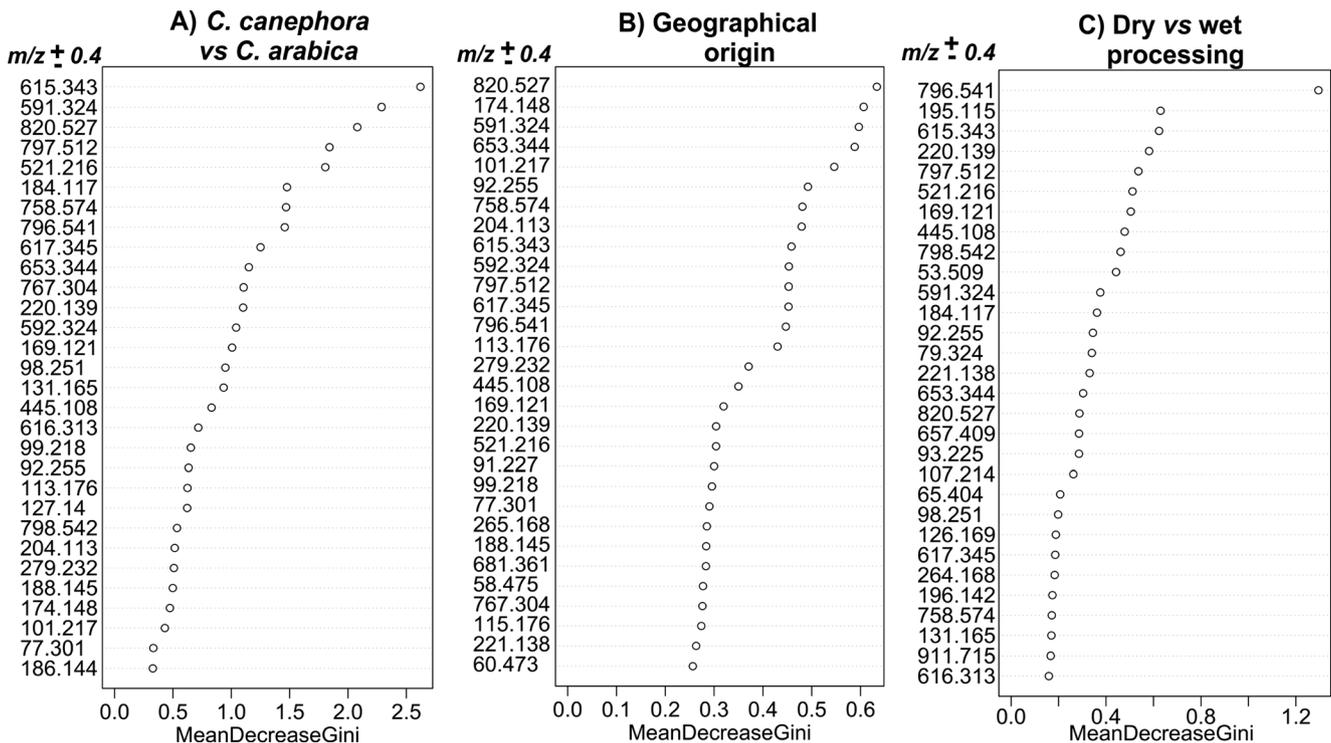


Fig. 2 Variable (ion) importance ranking for differentiation between coffee sample processes with DIESI-MS data. Thirty most important ions are shown for discriminating between **a** *C. canephora* and *C. arabica* species, **b** geographical origin and **c** wet and dry post-harvest processing

in *C. arabica* and contributing to species differentiation are listed in Table S2.

Metabolic fingerprinting data of *C. canephora* and *C. arabica* species, obtained by DIESI-MS, therefore are sufficient to correctly classify coffee samples with high accuracy, employing a Random Forest model.

Classification of Roasted Beans According to Geographical Origin

To evaluate the capability of DIESI-MS to discriminate roasted coffee samples having different geographical origins, including four states in Mexico (Chiapas, Oaxaca, Veracruz, Puebla) and Vietnam, we generated Random Forest models. In total, 61 samples were included: 36 from Chiapas, 8 from Oaxaca, 8 from Veracruz, 8 from Puebla and 1 from Vietnam. The results obtained showed a classification error rate of 19.4 % with the final predictive model obtained. But applying this model to a testing sample group, 94 % of samples were assigned correctly. Figure 2b shows the 30 most important ions differentiating between different geographical origins. Of all tested ions for differentiating coffee processing or species, six ions were exclusive to separation of samples with different geographical origin: *m/z* 60.473 (trimethylamine), *m/z* 91.227 (lactic acid), *m/z* 115.176 (glycolic acid), *m/z* 265.168 (pentadecanoic acid) and ions *m/z* 681.361 and *m/z* 758.574. Some researchers have proposed the use of individual metabolites and compound classes to identify

geographical origin of coffee samples, including GABA, acetic acid, trigonelline, caffeine, choline, sucrose, free aminoacids, chlorogenic acids and methylxanthines (Parras et al. 2007; Alonso-Salces et al. 2009; Choi et al. 2010; Wei et al. 2012). But employing a set of metabolites demonstrates an increased discriminative power.

Although the sample set only included 61 samples from Mexico and Vietnam, our results indicate the possibility to use direct MS in combination with data mining to determine the geographical origin of roasted coffee samples and coffee blends and to identify relevant metabolic markers.

Discrimination of Roasted Beans From Wet and Dry Processing

After the harvest, cherry coffee fruits are processed using dry, semi-dry or wet methods for removing the pulp and hull. The green beans are then roasted and ground to obtain the coffee product. Post-harvest processing of coffee beans has been shown to have a great impact on coffee beverage quality (Tarzia et al. 2010). Thus, we employed DIESI-MS fingerprinting on roasted coffee beans to find the ions that indicate the chemical differences between wet and dry post-harvest processing (Fig. 2c). The Random Forest model generated permitted us to classify coffee samples from wet or dry processing with an error rate of 5.29 % in a testing dataset. The most important ions distinguishing coffee processing represent several compound classes, as well as chemical processes

Table 1 The most important compounds detected in coffee samples by DIESI-MS differentiating between species, geographical origin and coffee processing, putatively identified by a custom SpiderMass database

<i>m/z</i>	Monoisotopic mass (Da)	Ionization mode	Compound name
60.473	59.073498	[M+H] ⁺	Trimethylamine
77.301	76.016045	[M+H] ⁺	Glycolic acid
79.324	79.042198	+/- 0	Pyridine
91.227	90.031693	[M+H] ⁺	Lactic acid
92.255	69.057846	[M+Na] ⁺	1-Pyrroline
98.251	59.073498	[M+K] ⁺	Trimethylamine
99.218	76.016045	[M+Na] ⁺	Glycolic acid
101.217	100.052429	[M+H] ⁺	Tiglic acid
113.176	89.995308	[M+Na] ⁺	Oxalic acid
115.176	76.016045	[M+K] ⁺	Glycolic acid
127.14	104.010956	[M+Na] ⁺	Malonic acid
131.165	108.080772	[M+Na] ⁺	1,4-Dimethylpyridinium
169.121	130.026611	[M+K] ⁺	Mesaconic acid
174.148	173.045547	[M+H] ⁺	Shikimate
184.117	161.060806	[M+Na] ⁺	Methyl cinnamate
195.115	194.080383	[M+H] ⁺	Caffeine
196.142	173.045547	[M+Na] ⁺	Shikimate
220.139	219.112015	[M+H] ⁺	Zeatin
221.138	198.052826	[M+Na] ⁺	Feruloyl-caffeoylquinic acid
264.168	264.208923	+/- 0	Farnesyl acetate
265.168	242.224579	[M+Na] ⁺	Pentadecanoic acid
279.232	278.224579	[M+H] ⁺	Linoleic acid
281.202	280.240234	[M+H] ⁺	Linoleic acid
284.203	284.271515	+/- 0	Stearic acid
324.243	324.024628	+/- 0	5-Enolpyruvylshikimate-3-phosphate
381.136	342.116211	[M+K] ⁺	sucrose
521.216	498.116211	[M+Na] ⁺	3,4-Dicaffeoylquinic-1,5-lactone
537.317	498.116211	[M+K] ⁺	3,4-Dicaffeoylquinic-1,5-lactone
553.264	530.142456	[M+Na] ⁺	4-Caffeoyl-5-feruloylquinic acid
579.224	540.469482	[M+K] ⁺	Zeta-carotene
591.324	568.42804	[M+Na] ⁺	Zeaxanthin
595.234	594.137329	[M+H] ⁺	Cyanidin 3-(p-coumaroyl)-glucoside
617.345	578.142456	[M+K] ⁺	Procyanidin B-3
796.541	757.197449	[M+K] ⁺	Pelargonidin 3-O-(6-caffeoyl-β-D-glucoside) 5-O-β-D-glucoside
798.542	759.57782	[M+K] ⁺	1-16:0-2-18:1-Phosphatidylcholine
892.642	892.535339	+/- 0	Chlorophyll
893.705	870.565918	[M+Na] ⁺	Pheophytin a
909.713	908.529114	[M+H] ⁺	Chlorophyll b
915.716	892.535339	[M+Na] ⁺	a-Chlorophyll
927.756	926.631836	[M+H] ⁺	All-trans-undecaprenyl diphosphate
935.729	897.157043	[M+K] ⁺	(E)-cinnamoyl-CoA
949.768	926.631836	[M+Na] ⁺	All-trans-undecaprenyl diphosphate
Unknown [<i>m/z</i>]	53.509, 126.169, 186.144, 188.145, 204.113, 376.194, 445.108, 509.359, 2.324, 615.343, 616.313, 619.317, 647.339, 653.344, 655.437, 657.409, 671.353, 681.361, 758.574, 767.304, 797.512, 820.527, 869.718, 885.699, 907.682, 910.716, 911.715, 919.723, 922.754		

Table 2 The 30 most important compounds detected in coffee samples by LTP-MS differentiating between roasted and lyophilised coffee, putatively identified by a custom SpiderMass database

<i>m/z</i>	Monoisotopic mass (Da)	Ionization mode	Compound name
60.157	60.02113	+/- 0	Acetic acid
68.112	67.042198	[M+H] ⁺	Pyrrrole
80.996	80.037445	[M+H] ⁺	Pyrazine
93.970	94.053101	+/- 0	Methylpyrazine
95.918	96.021126	+/- 0	Furfural
96.947	74.073166	[M+Na] ⁺	2-Butanol
107.946	108.080772	+/- 0	1,4-Dimethylpyridinium
109.895	110.036781	+/- 0	Catechol
127.943	128.047348	+/- 0	Furaneol
133.929	134.02153	+/- 0	Malic acid
135.907	136.100052	+/- 0	2-Ethyl-3,5-dimethylpyrazine
137.913	137.047684	[M+H] ⁺	Trigonelline
143.928	144.115036	+/- 0	Caprylic acid
151.949	152.047348	+/- 0	Vanillin
153.953	154.135757	+/- 0	A-Terpineol
159.912	137.047684	[M+Na] ⁺	Trigonelline
163.924	163.975494	+/- 0	Methyl trifluoromethanesulfonate
165.929	166.049072	+/- 0	7-Methylxanthine
174.924	152.083725	[M+Na] ⁺	p-Ethylguaiacol
178.936	178.078247	[M+H] ⁺	Phenanthrene
192.944	170.130676	[M+Na] ⁺	cis-Furan linalool oxide
193.915	194.057907	+/- 0	Ferulic acid
Unknown [<i>m/z</i>]	70.116, 181.942, 205.923, 239.013, 445.019, 683.99, 757.954, 759.960		

occurring during coffee roasting. Trigonelline degradation was shown by compounds such as *m/z* 79.324 (pyridine), *m/z* 92.255 (1-pyrroline) and *m/z* 131.165 (1,4-dimethylpyridinium). 1,4-Dimethylpyridinium was described in a study that investigated the changes in levels of furans, ketones, pyrazines, acids, pyridines and pyrroles (Gonzalez-Rios et al. 2007). We also identified compounds important in chlorogenic acid formation (*m/z* 184.117, methyl cinnamate; *m/z* 196.142, shikimate) and lactone (*m/z* 521.216, 3,4-dicaffeoylquinic-1,5-lactone), all of which result from modifications of chlorogenic acids during roasting (Rakesh Jaiswal 2014).

Other ions which were present in higher levels in samples with wet processing compared to dry processing were *m/z* 591.324 (zeaxanthin), *m/z* 617.345 (procyanidin B-3), *m/z* 796.541 (pelargonidin 3-O-(6-caffeoyl-β-D-glucoside)5-O-β-D-glucoside) and *m/z* 798.542 (1-16:0-2-18:1-phosphatidylcholine). In contrast, *m/z* 221.138 (3,5-dimethyl-4-hydroxybenzoic acid), *m/z* 195.115 (caffeine), *m/z* 169.121 (mesaconic acid) and *m/z* 220.139 (zeatin) were found in more abundance in samples from dry processing. Previous reports have proposed ions from fatty acids, chlorogenic acids, sugars and carboxylic acids as being the most important to separate between dry, wet and semi-wet post-harvest processing, employing principal component analysis (PCA) (Amorim et al. 2009). In contrast, predictive data mining models

correctly classify samples according to wet or dry post-harvest processing from MS fingerprinting data, but also indicate relevant compounds and reactions in the coffee roasting process.

Rapid Classification of Roasted and Lyophilised Coffee Products

The lyophilisation process involves the treatment of ground roasted coffee with hot water, cooling, centrifugation, concentration by heating and drying through freeze-drying, until only approximately 5 % of moisture is left. During this process, only water-soluble compounds are extracted; the lipophilic fraction and volatile compounds are reduced in the final coffee product (Farah 2012).

Roasted beans and lyophilised coffee were available for five batches (IDs 91–100). These samples were analysed using DIESI-MS and LTP-MS. Both methods clearly discriminated between the two different products. Moreover, triplicate samples clustered together with one exception, which demonstrates the high repeatability of both methods (Fig. 3a, b).

The Random Forest model, generated with DIESI-MS data, perfectly (0 % error rate) discriminated between roasted and lyophilised coffee (Fig. 4a) when the model was applied to the testing dataset. Sixteen of the 30 most important ions

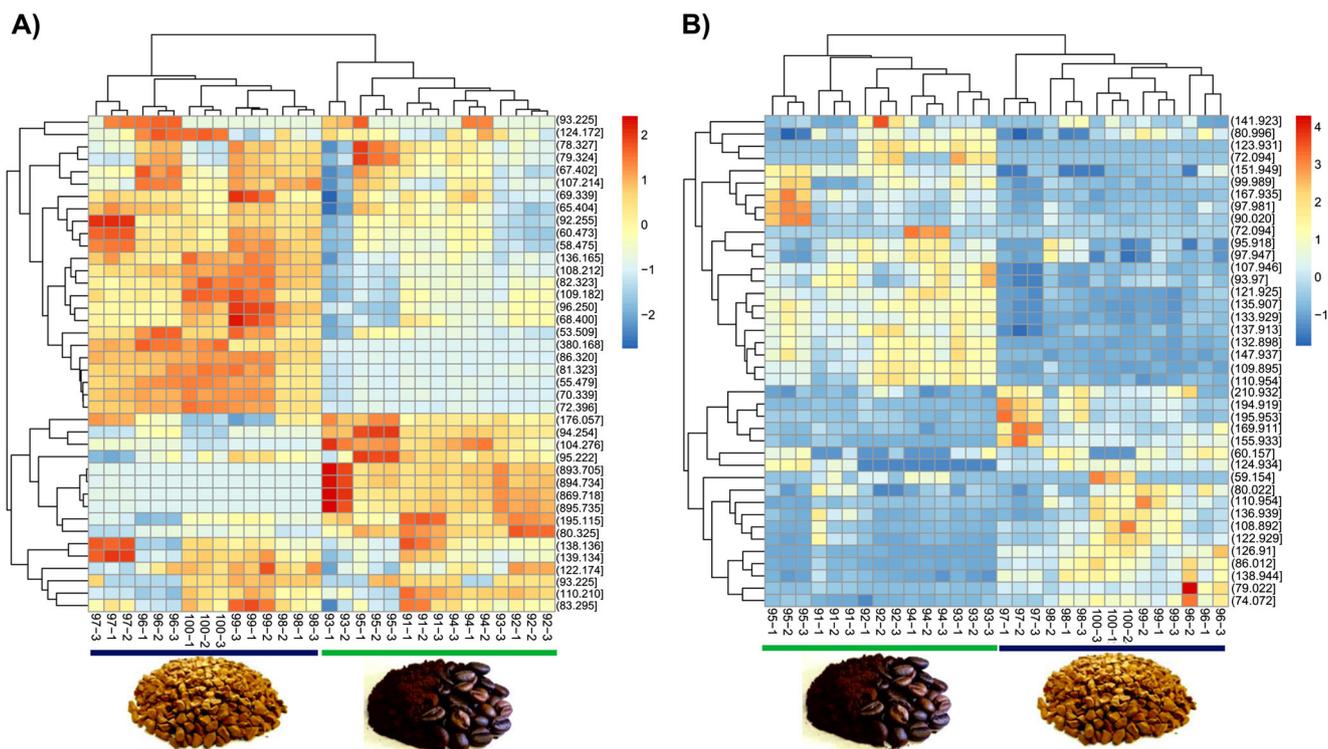


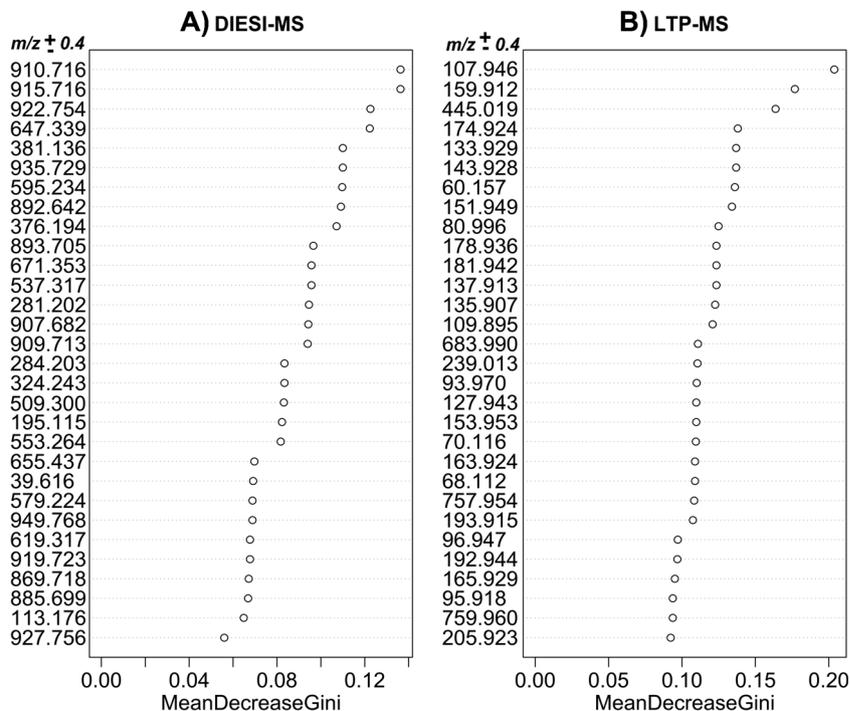
Fig. 3 Metabolic heatmap generated with the 40 most intense ions detected by DIESI-MS (a) and LTP-MS (b). Ten samples, corresponding to five roasted coffee samples (91–95) and five lyophilised coffee samples

(96–100) with their respective technical triplicates, were included. All triplicates cluster together. The main branches reflect the processing state

that indicate difference in processing were present in roasted, but not in lyophilised samples (Fig. S2). These ions were mainly signals with a mass-to-charge ratio above m/z 600.

On the other hand, some ions presented higher signal intensity in roasted coffee, compared to lyophilised samples e.g. m/z 39.616, m/z 195.115 (caffeine), m/z 281.202 (linoleic

Fig. 4 Variable (ion) importance ranking for differentiation between roasted and lyophilised coffee samples with DIESI-MS (a) and LTP-MS (b) data. The 30 most important ions are shown



acid) and m/z 284.203 (stearic acid). Since little sample preparation is necessary before measurement, DIESI-MS represents a suitable mass spectrometry technique for rapid analysis and classification of samples. Nevertheless, new strategies which require no sample preparation still would be advantageous.

Using ionisation/desorption with LTP, it is possible to cover a large range of compounds, including volatiles, which correspond to the mainly observed ions of roasted and lyophilised coffee. The predictive Random Forest model generated with LTP-MS data correctly distinguished (error rate of 0 %) between roasted or lyophilised coffee samples in the testing sample set. Between the 30 most important ions that contribute to the classification, we found 17 ions presenting higher abundance in roasted coffee in comparison with lyophilised coffee (Fig. S3). Seven ions (m/z 93.97, 107.946, 109.895, 133.929, 135.907, 143.928 and 159.912) were only detected in roasted coffee by LTP-MS, but not in lyophilised samples (Fig. S3). In contrast, ions m/z 239.013 and higher mass-to-charge ratio such as m/z 445.019, m/z 683.990, m/z 757.954 and m/z 759.960 were not found in roasted samples (Fig. S3). This is consistent with the fact that during the lyophilisation process compound transformations occur, as well as qualitative and quantitative changes of several volatiles, which in turn influence the sensory quality of soluble coffee (Flament and Bessi re-Thomas 2002). In fact, several methodologies focussed on retention of aroma quality-related compounds have been reported (Baggenstoss et al. 2010; Jeszka-Skowron et al. 2014; Padma Ishwarya and Anandharamakrishnan 2015; Weschenfelder et al. 2015; Freiberger et al. 2015; Moreno et al. 2015).

Roasted and soluble coffee could be easily distinguished by organoleptic inspection. But since the sample preparation for LTP-MS is negligible, and data are delivered in a few seconds, the method demonstrates a potential in process development and control, for optimizing the aroma and flavour of products, and in the quality control of soluble coffee products, such as proving the origin of raw materials or the authenticity of a producer.

Conclusions

Applying data mining strategies to direct mass spectrometry (MS) data of coffee products resulted in predictive models for reliable classification according to species, geographic origin and processing. Although our sample set was limited with respect to the number of included geographic regions, final coffee products, etc., our results demonstrated the impressive discriminative power of data mining models from direct-injection electrospray (DIESI) and low-temperature plasma (LTP) ionisation mass spectrometry (MS) data. LTP-MS is able to trace volatile compounds, which are important for the aroma of the final product, without sample pretreatment.

Therefore, LTP-MS shows high potential for the rapid and cost-efficient analysis of coffee products in process monitoring and quality control. Analysing chemical profile data with a Random Forest strategy delivers testable classification models, as well as important compounds, and therefore supports the study of process reactions.

Compliance with Ethical Standards

Funding This study was funded by Secretar a de Agricultura, Ganader a, Desarrollo Rural, Pesca y Alimentaci n (SAGARPA) and the Consejo Nacional de Ciencia y Tecnolog a (CONACyT), Mexico, with grants AVANCE N  127280, INNOVATEC PEI 451/2012, FIT S0017-2015-01, I0017/CB2010-01/151596, FINNOVA I010/260/2014 and FRONTERAS 2015-2/814. RGB, JMMV, SMJ and AMP thank CONACyT for their postgraduate fellowships.

Conflict of Interest SMJ and RW are inventors of the patent application ‘Non-thermal plasma jet device as source of spatial ionization for ambient mass spectrometry and method of application’ (WO 2014/057409, United States Patent 9362100). The micromass ZQ was borrowed from Waters, Mexico. Agroindustrias Unidas de Mexico SA de CV and Garcomex SA de CV provided samples. The sponsors were not involved in the study design, collection, analysis or interpretation of data.

Informed Consent Not applicable.

References

- Alonso-Salces RM, Serra F, Reniero F, K  H (2009) Botanical and geographical characterization of green coffee (*Coffea arabica* and *Coffea canephora*): chemometric evaluation of phenolic and methylxanthine contents. *J Agric Food Chem* 57:4224–4235. doi:10.1021/jf8037117
- Amorim ACL, Hovell AMC, Pinto AC et al (2009) Green and roasted arabica coffees differentiated by ripeness, process and cup quality via electrospray ionization mass spectrometry fingerprinting. *J Braz Chem Soc* 20:313–321. doi:10.1590/S0103-50532009000200017
- Baggenstoss J, Thomann D, Perren R, Escher F (2010) Aroma recovery from roasted coffee by wet grinding. *J Food Sci* 75:C697–C702. doi:10.1111/j.1750-3841.2010.01822.x
- Buffo RA, Cardelli-Freire C (2004) Coffee flavour: an overview. *Flavour Fragr J* 19:99–104. doi:10.1002/ffj.1325
- Choi M-Y, Choi W, Park JH et al (2010) Determination of coffee origins by integrated metabolomic approach of combining multiple analytical data. *Food Chem* 121:1260–1268. doi:10.1016/j.foodchem.2010.01.035
- Eakin H, Winkels A, Sendzimir J (2009) Nested vulnerability: exploring cross-scale linkages and vulnerability teleconnections in Mexican and Vietnamese coffee systems. *Environ Sci Pol* 12:398–412
- FAOSTAT (2013) <http://faostat.fao.org/>
- Farah A (2012) Coffee constituents. In: Chu Y-F (ed) *Coffee*. Wiley, Hoboken, pp. 21–58
- Farah A, de Paulis T, Trugo LC, Martin PR (2005) Effect of roasting on the formation of chlorogenic acid lactones in coffee. *J Agric Food Chem* 53:1505–1513. doi:10.1021/jf048701t
- Flament I, Bessi re-Thomas Y (2002) *Coffee flavor chemistry*. Wiley, Chichester; New York
- Freiberger EB, Kaufmann KC, Bona E et al (2015) Encapsulation of roasted coffee oil in biocompatible nanoparticles. *LWT-Food Sci Technol* 64:381–389. doi:10.1016/j.lwt.2015.05.039

- García-Flores M, Juárez-Colunga S, García-Casarrubias A et al (2015) Metabolic profiling of plant extracts using direct-injection electrospray ionization mass spectrometry allows for high-throughput phenotypic characterization according to genetic and environmental effects. *J Agric Food Chem* 63:1042–1052. doi:10.1021/jf504853w
- García-Flores M, Juárez-Colunga S, Montero-Vargas JM et al (2012) Evaluating the physiological state of maize (*Zea mays* L.) plants by direct-injection electrospray mass spectrometry (DIESI-MS). *Mol BioSyst* 8:1658–1660. doi:10.1039/c2mb25056j
- Garrett R, Vaz BG, Hovell AMC et al (2012) Arabica and Robusta coffees: identification of major polar compounds and quantification of blends by direct-infusion electrospray ionization–mass spectrometry. *J Agric Food Chem* 60:4253–4258. doi:10.1021/jf300388m
- Gonzalez-Rios O, Suarez-Quiroz ML, Boulanger R et al (2007) Impact of “ecological” post-harvest processing on coffee aroma: II. Roasted coffee. *J Food Compos Anal* 20:297–307. doi:10.1016/j.jfca.2006.12.004
- Harper JD, Charipar NA, Mulligan CC et al (2008) Low-temperature plasma probe for ambient desorption ionization. *Anal Chem* 80:9097–9104
- Jeszka-Skowron M, Zgoła-Grześkowiak A, Grześkowiak T (2014) Analytical methods applied for the characterization and the determination of bioactive compounds in coffee. *Eur Food Res Technol* 240:19–31. doi:10.1007/s00217-014-2356-z
- Jha S, Bacon CM, Philpott SM, et al (2011) A review of ecosystem services, farmer livelihoods, and value chains in shade coffee agroecosystems. In: Integrating agriculture, conservation and ecotourism: examples from the field. Springer, pp 141–208
- Lewin B, Giovannucci D, Varangis P (2004) Coffee markets: new paradigms in global supply and demand.
- Martínez Jarquín S, Winkler R (2016) Non-thermal plasma jet device as source of spatial ionization for ambient mass spectrometry and method of application. United States Patent 9362100
- Martínez-Jarquín S, Moreno-Pedraza A, Guillén-Alonso H, Winkler R (2016) Template for 3D printing a low-temperature plasma probe. *Anal Chem*. doi:10.1021/acs.analchem.6b01019
- Martínez-Jarquín S, Winkler R (2013) Design of a low-temperature plasma (LTP) probe with adjustable output temperature and variable beam diameter for the direct detection of organic molecules. *Rapid Commun Mass Spectrom* 27:629–634. doi:10.1002/rcm.6494
- Martín MJ, Pablos F, González AG (1998) Discrimination between arabica and robusta green coffee varieties according to their chemical composition. *Talanta* 46:1259–1264. doi:10.1016/S0039-9140(97)00409-8
- Martín MJ, Pablos F, González AG et al (2001) Fatty acid profiles as discriminant parameters for coffee varieties differentiation. *Talanta* 54:291–297. doi:10.1016/S0039-9140(00)00647-0
- Mendonça JCF, Franca AS, Oliveira LS, Nunes M (2008) Chemical characterisation of non-defective and defective green arabica and robusta coffees by electrospray ionization-mass spectrometry (ESI-MS). *Food Chem* 111:490–497. doi:10.1016/j.foodchem.2008.03.080
- Montero-Vargas JM, González-González LH, Gálvez-Ponce E et al (2013) Metabolic phenotyping for the classification of coffee trees and the exploration of selection markers. *Mol BioSyst* 9:693–699. doi:10.1039/C3MB25509C
- Moreno FL, Quintanilla-Carvajal MX, Sotelo LI et al (2015) Volatile compounds, sensory quality and ice morphology in falling-film and block freeze concentration of coffee extract. *J Food Eng* 166: 64–71. doi:10.1016/j.jfoodeng.2015.05.018
- Nascimento EA, Morais SA, Chang R, et al (2015) Looking for unknown molecules in Arabica and Robusta coffee brews by high performance liquid chromatography coupled to high resolution mass spectrometry and differentiating both coffee species by this technique.
- Oestreich-Janzen S (2010) Chemistry of coffee In: Comprehensive natural products II: Chemistry and Biology: 10 Volume Set. In: Mander L, Liu H-W (eds). Elsevier Science
- Padma Ishwarya S, Anandharamakrishnan C (2015) Spray-freeze-drying approach for soluble coffee processing and its effect on quality characteristics. *J Food Eng* 149:171–180. doi:10.1016/j.jfoodeng.2014.10.011
- Parras P, Martínez-Tomé M, Jiménez AM, Murcia MA (2007) Antioxidant capacity of coffees of several origins brewed following three different procedures. *Food Chem* 102:582–592
- Pizarro C, Esteban-Díez I, González-Sáiz JM (2007) Mixture resolution according to the percentage of robusta variety in order to detect adulteration in roasted coffee by near infrared spectroscopy. *Anal Chim Acta* 585:266–276. doi:10.1016/j.aca.2006.12.057
- Ponte S (2002) The latte revolution? Regulation, markets and consumption in the global coffee chain. *World Dev* 30:1099–1122
- Rakesh Jaiswal MFM (2014) Does roasted coffee contain chlorogenic acid lactones or/and cinnamoylshikimate esters? *Food Res Int* 61: 214–227. doi:10.1016/j.foodres.2013.09.040
- Rogers WJ, Michaux S, Bastin M, Bucheli P (1999) Changes to the content of sugars, sugar alcohols, myo-inositol, carboxylic acids and inorganic anions in developing grains from different varieties of Robusta (*Coffea canephora*) and Arabica (*C. arabica*) coffees. *Plant Sci* 149:115–123. doi:10.1016/S0168-9452(99)00147-8
- Romano R, Santini A, Le Grottaglie L et al (2014) Identification markers based on fatty acid composition to differentiate between roasted Arabica and Canephora (Robusta) coffee varieties in mixtures. *J Food Compos Anal* 35:1–9. doi:10.1016/j.jfca.2014.04.001
- Sturm M, Bertsch A, Gröpl C et al (2008) OpenMS—an open-source software framework for mass spectrometry. *BMC Bioinformatics* 9:163. doi:10.1186/1471-2105-9-163
- Tarzia A, Dos Santos Scholz MB, De Oliveira Petkowicz CL (2010) Influence of the postharvest processing method on polysaccharides and coffee beverages. *Int J Food Sci Technol* 45:2167–2175. doi:10.1111/j.1365-2621.2010.02388.x
- Wei F, Furihata K, Koda M et al (2012) ¹³C NMR-based metabolomics for the classification of green coffee beans according to variety and origin. *J Agric Food Chem* 60:10118–10125. doi:10.1021/jf3033057
- Weschenfelder TA, Lantin P, Viegas MC et al (2015) Concentration of aroma compounds from an industrial solution of soluble coffee by pervaporation process. *J Food Eng* 159:57–65. doi:10.1016/j.jfoodeng.2015.03.018
- Williams G (2011) Data mining with rattle and R. Springer New York, New York, NY
- Winkler R (2015a) SpiderMass: semantic database creation and tripartite metabolite identification strategy. *J Mass Spectrom* 50:538–541. doi:10.1002/jms.3559
- Winkler R (2015b) An evolving computational platform for biological mass spectrometry: workflows, statistics and data mining with MASSyPup64. *Peer J* 3:e1401. doi:10.7717/peerj.1401



Cinvestav

CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL IPN

SECRETARIA ACADÉMICA CINVESTAV IRAPUATO

KM 9.6 LIBRAMIENTO NORTE CARRETERA IRAPUATO-LEÓN APARTADO POSTAL 629, 36821 IRAPUATO, GTO.

TEL: (462) 6239606, FAX: (462) 6245846 secretaria@ira.cinvestav.mx

08 diciembre, 2016

DR. RUAIRIDH SAWERS
Coordinador Académico
Cinvestav Irapuato

De nuestra mayor consideración:

Por este conducto, hacemos de su conocimiento que hemos leído la tesis titulada: "Estudio integral metabólico en plantas de *Coffea canephora* para el mejoramiento sensorial del café", que presentará el **M.C. ROBERTO GAMBOA BECERRA**, estudiante del Programa de Doctorado en Ciencias del Departamento de Biotecnología y Bioquímica de esta Unidad, el día **martes 24 de enero del año 2017**

Como integrantes de su jurado en el examen de grado, manifestamos que se han llevado a cabo los cambios en las sugerencias y correcciones hechas al trabajo de tesis, y que autorizamos que proceda a realizar los trámites administrativos para dicho examen.

Sin más por el momento, agradecemos su amable atención y aprovechamos para enviarle un cordial saludo.

Atentamente
Los integrantes del Jurado

DR. ROBERT WINKLER

DR. JORGE MOLINA TORRES

DRA. NAYELLI MARSCH MARTINEZ DR. AGUSTINO MARTINEZ ANTONIO

DR. CHARLES STEWART GILLMOR III

DR. OSCAR GONZALEZ RIOS